Decentralized Fictitious Play Converges Near a Nash Equilibrium in Near-Potential Games

Sarper Aydın, Sina Arefizadeh and Ceyhun Eksin

Abstract— **We investigate convergence of decentralized fictitious play (DFP) in near-potential games, wherein agents preferences can almost be captured by a potential function. In DFP agents keep local estimates of other agents' empirical frequencies, best-respond against these estimates, and receive information over a time-varying communication network. We prove that empirical frequencies of actions generated by DFP converge around a single Nash Equilibrium (NE) assuming that there are only finitely many Nash equilibria, and the difference in utility functions resulting from unilateral deviations is close enough to the difference in the potential function values. This result assures that DFP has the same convergence properties of standard Fictitious play (FP) in near-potential games.**

I. INTRODUCTION

Game theory deals with systems having multiple decision-makers. In non-cooperative games, agents take actions to maximize their individual utility functions that depend on the actions of other agents. Potential games is a special class of games that capture scenarios where there exists a common function modeling the change in individual utilities, named as potential function. Applications of potential games appear in various large-scale networked systems including transportation systems [1], mobile robotic systems [2], and communication networks [3]. Decentralized decision-making protocols, e.g., best-response [4], [5], fictitious play (FP) [6], [7], are used to understanding emerging behavior or to design individual actions in such large-scale systems. A common assumption in the convergence of these protocols is that agents have full or common information about their utility functions or the potential function. Here, we lift this assumption by allowing the game agents are playing to deviate from an exact potential game.

Near-potential games [8] extend potential games, by defining games as a deviation from a potential game. This deviation may stem from incomplete information about payoff-relevant environment parameters. Specifically, the deviation between two games is defined in terms of unilateral change of actions, where only one agent changes its and others stay in the same profile. If this deviation is bounded, traditional decision-making protocols, e.g., best-response, or FP, converge to a region around NE, i.e., an approximate-NE [8]. In this paper, we analyze convergence properties of a decentralized version of FP (DFP) where agents can only exchange information with a subset of their neighbors after

each decision epoch, which is in contrast to standard FP that assumes agents have perfect information about the past actions of other agents.

In particular, we consider agents taking actions with respect to DFP in time-varying communication networks as in [9]. In [9], we had shown that DFP converges to a set of strategies that obtain potential function values comparable to the set of approximate Nash equilibria. Here, we extend these results to show that empirical frequencies of agents converge around a single NE (Theorem 1) given two additional assumptions: *i)* number of Nash equilibria is finite, and *ii)* the near-potential game is close enough to a potential game. Numerical experiments on a target assignment game with unknown payoffs show that the action profiles can actually converge to the exact NE of the closest potential game (the game with known payoffs). Together these results show that DFP can be used to model or design team behavior in large-scale networked systems where agents communicate over a time-varying network, and have different information about a given common goal.

II. NEAR-POTENTIAL GAMES

In a game Γ, agents defined by the set $\mathcal{N} = \{1, \cdots, N\}$, select actions over a common finite set $a_i \in \mathcal{A}_i = \mathcal{A}$, where $i \in \mathcal{N}$. Agents have individual utility values $u_i : \mathcal{A}^N \to \mathbb{R}$ corresponding to joint action profiles $(a_i, a_{-i}) \in \mathcal{A}^N$ where $-i := \{j \in \mathcal{N} \setminus \{i\}\}$ denotes the set of all agents except i. The game is formally defined by the tuple $\Gamma := (\mathcal{N}, \mathcal{A}^N, \{u_i\}_{i \in \mathcal{N}})$. Potential games can be defined as follows [4].

Definition 1 (Potential Games) *A game Γ is a potential game, if there exists a potential function $u : \mathcal{A}^N \to \mathbb{R}$ such 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 $a_{-i} \in \mathcal{A}_{-i}$.

Potential function $u : \mathcal{A}^N \to \mathbb{R}$ mirrors the change in individual utilities $u_i : \mathcal{A}^N \to \mathbb{R}$ as a result of unilateral changes between two actions $a_i' \in \mathcal{A}$ and $a_i \in \mathcal{A}$ taken by each agent i. We are going to define the class of near-potential games, using a metric of closeness between games in the following.

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 $\mathcal{N}$ and the joint action sets $\mathcal{A}^N$ but possibly different utilities $\{u_i\}_{i \in \mathcal{N}}$ and $\{\hat{u}_i\}_{i \in \mathcal{N}}$. Further,*

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.

