2022 IEEE 61st Conference on Decision and Control (CDC)
December 6-9, 2022. Cancún, Mexico

Convergence Bounds of Decentralized Fictitious Play Around a Single Nash Equilibrium in Near-Potential Games

Sarper Aydın, Sina Arefizadeh and Ceyhun Eksin

2022 IEEE 61st Conference on Decision and Control (CDC) | 978-1-6654-6761-2/22/$31.00 ©2022 IEEE | DOI: 10.1109/CDC51059.2022.9993189

Abstract— **We analyze convergence of decentralized fictitious play (DFP) in near-potential games, where agents participate in a game in which the change in utility functions are closely aligned with a potential function. In DFP, agents take actions that maximize their expected utilities computed based on local estimates of empirical frequencies of other agents. These local estimates are updated by averaging estimates received from neighbors in a time-varying communication network. Assuming that a near-potential game has finitely many Nash equilibria that are distant enough from each other, we show that the empirical frequencies converge near a single Nash Equilibrium. This result establishes that DFP maintains the properties of standard fictitious play (FP) in near-potential games.**

I. Introduction

In non-cooperative games, agents select among available actions to maximize their individual utilities given the actions of other agents. Potential games [1] assume the existence of a potential function that captures the changes in individuals' utilities due to unilateral changes in actions. Various autonomous systems including transportation [2], robotic [3], communication [4] and energy systems [5] can be modeled with potential games. In potential games, convergence to a Nash equilibrium (NE) is shown via different dynamics such as best-response [1], fictitious play (FP) [3], and log-linear learning [6]. The fact that NE are equal to the optimal set of potential functions provides the justification to use NE to optimize system-level performance. However, assuming the existence of exact or ordinal potential games may be unrealistic in such systems in which there often exist unknown payoff relevant parameters.

Near-potential games allow a deviation between individual utilities and a potential function in terms of unilateral changes. This deviation can stem from estimation error about the environment, quantization, or the agents' abilities to execute actions with precision. In more detail, unilateral change is defined as the change in joint action profile where only one agent changes its action and others continue to select the same actions. Given the bounded deviation, canonical decision-making protocols, e.g., best-response, fictitious play (FP) and log-linear learning, converge to an approximate-Nash Equilibrium (NE) [7].

FP is an iterative multi-agent decision-making mechanism based on the premise that agents select actions that maximize their expected utilities assuming other agents follow stationary strategies. The stationary strategy of an agent is given by the empirical frequency of its past actions. Computation of this stationary strategy requires that agents have perfect information about the entire history of the plays. For the case of autonomous systems in real-life applications, this may not be possible. More realistically, each agent may only have limited access to past history of play via a communication network. In such a scenario, agents need to form estimates about the histogram of past actions of other agents using their local information. Here, we consider one such iterative learning mechanism, which can be interpreted as a decentralized version of FP (DFP) for near-potential games. FP is also a commonly employed iterative learning mechanism in other games such as stochastic zero-sum [8], [9] and mean-field games [10] in addition to near-potential and potential games.

Specifically, we consider agents that employ DFP in time-varying communication networks as in [11]. In [11], we showed that DFP converges to a set of strategies having better potential function values compared to the minimum potential value of approximate Nash equilibria. In [12], we characterized that empirical frequencies of agents converge around a single NE under two additional assumptions: *i)* finite number of Nash equilibria and *ii)* near-potential game is close enough to a potential game. Here, we further extend the results by deriving an upper bound on the distance between empirical frequencies and a single NE (Theorem 1). This result shows that DFP maintains the convergence properties of standard FP established in [7] for near-potential games. Considering a target assignment game with unknown target locations, we show that the action profiles can actually converge to the exact NE of the potential game when the uncertainty around the target locations is small. These results demonstrate that DFP can be used to model and design system behavior in large-scale autonomous teams in which agents have imperfect information about the objective and the actions of other agents.

II. Near-Potential Games

A game is defined by the tuple $\Gamma := (\mathcal{N}, \mathcal{A}^N, \{u_i\}_{i\in\mathcal{N}})$, where $\mathcal{N} = \{1, \cdots, N\}$ is the set of agents (players), $\mathcal{A}$ is a common action set, and $u_i : \mathcal{A}^N \to \mathbb{R}$ is the utility function of agent i that maps a joint action profile $(a_i, a_{-i}) \in \mathcal{A}^N$ to a real value. We denote the set of agents excluding i by $-i := \{j \in \mathcal{N} \setminus \{i\}\}$.

In a potential game $\hat{\Gamma}$, there exists a potential function u that captures the changes in individual utility functions given unilateral deviations [1].

Definition 1 (Potential Games) *A game $\hat{\Gamma}$ is a potential game, if there exists a potential function $u : \mathcal{A}^N \to \mathbb{R}$ such*

S. Aydin and C. Eksin are with the Industrial and Systems Engineering Department, Texas A&M University, College Station, TX 77843. E-mail: sarper.aydin@tamu.edu; eksinc@tamu.edu. This work was supported by NSF CCF-2008855.

that the following relation holds for all agents $i \in \mathcal{N}$,

$$u(a_i', a_{-i}) - u(a_i, a_{-i}) = u_i(a_i', a_{-i}) - u_i(a_i, a_{-i}) \quad (1)$$

where $a_i' \in \mathcal{A}$ and $a_i \in \mathcal{A}$ and $a_{-i} \in \mathcal{A}_{-i}$.