Sina Arefizadeh was with the Industrial and Systems Engineering Department, Texas A&M University, College Station, TX 77843.

let $d^{\Gamma}_{(a'_i,a)} := u_i(a'_i, a_{-i}) - u_i(a_i, a_{-i})$ *be the unilateral change in utility of an agent i by switching to an action $a'_i \in \mathcal{A}_i$, given joint action profile $a = (a_i, a_{-i}) \in \mathcal{A}^N$ in a game Γ. Then, the maximum pairwise difference $d(\Gamma, \hat{\Gamma})$ 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 unilateral changes of actions. Using MPD we define near-potential games as in [8].

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 $\delta \in \mathbb{R}^+$.*

Near potential games are another relaxation of potential games. Other well-known extensions of potential games are ordinal [4], weighted [4], and best-response potential games [10]. Finding the potential function of near-potential games is studied at [11] and a convex mathematical program is proposed. In this paper, our goal is to show the convergence of the decentralized FP in near-potential games.

III. DECENTRALIZED FICTITIOUS PLAY

Fictitious play is an individual decision-making rule, where agents are assumed to select actions according to a stationary distribution (strategy) $\sigma_i \in \Delta\mathcal{A}$, where $\Delta\mathcal{A}$ is the set of probability distributions over the action set $\mathcal{A}$. From now on, we suppose that each action $a \in \mathcal{A}$ is represented with an unit vector $\mathbf{e}_k \in \{0,1\}^K$ where $|\mathcal{A}| = K$. The stationary distribution of agent i at time t, denoted with $f_{i,t} \in \Delta\mathcal{A}$, is given by the empirical frequency of past actions taken by agent i,

$$f_{i,t} = \frac{t-1}{t} f_{i,t-1} + \frac{1}{t} a_{i,t}, \quad (3)$$

The individual actions of agent i is a result of it best responding to the stationary action distribution of other agents $f_{-i,t-1} := \{f_{j,t-1}\}_{j \in \mathcal{N} \setminus i}$,

$$a_{i,t} \in \arg \max_{a_i \in \mathcal{A}_i} u_i(a_i, f_{-i,t-1}), \quad (4)$$

where $u_i : \Delta\mathcal{A}^N \to \mathbb{R}$ can be defined as expected utility function on the set of probability distributions $\Delta\mathcal{A}^N$ with little abuse of notation. Given this definition, it also holds $(f_{i,t}, f_{-i,t}) \in \Delta\mathcal{A}^N$ and $a_i \in \Delta\mathcal{A}$, as a_i also becomes a degenerate distribution giving probability 1 to a selected action.

In a decentralized setting, it is not possible that the past actions of agent $j \in \mathcal{N} \setminus i$ ($f_{j,t}$) is available to agent i. An extension to FP assumes agents communicate over a network and exchange their beliefs about the empirical frequencies of other agents [9].

In particular, we assume agents communicate over a time-varying network $\mathcal{G} = (\mathcal{N}, \mathcal{E}_t)$, where each agent i only communicates with its current neighbors $\mathcal{N}_{i,t} := \{j : (i,j) \in \mathcal{E}_t\}$. In such a case, we replace the $f_{j,t} \in \Delta\mathcal{A}$ with a local

copy (belief) $v^i_{j,t} \in \Delta\mathcal{A}_j$ kept at agent i. Given $v^i_{i,t} = f_{i,t}$ agent i updates its local estimate with the local copies from its neighbors,

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

where $w^i_{jl,t} \geq 0$ is the weight that agent i puts on agent l's estimate of agent j such that $w^i_{j,l} \geq \eta$ for some $\eta > 0$ only if $l \in \mathcal{N}_i \cup \{i\}$, otherwise $w^i_{j,l} = 0$, and $\sum_{l \in \mathcal{N}_i \cup \{i\}} w^i_{j,l} = 1$ for all i,j.

In DFP, agent i selects an action $a_{i,t}$ to maximize its expected utility computed using its local beliefs $v^i_{-i,t-1} := \{v^i_{j,t-1}\}_{j \in \mathcal{N} \setminus i}$,

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

We summarize the steps of DFP algorithm below.

Algorithm 1 DFP for Agent i

1: **Input:** Local estimates v^i_{-i0} and time-varying networks $\{\mathcal{G}_t = (\mathcal{N}, \mathcal{E}_t)\}_{t \geq 1}$.
2: **for** $t = 1, 2, \cdots$ **do**
3:     Select an action a_{it} (6) and update $f_{i,t}$ (3).
4:     Share and update local copies $v^i_{j,t}$ (5).
5: **end for**

In the next section, we are going to show that joint empirical frequencies $f_t = (f_i, f_{-i,t}) \in \Delta\mathcal{A}^N$ converges around only one Nash Equilibrium (NE) of the closest potential game.

IV. CONVERGENCE OF DFP IN NEAR-POTENTIAL GAMES

A. Preliminaries

The joint strategy profile σ^* is an approximate NE, if agent i can obtain at most $\epsilon \geq 0$ utility value, by changing its strategy to another strategy.

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

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

We denote the set of ϵ-Nash equilibria in a game Γ with the notation Σ_ϵ. If $\epsilon = 0$ in (7), then σ^* is a NE strategy profile. Next, we provide the notion of upper semi-continuity.

Definition 5 (Upper Semi-continuous Correspondence) *A correspondence $h : X \Rightarrow Y$ 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.*

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 following assumptions,*
   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 $(i,j) \in \bigcup_{\tau=0}^{T_B-1} \mathcal{E}_{t+\tau}$.*

Assumption 1 $i) - ii)$ are referred as *connectivity* and *bounded communication interval* in order. This assumptions assures that any information about an agent j reaches agent i in some bounded time.

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 $t = 1, 2, \dots,$*
   (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) *$\sum_{l \in \mathcal{N}_{i,t} \cup \{i\}} w^i_{jl,t} = 1$.*

Assumption 2*(i)* indicates that agents place positive weight on estimates they receive from their current neighbors in (5). Assumption 2*(ii)* guarantees $\nu^i_{i,t} = f_{i,t}$ for all $t > 0$. Assumption 2*(iii)* assures that the weights $W_{j,t}$, defined as the collective update weights on agent j's empirical frequency $[W_{j,t}]_{i,l} = w^i_{jl,t}$, is row stochastic for all times. Now, we state the convergence of local estimates $\nu^i_{j,t}$ to empirical frequencies $f_{j,t}$—see [12] for the proof.

Lemma 1 (Proposition 1, [12]) *Suppose Assumptions 1-2 hold. If $f_{j0} = \nu^i_{j,0}$ holds for all pairs of agents $j \in \mathcal{N}$ and $i \in \mathcal{N}$, then the local copies $\{\nu^i_t\}^{i\in\mathcal{N}}_{t\geq 0}$ converge to the empirical frequencies $\{f_t\}_{t\geq 0}$ with rate $O(\log t/t)$, i.e., $\|\nu^i_{j,t} - f_{j,t}\| = O(\log t/t)$ for all $j \in \mathcal{N}$ and $i \in \mathcal{N}$.*

The proof mainly exploits the properties of row-stochastic matrices as per Assumption 2. Next result provides the difference potential value of empirical frequencies between consecutive time steps—see [9] for the proof.

Lemma 2 (Lemma 2, [9]) *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*

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

Lemma 2 suggests that after long enough time, if the empirical frequencies are outside the approximate NE region $N\delta$, the potential value of a close potential game increases. We next characterize the potential change if empirical frequencies follow an excursion path where they go outside of approximate-NE regions $N\delta + \epsilon_1$ and $N\delta + \epsilon_2$ in order, and return back firstly to $N\delta + \epsilon_2$ and then $N\delta + \epsilon_1$ given $0 < \epsilon_1 < \epsilon_2$.

Lemma 3 *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 that holds $f_{T_1-1} \in \Sigma_{N\delta+\epsilon_1}$ and $f_t \notin \Sigma_{N\delta+\epsilon_1}$, for all $T_1 \leq t < T'_1$,*
   - *T_2 is a time step that holds $f_{T_2-1} \in \Sigma_{N\delta+\epsilon_2}$ and $f_t \notin \Sigma_{N\delta+\epsilon_2}$, for all $T_2 \leq t < T'_2$,*
   - *T'_2 is a time step that holds $f_{T'_2-1} \notin \Sigma_{N\delta+\epsilon_2}$ and $f_{T'_2} \in \Sigma_{N\delta+\epsilon_2}$,*
   - *T'_1 is a time step that holds $f_{T'_1-1} \notin \Sigma_{N\delta+\epsilon_1}$ 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,