We will use the maximum pairwise difference (MPD) as the distance metric in defining the class of near-potential games [7].

Definition 2 (Maximum Pairwise Difference) *Let $\Gamma = (\mathcal{N}, \mathcal{A}^N, \{u_i\}_{i\in\mathcal{N}})$ and $\hat{\Gamma} = (\mathcal{N}, \mathcal{A}^N, \{\hat{u}_i\}_{i\in\mathcal{N}})$ be two games with the same set of agents and the joint action sets, and utilities respectively given as $\{u_i\}_{i\in\mathcal{N}}$ and $\{\hat{u}_i\}_{i\in\mathcal{N}}$. Further, let $d^{\Gamma}_{(a_i',a)} := u_i(a_i', a_{-i}) - u_i(a_i, a_{-i})$ be the difference in utility of agent i if only agent i changes its action to $a_i' \in \mathcal{A}_i$ from the joint action profile $a = (a_i, a_{-i}) \in \mathcal{A}^N$ in the game Γ. Then, MPD between the games Γ and $\hat{\Gamma}$ is defined as,*

$$d(\Gamma, \hat{\Gamma}) := \max_{i\in\mathcal{N},\, a_i'\in\mathcal{A},\, a\in\mathcal{A}^N} |d^{\Gamma}_{(a_i',a)} - d^{\hat{\Gamma}}_{(a_i',a)}|. \quad (2)$$

The maximum pairwise difference (MPD) defines the distance between two games based on the effects of unilateral deviations on individual utility functions. We can now formally state near-potential games as in [7].

Definition 3 (Near-Potential Games) *A game Γ is a near-potential game if there exists a potential game $\hat{\Gamma}$ within a maximum-pairwise distance (MPD), $d(\Gamma, \hat{\Gamma}) \leq \delta$ where $\delta \geq 0$.*

Near potential games relax the definition of potential games (Definition 1), similar to weakly acyclic [13], ordinal [1], weighted [1], and best-response potential games [14]. Determining the potential function for a near-potential games requires solving an optimization problem that finds the function minimizing (2). This optimization problem is shown to be convex in [15]. In this paper, we focus on the convergence properties of DFP in near-potential games.

III. Decentralized Fictitious Play

Fictitious play is a game-theoretical learning algorithm where agents assume that each agent selects its actions according to a stationary distribution (strategy) $\sigma_i \in \Delta\mathcal{A}$, where $\Delta\mathcal{A}$ denotes the set of probability distributions over the action space $\mathcal{A}$. Individual actions $a_i \in \mathcal{A}$ are represented with unit vectors $\mathbf{e}_k \in \{0,1\}^K$ where $|\mathcal{A}| = K$. In defining each action as an unit vector, we can consider each action as a degenerate distribution on $\Delta\mathcal{A}$. Then, the expected utility of agent i can be defined as follows,

$$u_i(\sigma_i, \sigma_{-i}) = \sum_{a\in\mathcal{A}^N} u_i(a_i, a_{-i})\sigma(a), \quad (3)$$

where $\sigma(a) = [\sigma_1(a_1), \ldots, \sigma_N(a_N)]$ is the joint strategy profile, and σ_{-i} is strategy profile of all agents excluding agent i.

The empirical frequency of past $t \in \mathbb{N}$ actions of agent i $(a_{i,1}, a_{i,2}, \ldots, a_{i,t})$, is denoted with $f_{i,t} \in \Delta\mathcal{A}$. We can write $f_{i,t}$ using the following recursion,

$$f_{i,t} = \frac{t-1}{t} f_{i,t-1} + \frac{1}{t} a_{i,t}, \quad (4)$$

where we note that $a_{i,t}$ is a degenerate distribution in $\Delta\mathcal{A}$.

In networked interactions, agents may not have perfect information on other agents' empirical frequencies $(f_{j,t})$. In such a setting, agents may form beliefs about the empirical distribution of other agents' actions. Specifically, agents communicate over a time-varying network $\mathcal{G}_t = (\mathcal{N}, \mathcal{E}_t)$, where neighbors of agent i, denoted with $\mathcal{N}_{i,t} := \{j : (i,j) \in \mathcal{E}_t\}$, may change over each time step t. In this setting, agent i keeps a local copy (belief) $\upsilon^i_{j,t} \in \Delta\mathcal{A}_j$ of true empirical frequencies of agent j $(f_{j,t} \in \Delta\mathcal{A})$. Local beliefs $\upsilon^i_{j,t}$ are updated by weighted averaging of local copies received from its neighbors,

$$\upsilon^i_{j,t} = \sum_{l\in\mathcal{N}_{i,t}\cup\{i\}} w^i_{jl,t} \upsilon^l_{j,t}, \quad (5)$$

where $w^i_{jl,t} \geq 0$ is the weight of agent l's estimate of agent j to update the local belief of agent i.

In DFP, agent i assumes other agents' strategies follow stationary distributions given by its local beliefs $\upsilon^i_{-i,t-1} := \{\upsilon^i_{j,t-1}\}_{j\in\mathcal{N}\setminus i}$, and takes action $a_{i,t}$ to maximize its utility,

$$a_{i,t} \in \arg\max_{a_i\in\mathcal{A}_i} u_i(a_i, \upsilon^i_{-i,t-1}). \quad (6)$$

The decision-making and information exchange steps of agent i are summarized in Algorithm 1.