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

Proof: From (8), if $f_t \notin \Sigma_{N\delta+\epsilon_2}$, i.e., $T < T_2 \leq t < T'_2$, the following holds,

$$u(f_{t+1}) - u(f_t) \geq \frac{\epsilon_2}{(t+1)} - O\Big(\frac{\log t}{t^2}\Big) \geq \frac{2\epsilon_2}{3(t+1)}. \quad (10)$$

Then, (10) implies,

$$u(f_{T'_2}) - u(f_{T_2}) = \sum_{t=T_2}^{T'_2-1} u(f_{t+1}) - u(f_t) \geq \sum_{t=T_2}^{T'_2-1} \frac{2\epsilon_2}{3(t+1)}. \quad (11)$$

Similarly, for the time intervals, $T_1 \leq t < T_2$ or $T'_2 \leq t < T'_1$, the following inequalities also hold,

$$u(f_{T'_1}) - u(f_{T'_2}) = \sum_{t=T'_2}^{T'_1-1} u(f_{t+1}) - u(f_t) \geq \sum_{t=T'_2}^{T'_1-1} \frac{2\epsilon_1}{3(t+1)}, \quad (12)$$

$$u(f_{T_2}) - u(f_{T_1}) = \sum_{t=T_1}^{T_2-1} u(f_{t+1}) - u(f_t) \geq \sum_{t=T_1}^{T_2-1} \frac{2\epsilon_1}{3(t+1)}. \quad (13)$$

Thus, the result in (9) follows by (11), (12), (13), and the fact that $u(f_{T'_1}) - u(f_{T_1}) = (u(f_{T'_1}) - u(f_{T'_2})) + (u(f_{T'_2}) - u(f_{T_2})) + (u(f_{T_2}) - u(f_{T_1}))$. ∎

This result puts a lower bound on the excursion away from an approximate NE. We need the following additional assumptions to hold for the main convergence result.

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, $\Sigma_0 = \{\sigma^{*(1)}, \sigma^{*(2)}, \cdots, \sigma^{*(M)}\}$ for $M \in \mathbb{Z}^+$.*

Assumptions 3 asserts that there needs to be only a finite number of Nash equilibria.

Lemma 4 (Theorem 5.2, [8]) *Suppose Assumption 3 holds. Let $q : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ be a function defined as follows,*

$$q(\alpha) = \max_{\sigma \in \Sigma_\alpha} \min_{m \in \{1, \cdots, M\}} \|\sigma - \sigma^{*(m)}\|, \quad (14)$$

1000

where $\sigma^{*(m)}$ is a NE of the game Γ as defined in Assumption 3. Then, the function $q(\cdot)$ is i) weakly increasing, ii) upper semi-continuous, iii) satisfies $q(0) = 0$ and $\lim_{\alpha \to 0} q(\alpha) = 0$.

The function $q : \mathbb{R}_+ \to \mathbb{R}_+$ defines the largest distance between the set of approximate Nash equilibria and the set of Nash equilibria.

Assumption 4 *Let $q : \mathbb{R}_+ \to \mathbb{R}_+$ be the function as defined in (14). The maximum pairwise distance 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, i.e., $d^* = \min_{m' \neq m''} ||\sigma^{*(m')} - \sigma^{*(m'')}||$, where $m', m'' \in \{1, \cdots, M\}$.*

Assumption 4 provides a relation between the minimum distance between two different Nash equilibria of the near potential game, and the distance to a given potential game. Next, we state the main theorem of this study.

Theorem 1 *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 time enough $t > T$.*

Proof: Suppose that $\epsilon_1 = \bar{\epsilon}$ and $\epsilon_2 = \bar{\alpha} - N\bar{\delta}$ such that $0 < \epsilon_1 < \epsilon_2$ is satisfied as stated in Lemma 3. Further, using $\bar{\delta} > \delta$ by Assumption 4, it holds $N\delta + \epsilon_1 < N\delta + \epsilon_2 < \bar{\alpha} = N\bar{\delta} + \epsilon_2$. Then, by this relation, it also holds, $\Sigma_{N\delta+\epsilon_1} \subset \Sigma_{N\delta+\epsilon_2} \subset \Sigma_{N\bar{\delta}+\epsilon_2}$. Therefore, the sets $\Sigma_{N\delta+\epsilon_1}$ and $\Sigma_{N\delta+\epsilon_2}$ consist of disjoint neighborhoods of finitely many equilibrium points by Step 2 of Theorem 5.2 [8]. This implies that for any $\sigma \in \Sigma_{N\delta+\epsilon_1}$ or $\sigma \in \Sigma_{N\delta+\epsilon_2}$, the definition of q provides $||\sigma - \sigma^{*(m)}|| \leq q(N\delta + \epsilon_1)$ or $||\sigma - \sigma^{*(m)}|| \leq q(N\delta + \epsilon_2)$ in order, for exactly one equilibrium $\sigma^{*(m)}$.