Algorithm 1 DFP for Agent i

Input: Local estimates υ^i_{-i0} and time-varying networks $\{\mathcal{G}_t = (\mathcal{N}, \mathcal{E}_t)\}_{t\geq 1}$.
for $t = 1, 2, \cdots$ **do**
  Best respond with a_{it} (6) and update $f_{i,t}$ (4).
  Communicate with $\mathcal{N}_{i,t}$ and update beliefs $\upsilon^i_{j,t}$ (5).
end for

IV. Convergence of DFP in Near-Potential Games

A. Preliminaries

The joint strategy profile σ^* is an approximate NE, if no agent can benefit more than $\epsilon \geq 0$ by switching to another strategy.

Definition 4 (Approximate Nash Equilibrium) *The joint strategy profile $\sigma^* = (\sigma_i^*, \sigma_{-i}^*) \in \Delta\mathcal{A}^N$ is an ϵ-Nash equilibrium of the game Γ for $\epsilon \geq 0$ if and only if for all $i \in \mathcal{N}$,*

$$u_i(\sigma_i^*, \sigma_{-i}^*) - u_i(\sigma_i, \sigma_{-i}^*) \geq -\epsilon, \quad \textit{for all } \sigma_i \in \Delta\mathcal{A}_i. \quad (7)$$

We use the notation Σ_ϵ to denote the set of ϵ-Nash equilibria in game a Γ.

B. Convergence Analysis

We state the assumption on the structure of the time-varying communication network $\{\mathcal{G}_t\}_{t\geq 1}$ in the following.

Assumption 1 *Time-varying communication networks $\{\mathcal{G}_t\}_{t\geq 1}$ satisfy the properties below,*

i) The network $\mathcal{G} = (\mathcal{N}, \mathcal{E}_\infty)$ is connected, where $\mathcal{E}_\infty = \{(i,j) | (i,j) \in \mathcal{E}_t, \text{for infinitely many } t \in \mathbb{N}\}$.

ii) There exists a time step $T_B > 0$, such that for any edge $(i,j) \in \mathcal{E}_\infty$ and $t \geq 1$, it holds $(i,j) \in \bigcup_{\tau=0}^{T_B-1} \mathcal{E}_{t+\tau}$.

The properties $i)-ii)$ are named as *connectivity* and *bounded communication interval* such that any information between agents i and j is sent in a bounded time interval. These assumptions are standard in distributed optimization algorithms for time-varying networks [16].

Assumption 2 *There exists a scalar $0 < \eta < 1$, such that the followings hold for all $i \in \mathcal{N}$, $j \in \mathcal{N}$ and $t = 1, 2, \ldots$,*

(i) If $l \in \mathcal{N}_{i,t} \cup \{i\}$, then $w^i_{jl,t} \geq \eta$. Otherwise, $w^i_{jl,t} = 0$,

(ii) $w^i_{ii,t} = 1$,

(iii) $\sum_{l \in \mathcal{N}_{i,t} \cup \{i\}} w^i_{jl,t} = 1$.

According to Assumption 2*(i)*, agents put positive weights on their current neighbors' beliefs (5). Assumption 2*(ii)* ensures that agents beliefs about their empirical frequencies is correct, i.e., so that $\nu^i_{i,t} = f_{i,t}$ for all $t > 0$. Assumption 2*(iii)* ensures that the weights sum to one. The following lemma characterizes the convergence rate of local estimates $\nu^i_{j,t}$ to empirical frequencies $f_{j,t}$—see [17] for proof.

Lemma 1 (Proposition 1, [17]) *Suppose Assumptions 1-2 hold. If $f_{j0} = \upsilon^i_{j,0}$ holds for all pairs of agents $j \in \mathcal{N}$ and $i \in \mathcal{N}$, then the local copies $\{\upsilon^i_t\}^{i\in\mathcal{N}}_{t\geq 0}$ converge to the empirical frequencies $\{f_t\}_{t\geq 0}$ with rate $O(\log t/t)$, i.e., $||\upsilon^i_{j,t} - f_{j,t}|| = O(\log t/t)$ for all $j \in \mathcal{N}$ and $i \in \mathcal{N}$.*

The proof mainly relies on the properties of row-stochastic weight matrices given by Assumption 2. Next result provides a lower bound on the difference in potential value of empirical frequencies between consecutive time steps—see [11] for the proof.

Lemma 2 (Lemma 2, [11]) *Suppose Assumptions 1-2 hold. Let Γ be a δ near-potential game for some $\delta \geq 0$. The potential function is given by $u(\cdot)$. We denote the empirical frequency sequence generated by the DFP algorithm as $\{f_t\}_{t\geq 1}$. If the empirical frequency f_t is outside the ϵ-NE set for $\epsilon \geq 0$, then given a long enough $T > 0$ we have*

$$u(f_{t+1}) - u(f_t) \geq \frac{\epsilon - N\delta}{t+1} - O\Big(\frac{\log t}{t^2}\Big) \text{ for all } t \geq T. \quad (8)$$

This results assures that when the empirical frequencies are outside the approximate NE region $N\delta$, the potential value of a close potential game increases as the second term $O\Big(\frac{\log t}{t^2}\Big)$ goes to 0. In the following, we analyze the potential change if empirical frequencies go through a path where they start outside of approximate-NE regions $N\delta + \epsilon_1$ and $N\delta + \epsilon_2$ in order, and then go back firstly to $N\delta + \epsilon_2$ and then to $N\delta + \epsilon_1$ given $0 < \epsilon_1 < \epsilon_2$.

Lemma 3 (Lemma 3, [12]) *Suppose Assumptions 1- 2 hold Let $\{f_t\}_{t\geq 1}$ be the sequence generated by Algorithm 1. Further, let T_1, T_2, T'_2, T'_1 be time steps such that $T < T_1 \leq T_2 < T'_2 \leq T'_1$, for large enough $T > 0$ defined as follows,*

- *T_1 is a time step when the empirical frequencies leave from $(N\delta + \epsilon_1)$-NE region, i.e. $f_{T_1 - 1} \in \Sigma_{N\delta+\epsilon_1}$ and $f_t \notin \Sigma_{N\delta+\epsilon_1}$, for all $T_1 \leq t < T'_1$,*
- *T_2 is a time step when the empirical frequencies leave from $(N\delta + \epsilon_2)$-NE region, i.e. $f_{T_2 - 1} \in \Sigma_{N\delta+\epsilon_2}$ and $f_t \notin \Sigma_{N\delta+\epsilon_2}$, for all $T_2 \leq t < T'_2$,*
- *T'_2 is a time step when the empirical frequencies enters again into $(N\delta + \epsilon_2)$-NE region, i.e. $f_{T'_2 - 1} \notin \Sigma_{N\delta+\epsilon_2}$ and $f_{T'_2} \in \Sigma_{N\delta+\epsilon_2}$,*
- *T'_1 is a time step when the empirical frequencies enters again into $(N\delta + \epsilon_1)$-NE region, i.e. $f_{T'_1 - 1} \notin \Sigma_{N\delta+\epsilon_1}$ and $f_{T'_1} \in \Sigma_{N\delta+\epsilon_1}$,*

where $\epsilon_1 > 0$, $\epsilon_2 > 0$. Then, there exist $0 < \epsilon_1 < \epsilon_2$ such that the following holds,

$$u(f_{T'_1}) - u(f_{T_1}) \geq \sum_{t=T_2}^{T'_2 - 1} \frac{2\epsilon_2}{3(t+1)}. \quad (9)$$

The proof follows from the statement of Lemma 2 with the fact that the potential increases when empirical frequencies are outside the given approximate-NE regions. The result in (9) is a lower bound on the excursion away from an approximate NE. We make two additional assumptions identical to those made in [7].

Assumption 3 *The game $\Gamma := (\mathcal{N}, \mathcal{A}^N, \{u_i\}_{i\in\mathcal{N}})$ has only a nonempty set of finitely many Nash equilibria, $\Sigma_0 = \{\sigma^{*(1)}, \sigma^{*(2)}, \cdots, \sigma^{*(M)}\}$ where $|\Sigma_0| = M$ and $M \in \mathbb{Z}^+$.*

Assumptions 3 states that the game Γ can only have a finite number of Nash equilibria.

Assumption 4 *Let $q : \mathbb{R}_+ \to \mathbb{R}_+$ be a function defined as follows,*

$$q(\alpha) = \max_{\sigma \in \Sigma_\alpha} \min_{m \in \{1,\cdots,M\}} ||\sigma - \sigma^{*(m)}||. \quad (10)$$

where $\sigma^{*(m)}$ is a NE of the game Γ as defined in Assumption 3. The MPD between two games $d(\Gamma, \hat{\Gamma}) \leq \delta < \bar{\delta}$ is small enough such that there exists $\bar{\alpha} > 0$ that satisfies $N\delta < N\bar{\delta} < \bar{\alpha}/2$ and $q(\bar{\alpha}) < d^*/4$, where d^* is the minimum distance between any two equilibria the game Γ, i.e., $d^* = \min_{m' \neq m''} ||\sigma^{*(m')} - \sigma^{*(m'')}||$, where $m', m'' \in \{1, \cdots, M\}$.

Assumption 4 requires that the minimum distance between any two different Nash equilibria of the near potential game has to be large enough with respect to the MPD between the near potential game and a given potential game. In particular, the function $q(\bar{\alpha})$ provides the largest distance within $\bar{\alpha}$-Nash equilibrium strategies to a Nash equilibrium. The assumption requires the existence of a constant $\bar{\alpha} > 0$ such that MPD distance between the two games is smaller than $\bar{\alpha}/2N$, and the largest distance between approximate Nash and Nash equilibrium strategies is bounded by a quarter of the minimum distance between any two equilibrium strategies. This assumption is critical for our main result provided in Theorem 1. Next result states that agents stay in an approximate NE region of a single NE after long enough time steps.

Lemma 4 (Theorem 1, [12]) *Suppose Assumptions 1- 4 hold. Let* $\{f_t\}_{t\geq 1}$ *be the sequence generated by Algorithm 1. The empirical frequencies* $\{f_t\}_{t\geq 1}$ *converge to an approximate equilibrium set around a single equilibrium point, after long enough time* $t > T$.

The proof, provided in [12], is based on the idea that the potential values around different NE points are strictly ordered. Note that if the empirical frequencies moves away from an approximate NE region, then agents increase their utilities according to Lemma 2. These two observations are used to show that the empirical frequencies stay around a single NE–see [12] for proof. Now, we are ready to state our main result.