To prove the statement, we are going to use a contradiction. Given the definitions of T_1, T_2, T_1', T_2' as in Lemma 3, we assume that the empirical frequencies f_t leaves from $N\delta + \epsilon_1$-NE set around a equilibrium point $\sigma^{*(m')}$, and then enters again into $N\delta + \epsilon_1$-NE set, but around another equilibrium point $\sigma^{*(m'')}$ so that the following relations are given as below,

$$f_{T_1-1} \in \Sigma_{N\delta+\epsilon_1} \quad \text{and} \quad ||f_{T_1-1} - \sigma^{*(m')}|| \leq q(N\delta + \epsilon_1), \tag{15}$$

$$f_{T_1'} \in \Sigma_{N\delta+\epsilon_1} \quad \text{and} \quad ||f_{T_1'} - \sigma^{*(m'')}|| \leq q(N\delta + \epsilon_1), \tag{16}$$

$$f_{T_2-1} \in \Sigma_{N\delta+\epsilon_2} \quad \text{and} \quad ||f_{T_2-1} - \sigma^{*(m')}|| \leq q(N\delta + \epsilon_2), \tag{17}$$

$$f_{T_2'} \in \Sigma_{N\delta+\epsilon_2} \quad \text{and} \quad ||f_{T_2'} - \sigma^{*(m'')}|| \leq q(N\delta + \epsilon_2). \tag{18}$$

Given (17) and (18), $d^* = \min_{m' \neq m''} ||\sigma^{*(m')} - \sigma^{*(m'')}||$, and the fact that $\bar{\alpha} > N\delta + \epsilon_2$ implies $q(N\delta + \epsilon_2) < q(\bar{\alpha}) <$

$d^*/4$, the distance between empirical frequencies $f_{T_2'}$ and f_{T_2-1} can be bounded below

$$||f_{T_2'} - f_{T_2-1}|| > \frac{d^*}{2}. \tag{19}$$

Next, we obtain an upper bound on the distance between empirical frequencies at consecutive time steps, using the update rule (3) and triangle inequality,

$$||f_{t+1} - f_t|| = \frac{1}{t+1}||f_t - a_t|| \leq (||f_t|| + ||a_t||) \leq \frac{2N}{t+1} \tag{20}$$

since it holds $||a_{i,t}|| \leq 1$ and $||f_{i,t}|| \leq 1$ for all $i \in \mathcal{N}$. Then, (19) suggests $||f_{T_2'} - f_{T_2}|| > \frac{d^*}{2}$ also holds, as $||f_{T_2} - f_{T_2-1}||$ is sufficiently small for a sufficiently large time instant $T_2 > T$. Together, upper and lower bounds from T_2 to T_2' can be provided as,

$$\sum_{t=T_2}^{T_2'-1} \frac{2N}{t+1} \geq \sum_{t=T_2}^{T_2'-1} ||f_{t+1} - f_t||,$$

$$\geq ||(\sum_{t=T_2}^{T_2'-1} f_{t+1} - f_t)|| = ||f_{T_2'} - f_{T_2}|| > \frac{d^*}{2}. \tag{21}$$

Since $\sum_{t=T_2}^{T_2'-1} \frac{2N}{t+1} > \frac{d^*}{2}$, it holds that $\sum_{t=T_2}^{T_2'-1} \frac{d^*(t+1)}{4N} < 1$ dividing $\frac{d^*}{2}$ by $\sum_{t=T_2}^{T_2'-1} \frac{2N}{t+1}$. Then, again multiplying $\sum_{t=T_2}^{T_2'-1} \frac{2\epsilon_2}{3(t+1)}$ by $\sum_{t=T_2}^{T_2'-1} \frac{d^*(t+1)}{4N}$, we obtain the following lower bound,

$$u(f_{T_1'}) - u(f_{T_1}) \geq \sum_{t=T_2}^{T_2'-1} \frac{2\epsilon_2}{3(t+1)} \geq \frac{\epsilon_2 d^*}{6N}. \tag{22}$$