Theorem 1 *Suppose Assumptions 1- 4 hold. Let* $\{f_t\}_{t\geq 1}$ *be the sequence generated by Algorithm 1. Then, there exist* $\bar{\delta} > \delta > 0$ *and* $\bar{\epsilon} > 0$, *such that the empirical frequencies* f_t *converge to the set around a single equilibrium point* $\sigma^{*(m)}$ *for some* $m \in \{1, \ldots, M\}$,

$$\mathcal{S}^* = \left\{ \sigma \in \Delta\mathcal{A}^N \middle| ||\sigma - \sigma^{*(m)}|| \leq \frac{4q(N\delta)NL}{\epsilon} + q(N\delta + \epsilon) \right\} \quad (11)$$

where $q : \mathbb{R}_+ \to \mathbb{R}_+$ *is the function defined as in* (10), $\bar{\epsilon}$ *is defined in Lemma 7 such that it satisfies* (24) *and* (25), *and* ϵ *is a constant where* $\bar{\epsilon} \geq \epsilon > 0$.

Proof: Observe that the following relation is satisfied,

$$||f_{t+1} - f_t|| = \frac{1}{t+1}||f_t - a_t|| \leq \frac{1}{t+1}(||f_t|| + ||a_t||) \leq \frac{2N}{t+1}, \quad (12)$$

since it holds $||f_{i,t}|| \leq 1$ and $||a_{i,t}|| \leq 1$ for any $i \in \mathcal{N}$ and $t \in \mathbb{N}$.

Let T_1, T_2, T_2', T_1' be time steps defined as previously. The lower and upper bounds are going to be derived for the difference $u(f_{T_1'}) - u(f_{T_1})$, in which f_t traverses around a single equilibrium point $\sigma^{*(m)}$. By letting $T_2 \geq T_1 > T + 1$, $\bar{d}_f$ is defined as the maximum distance between f_t during time interval $\mathcal{T} = \{T_2, T_2 + 1, \cdots, T_2' - 1\}$ and the approximate equilibrium set $\Sigma_{N\delta+\epsilon_2}$,

$$\bar{d}_f = \max_{t\in\mathcal{T}} \min_{\sigma\in\Sigma_{N\delta+\epsilon_2}} ||f_t - \sigma||. \quad (13)$$

Noting $(f_{T_2-1}, f_{T_2'}) \in \Sigma_{N\delta+\epsilon_2} \times \Sigma_{N\delta+\epsilon_2}$, the following bound holds for the sum of distances traversed between each time step for the interval $\{T_2 - 1\} \cup \mathcal{T}$,

$$2\bar{d}_f \leq \sum_{t=T_2-1}^{T_2'-1} ||f_{t+1} - f_t||. \quad (14)$$

Using (12), the inequality (14) can be rewritten as,

$$2\bar{d}_f \leq \sum_{t=T_2-1}^{T_2'-1} \frac{2N}{t+1} = \sum_{t=T_2}^{T_2'-1} \frac{2N}{t+1} + \frac{2N}{T_2}. \quad (15)$$

Hence, we can obtain a lower bound for $u(f_{T_1'}) - u(f_{T_1})$ with (15) and Lemma 3,

$$u(f_{T_1'}) - u(f_{T_1}) \geq \sum_{t=T_2}^{T_2'-1} \frac{2\epsilon_2}{3(t+1)} \geq \left(\bar{d}_f - \frac{N}{T_2}\right)\frac{2\epsilon_2}{3N}. \quad (16)$$

Similarly, we can also derive an upper bound again for $u(f_{T_1'}) - u(f_{T_1})$. We achieve the upper bound, by the fact $(f_{T_2-1}, f_{T_2'}) \in \Sigma_{N\delta+\epsilon_1} \times \Sigma_{N\delta+\epsilon_1}$, it holds, $||f_{T_1'} - \sigma^{*(m)}|| \leq q(N\delta + \epsilon_1)$ and $||f_{T_1-1} - \sigma^{*(m)}|| \leq q(N\delta + \epsilon_1)$. By Lipschitz continuity, it holds $u(f_{T_1'}) - u(f_{T_1} - 1) \leq 2q(N\delta + \epsilon_1)L$. Again with Lipschitz continuity, (12) assures $u(f_{T_1}) - u(f_{T_1} - 1) \leq \frac{2NL}{T_1}$. Using both of these bounds, the following holds,

$$u(f_{T_1'}) - u(f_{T_1}) \leq 2q(N\delta + \epsilon_1)L + \frac{2NL}{T_1}. \quad (17)$$

Then, using the lower and upper bounds, we have

$$\left(\bar{d}_f - \frac{N}{T_2}\right)\frac{2\epsilon_2}{3N} \leq 2q(N\delta + \epsilon_1)L + \frac{2NL}{T_1}. \quad (18)$$

After sufficiently long enough time $T_2 > T_1 > T + 1 > T$, (18) implies that

$$\bar{d}_f \leq \frac{3q(N\delta + \epsilon_1)NL}{\epsilon_2} + \frac{3N^2L}{\epsilon_2 T_1} + \frac{N}{T_2} \leq \frac{4q(N\delta + \epsilon_1)NL}{\epsilon_2}. \quad (19)$$

As $\epsilon_1 > 0$ is an arbitrary positive number such that $0 < \epsilon_1 < \epsilon_2$, it consequently holds by upper semi-continuity (Lemma 6),

$$\bar{d}_f \leq \limsup_{\epsilon_1\to 0} \frac{4q(N\delta + \epsilon_1)NL}{\epsilon_2} \leq \frac{4q(N\delta)NL}{\epsilon_2}. \quad (20)$$

Hence, since $\bar{d}_f$ is the maximum distance between f_t and the set $\Sigma_{N\delta+\epsilon_2}$, there exists a $\tilde{\sigma} \in \Sigma_{N\delta+\epsilon_2} \subseteq \Delta\mathcal{A}^N$ such that the upper bound on the given distance below holds,

$$||f_t - \tilde{\sigma}|| \leq \frac{4q(N\delta)NL}{\epsilon_2}. \quad (21)$$

Thus, given $\tilde{\sigma} \in \Sigma_{N\delta+\epsilon_2}$ and $||\tilde{\sigma} - \sigma^{*(m)}|| \leq q(N\delta + \epsilon_1)$, using triangle inequality together with (21) proves the theorem's statement that f_t converges to the set $\mathcal{S}^*$ in (11) with any arbitrary number ϵ_2 satisfying the condition $0 < \epsilon_2 \leq \bar{\epsilon}$ around a single NE $\sigma^{*(m)}$. ■

Theorem 1 characterizes the region around a NE strategy that the empirical frequencies will stay in. The region grows with MPD between the two games and as the distance between $\bar{\alpha} - N\delta$ shrinks. Note that $\bar{\alpha}$ comes from Assumption 4 making sure that the MPD between the two games is relatively small compared to the distance between the Nash equilibrium strategies. If the game Γ is a potential game, i.e., $\delta = 0$, Theorem 1 shows the convergence to a NE by fact that $q(N\delta) = 0$ and, $\lim_{\epsilon\to 0} q(\epsilon) = 0$ with arbitrarily small chosen ϵ. This result confirms that DFP obtains identical bounds for convergence in near-potential games as the ones obtained by the standard FP stated in [7].

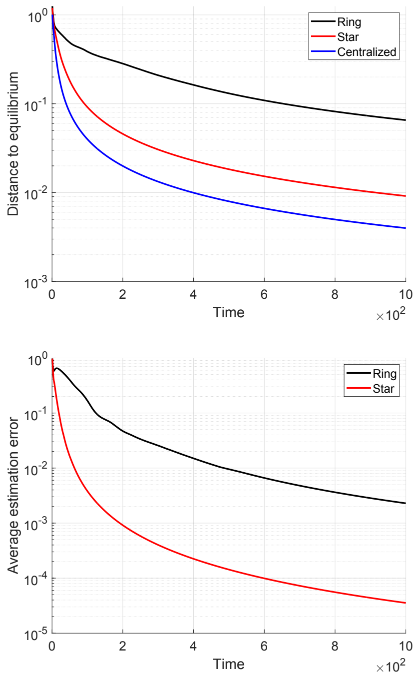

Fig. 1. FP and DFP in target assignment game with unknown target locations over 20 runs. (Top) Average distance to Nash equilibrium $\frac{1}{N}\sum_{i\in\mathcal{N}}||f_{it} - \sigma_i^*||$ (Bottom) Average estimation error $\frac{1}{N(N-1)}\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{N}\setminus\{i\}}||f_{it} - v_{it}^j||$.

V. Numerical Experiments

We consider a target assignment game with $N = 20$ autonomous agents and $K = 20$ targets. Agents select only one target $k \in \mathcal{K} := \{1, \cdots, K\}$ to maximize their utility functions defined as below,

$$u_i(a_i, a_{-i}) = \frac{a_i^T \mathbb{1}_{a_{-ik}=0}}{a_i^T d_i}, \tag{22}$$

in which $a_i = \mathbf{e}_k \in \mathbb{R}^K$ is an unit vector and $\mathbb{1}_{a_{-ik}=0} \in \{0,1\}^K$ is a binary vector whose k^{th} index is 1 if none of the other agents $j \in \mathcal{N} \setminus \{i\}$ select k, and otherwise the k^{th} index is equal to 0. The distance vector $d_i = [d_{i1}, \cdots, d_{ik}, \cdots, d_{iK}] \in \mathbb{R}_+^K$ gives the distance between agent i and targets. The denominator term in (22) is equal to the distance of the agent to its selected target. As per the utility function in (22), agent i obtains a positive utility value by selecting target k, when this target k is not selected by any of the other agents. The exact utility value that agent i receives is inversely proportional to the distance of the agent between the target selected and the position of agent i. This means that no agent can improve their utility by selecting another target, if all other targets are selected by other agents. Hence, this ensures that any joint action profile creating one-to-one assignment between agents and targets is a NE of the game.