For the next part, we are going to derive the lower bound on the difference between maximum and minimum potential values in the neighborhood of different NE points $\sigma^{*(m')}$ and $\sigma^{*(m'')}$. Let $\underline{\sigma}^{*(m')} = \arg\min_{\sigma \in ||\sigma - a^{*(m')}|| \leq N\delta+\epsilon_1} u(\sigma)$ and $\bar{\sigma}^{*(m'')} = \arg\max_{\sigma \in ||\sigma - a^{*(m')}|| \leq N\delta+\epsilon_1} u(\sigma)$ be the respective mixed action profiles that provides maximum and minimum values around two arbitrarily different equilibrium points. The difference between their potentials $u(\underline{\sigma}^{*(m')})$ and $u(\bar{\sigma}^{*(m'')})$ can be rewritten as,

$$u(\underline{\sigma}^{*(m')}) - u(\bar{\sigma}^{*(m'')})$$
$$= (u(\underline{\sigma}^{*(m')}) - u(f_{T_1'})) + (u(f_{T_1'}) - u(f_{T_1}))$$
$$+ (u(f_{T_1}) - u(\bar{\sigma}^{*(m'')})). \tag{23}$$

For each segment of (RHS) of (23), we can derive a bound. Using Lipschitz continuity and (15)-(18), it gives $u(\underline{\sigma}^{*(m')}) - u(f_{T_1'}) \geq -2Lq(N\delta + \epsilon_1)$ and $u(f_{T_1}) - u(\bar{\sigma}^{*(m'')}) \geq -2Lq(N\delta + \epsilon_1)$, where L is Lipschitz constant of utility functions. Moreover, using again Lipschitz continuity and (20), it follows that $u(f_{T_1}) - u(f_{T_1-1}) \geq -\frac{2NL}{T_1}$. Therefore, using (22), the following is provided as below,

$$u(\underline{\sigma}^{*(m')}) - u(\bar{\sigma}^{*(m'')}) \geq \frac{\epsilon_2 d}{6N} - 4Lq(N\delta+\epsilon_1) - \frac{2NL}{T_1}. \tag{24}$$

The values $\epsilon_2 = \bar{\alpha} - N\bar{\delta}$, $\epsilon_1 = \bar{\epsilon}$ can be replaced. Hence, given the facts that $\bar{\delta} > \delta$, and the function q is increasing,
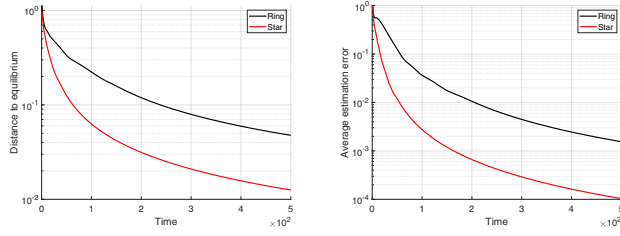
Fig. 1. DFP in target assignment game with unknown target locations over 20 runs. (Left) Average distance to Nash equilibrium $\frac{1}{N}\sum_{i\in\mathcal{N}}||f_{it}-\sigma_i^*||$ (Right) Average estimation error $\frac{1}{N(N-1)}\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{N}\setminus\{i\}}||f_{it}-v_{it}^j||$.

(24) takes the form,

$$u(\underline{\sigma}^{*(m')})-u(\bar{\sigma}^{*(m'')}) \geq \frac{(\bar{\alpha}-N\bar{\delta})d^*}{6N}-4Lq(N\bar{\delta}+\bar{\epsilon})-\frac{2NL}{T_1}. \quad (25)$$