The game with utility functions given in (22) is a potential game when the distance to each target for each agent is identical. Agents do not know target locations a priori. Agents receive individual signals ϑ_{kt}^i at each time step t about the position of each target k. The private signal $\vartheta_t^i = [\vartheta_1^i, \cdots, \vartheta_K^i]^T$ of each agent i follows a multivariate normal

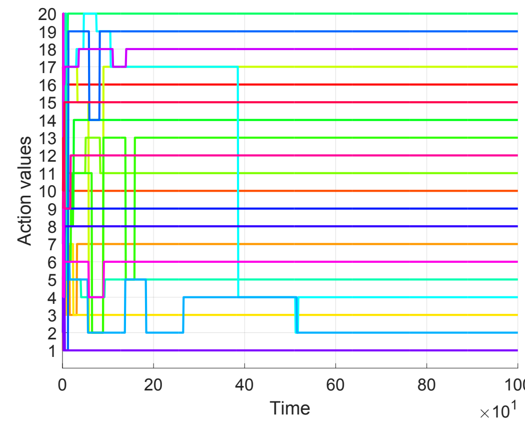

Fig. 2. DFP with ring network in target assignment game. Joint action profile over time from a single run. Actions converge to an one-to-one assignment between targets and agents.

distribution with mean values equal to the actual distance to each target $\theta = [\theta_1, \cdots, \theta_K]^T$ and a covariance matrix σI, where $\sigma = 0.5$ and $I \in \mathbb{R}^{K\times K}$ is the identity matrix. At time τ, the position of targets is estimated as $\hat{\theta}_k = (1/\tau)\sum_{t=1}^{\tau}\vartheta_t^i$ by each agent i. We assume agents only receive signals in the first $\tau = 5$ steps. Early stopping of signals and different distances to targets implies that the game agents take actions in is a near-potential game.

We consider ring, star, and fully-connected communication networks. Note that a fully-connected communication network means that the DFP corresponds to the centralized (standard) FP, in which agents have access to all available information immediately. In this case, agents can keep track of the empirical frequencies of all agents, i.e., there is no need for averaging local estimates. In the ring and star networks, we set self-weights are $w_{ji,t}^i = 0.75$ for the estimates of others, as weights of neighbor agents are $w_{jl,t}^i = 0.25/|\mathcal{N}_i|$, for all $j \in \mathcal{N}_i$ and for all $t \in \mathbb{N}$.

We consider 20 runs for each network and provide average estimation errors in Fig. 1 (Bottom). We observe that the average estimation error between local (estimated) and actual empirical frequencies aligns with the rate $O(\log t/t)$ given in Lemma 2. The estimation error shrinks faster in the star network compared to the ring network. For the centralized FP with perfect information, this error is always zero and, thus it does not appear in the figure.

Fig. 1 (Top) shows that empirical frequencies converge toward a NE in similar rates for all networks. In all cases, the action profile reaches a NE which is an one-to-one assignment of agents to targets by the final time $T_f = 1000$. Centralized FP is the fastest on average, while DFP on a ring network has the most distant results to a NE by the final time T_f. As shown by the action profile evolution on a single run instance given a ring network (Fig. 2), the joint action profile converges to a single NE (an one-to-one agent-target assignment). This observation also holds for the other runs and networks. Together with Fig. 1 (Top), this confirms that empirical frequencies $\{f_t\}$ converge to a region around a single NE and do not oscillate between different Nash equilibria after a long enough time.

VI. Conclusion

In this paper, we analyzed the convergence of DFP in near-potential games. We derived the upper bound on the distance between empirical frequencies and a NE of the closest potential game. This result shows that DFP preserves the convergence properties of the standard FP. It also implies that a team of networked agents can learn to behave approximately rational despite lasting disagreements about the team objective.

Appendix

Upper semi-continuous correspondences is an important notion for the results throughout the analysis. We are going to start with providing its definition.

Definition 5 (Upper Semi-Continuous Correspondence) *A correspondence $h : X \rightrightarrows Y$ is upper semi-continuous, if one of the following statements hold,*

- *For any $\bar{x} \in X$ and any open neighborhood V of $h(\bar{x})$, there exists a neighborhood U of $\bar{x}$, such that $h(x) \subset V$, and $h(x)$ is a compact set for all $x \in U$.*
- *Y is compact, and the set, i.e. its graph, $\{(x, y)|x \in X, y \in h(x)\}$ is closed.*

Lemma 5 (Lemma 4, [11]) *Let $h : \mathbb{R} \rightrightarrows \Delta\mathcal{A}^N$ be the correspondence representing the set of α-NE strategies, i.e.,*

$$h(\alpha) = \Sigma_\alpha = \{\sigma \in \Delta\mathcal{A}^N | \psi(\sigma) \geq -\alpha\} \tag{23}$$

where $\Psi : \Delta\mathcal{A}^N \rightarrow \mathbb{R}$ is defined as $\psi(\sigma) = -\max_{i\in\mathcal{N},\, a_i\in\mathcal{A}_i} (u_i(a_i, \sigma_{-i}) - u_i(\sigma_i, \sigma_{-i}))$. Then, the correspondence $h : \mathbb{R} \rightrightarrows \Delta\mathcal{A}^N$ is upper semi-continuous.

Lemma 6 *Suppose Assumption 3 holds. Let $q : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ be a function defined as in* (10)*. Then, the function $q : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is i) weakly increasing, ii) upper semi-continuous, iii) satisfies $q(0) = 0$ and $\lim_{\alpha\to 0} q(\alpha) = 0$.*

Proof: Before proving each statement, see that $\min_{m\in\{1,\cdots,M\}} ||\sigma - \sigma^{*(m)}||$ is a continuous function in σ, since it is the minimum of finitely many continuous functions. Further, by the definition of ϵ-NE set, for any value of α, the set Σ_α is compact. Thus, maximum over the compact set Σ_α exists, and the function q is well defined.