We also have that $q(N\delta+\bar{\epsilon}) < \frac{(\bar{\alpha}-N\delta)d^*}{24NL}$ (see Theorem 5.2 in [8]). This yields the fact that $u(\underline{\sigma}^{*(m')})-u(\bar{\sigma}^{*(m'')}) > 0$, after long enough time $T_1 > T$, as $\frac{2NL}{T_1} \to 0$. Then, (20) and (21) also indicate that, there exists $\tilde{d} > 0$ and $\tilde{T}$ such that $\tilde{d} \geq \frac{2NL}{T_1+2} > |u(f_{T_1+1})-u(f_{T_1})|$. Consequently, there exists $\bar{d} > 0$, that implies $u(\underline{\sigma}^{*(m')}) > u(f_{T_1'+1}) > u(\bar{\sigma}^{*(m'')})$. However, by Lemma 2, if f_t is outside the set $\Sigma_{N\delta+\epsilon_1}$, the potential value has to increase until again entering into the approximate equilibrium set. This creates a contradiction, since the maximum potential $u(\bar{\sigma}^{*(m'')})$ around another equilibrium $a^{*(m'')}$ is less than $u(f_t)$. Thus, this holds for any equilibrium pair, and guarantees that the sequence f_t only visits around a single equilibrium after long time enough $t > T$. ∎

Theorem 1 is a sequel to the results in [9]. DFP creates additional error rate of $O(\log t/t^2)$ in (8), while FP has an error rate of $O(1/t^2)$ [8]. Since in both cases, the error rate goes to 0, DFP recovers the convergence results of FP with the same set of assumptions on the game structure.

V. Numerical Experiments

We use target assignment game with $N = 10$ autonomous agents and $K = 10$ targets in the numerical experiments. Each agent $i \in \mathcal{N}$ selects only one target $k \in \mathcal{K} := \{1,\cdots,K\}$ so that the objective is to cover all targets with minimum effort as a team, given the utility functions defined as below,

$$u_i(a_i,a_{-i}) = \frac{a_i^T \mathbb{1}_{a_{-ik}=0}}{a_i^T d_i}, \quad (26)$$

where $a_i = \mathbf{e}_k \in \mathbb{R}^K$ 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 kth index is equal to 0. The distance vector $d_i = [d_{i1},\cdots,d_{ik},\cdots,d_{iK}] \in \mathbb{R}_+^K$ captures the distance between targets and agent i. We assume that target locations are unknown, and agents obtain noisy private signals about target locations coming from a normal distribution with the mean equal to the target locations and variance 0.1 for each dimension (x and y axes) of the target location independently. Then, agents take sample average of private signal for targets coming at each time step t. Agents stop receiving

these signals after time step $t = 10$. Target assignment game with equal distances between targets and agents is a potential game. We convert it into a near-potential game by introducing different estimates and distances to targets. We use ring and stars communication networks. Self-weights are set as $w_{i,l}^i = 0.75$ while $w_{j,l}^i = 0.25/|\mathcal{N}_i|$, for all $j \in \mathcal{N}_i$ is selected for neighbor agents.

We implemented 20 runs of numerical experiments with randomly created distances. Fig. 1 (Right) shows the average estimation error between local and real empirical frequencies. It corroborates the convergence rate of $O(\log t/t)$ given in Lemma 2. Further, we observe that agents' final action profiles converge to an one-to-one assignment between agents and targets in all cases. The star network converges faster than the ring network. Fig. 1 (Left) shows the average rate of convergence of the empirical frequencies to a NE of the target assignment game which is an one-to-one assignment of agents to targets. It also confirms the result of convergence around a single equilibrium point.

VI. Conclusion

In this paper, we studied the convergence of DFP in near-potential games. We proved empirical frequencies of actions converge to an approximate NE around a single NE point. That is, DFP preserves the convergence properties of FP, despite the lack of perfect information on others' past actions. This result implies that decentralized and repetitive best-response type behavior in large-scale networked systems, converge to rational behavior.

References

[1] 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.

[2] 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.

[3] U. O. Candogan, I. Menache, A. Ozdaglar, and P. A. Parrilo, "Near-optimal power control in wireless networks: A potential game approach," in *2010 Proceedings IEEE INFOCOM*. IEEE, 2010, pp. 1–9.

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

[5] B. Swenson, C. Eksin, S. Kar, and A. Ribeiro, "Distributed inertial best-response dynamics," *IEEE Transactions on Automatic Control*, vol. 63, no. 12, pp. 4294–4300, 2018.

[6] 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, 1996.

[7] B. Swenson, S. Kar, and J. Xavier, "Single sample fictitious play," *IEEE Transactions on Automatic Control*, vol. 62, no. 11, pp. 6026–6031, 2017.

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

[9] S. Aydın, S. Arefizadeh, and C. Eksin, "Decentralized fictitious play in near-potential games with time-varying communication networks," *IEEE Control Systems Letters*, 2021.

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

[11] 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.

[12] 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.