i) For any $(\alpha_1, \alpha_2) \in \mathbb{R}_+ \times \mathbb{R}_+$ such that $\alpha_1 \leq \alpha_2$, by definition again, it holds $\Sigma_{\alpha_1} \subseteq \Sigma_{\alpha_2}$. Then, the function also satisfies the condition $q(\alpha_1) \leq q(\alpha_2)$, for any $\alpha_1 \leq \alpha_2$. Hence, q is weakly increasing.

ii) Since again, $\min_{m\in\{1,\cdots,M\}} ||\sigma - \sigma^{*(m)}||$ is a continuous function in $\sigma \in \Delta\mathcal{A}^N$, and from Lemma 5 the correspondence $h(\alpha) = \Sigma_\alpha$ is upper semi-continuous. This gives that q is upper semi-continuous by Berge's maximum theorem.

iii) Since, Σ_0 is the set of NE, it holds $q(0) = 0$. By upper semi-continuity of q, for any $\epsilon > 0$, there exists a neighborhood V around 0, that holds $q(\alpha) \leq \epsilon$, for all $\alpha \in V$. Since it holds $q(\alpha) \geq 0$, for all α, the limit exists, and it holds $\lim_{\alpha\to 0} q(\alpha) = 0$.

■

Lemma 7 *Suppose Assumptions 3 and 4 hold. There exists $\bar{\epsilon} > 0$ and $C > 0$ such that the following inequalities are satisfied,*

$$N\delta + \bar{\epsilon} < \bar{\alpha}, \text{ and} \tag{24}$$

$$q(N\delta + \bar{\epsilon}) < \frac{(\bar{\alpha} - N\delta)d^*}{CNL}. \tag{25}$$

Proof: For sufficiently small enough $\delta > 0$ and $C > 0$ by Assumption 4, and $\bar{\epsilon} > 0$, it holds $N\delta + \epsilon < \alpha$ for any $\epsilon \leq \bar{\epsilon}$. Using the fact $\lim_{\alpha\to 0} q(\alpha) = 0$ from Lemma 6, it gives $q(N\bar{\delta} + \epsilon) < \frac{\bar{\alpha}d^*}{2CNL} < \frac{(\bar{\alpha} - N\bar{\delta})d^*}{CNL}$. ■

Note that $C < 24$ for Lemma 4 to hold. See [7], [12] for the details.

References

[1] D. Monderer and L. S. Shapley, "Potential games," *Games and economic behavior*, vol. 14, no. 1, pp. 124–143, 1996.

[2] 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, 2009.

[3] C. Eksin and A. Ribeiro, "Distributed fictitious play for multiagent systems in uncertain environments," *IEEE Transactions on Automatic Control*, vol. 63, no. 4, pp. 1177–1184, 2017.

[4] Q. D. Lã, Y. H. Chew, and B.-H. Soong, *Potential Game Theory: Applications in Radio Resource Allocation*. Springer, 2016.

[5] J. Zeng, Q. Wang, J. Liu, J. Chen, and H. Chen, "A potential game approach to distributed operational optimization for microgrid energy management with renewable energy and demand response," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 6, pp. 4479–4489, 2018.

[6] J. R. Marden and J. S. Shamma, "Revisiting log-linear learning: Asynchrony, completeness and payoff-based implementation," *Games and Economic Behavior*, vol. 75, no. 2, pp. 788–808, 2012.

[7] O. Candogan, A. Ozdaglar, and P. A. Parrilo, "Dynamics in near-potential games," *Games and Economic Behavior*, vol. 82, pp. 66–90, 2013.

[8] M. O. Sayin, F. Parise, and A. Ozdaglar, "Fictitious play in zero-sum stochastic games," 2021.

[9] D. S. Leslie, S. Perkins, and Z. Xu, "Best-response dynamics in zero-sum stochastic games," *Journal of Economic Theory*, vol. 189, p. 105095, 2020.

[10] S. Perrin, J. Perolat, M. Laurière, M. Geist, R. Elie, and O. Pietquin, "Fictitious play for mean field games: Continuous time analysis and applications," in *NeurIPS*, 2020. [Online]. Available: https://arxiv.org/abs/2007.03458

[11] S. Aydın, S. Arefizadeh, and C. Eksin, "Decentralized fictitious play in near-potential games with time-varying communication networks," *IEEE Control Systems Letters*, vol. 6, pp. 1226–1231, 2021.

[12] S. Aydin, S. Arefizadeh, and C. Eksin, "Decentralized fictitious play converges near a nash equilibrium in near-potential games," in *2021 55th Asilomar Conference on Signals, Systems, and Computers*. IEEE, 2022, pp. 998–1002.

[13] H. P. Young, "The evolution of conventions," *Econometrica: Journal of the Econometric Society*, pp. 57–84, 1993.

[14] M. Voorneveld, "Best-response potential games," *Economics letters*, vol. 66, no. 3, pp. 289–295, 2000.

[15] O. Candogan, I. Menache, A. Ozdaglar, and P. A. Parrilo, "Flows and decompositions of games: Harmonic and potential games," *Mathematics of Operations Research*, vol. 36, no. 3, pp. 474–503, 2011.

[16] A. Nedic and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," *IEEE Transactions on Automatic Control*, vol. 54, no. 1, pp. 48–61, 2009.

[17] S. Arefizadeh and C. Eksin, "Distributed fictitious play in potential games with time varying communication networks," in *2019 53rd Asilomar Conference on Signals, Systems, and Computers*. IEEE, 2019, pp. 1755–1759.