2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Learning in Near-Potential Games

Ozan Candogan, Asuman Ozdaglar and Pablo A. Parrilo

Abstract— **Except for special classes of games, there is no systematic framework for analyzing the dynamical properties of multi-agent strategic interactions. Potential games are one such special but restrictive class of games that allow for tractable dynamic analysis. Intuitively, games that are "close" to a potential game should share similar properties. In this paper, we formalize and develop this idea by quantifying to what extent the dynamic features of potential games extend to "near-potential" games. We first show that in an arbitrary finite game, the limiting behavior of better-response and best-response dynamics can be characterized by the approximate equilibrium set of a close potential game. Moreover, the size of this set is proportional to a closeness measure between the original game and the potential game. We then focus on logit response dynamics, which induce a Markov process on the set of strategy profiles of the game, and show that the stationary distribution of logit response dynamics can be approximated using the potential function of a close potential game, and its stochastically stable strategy profiles can be identified as the approximate maximizers of this function. Our approach presents a systematic framework for studying convergence behavior of adaptive learning dynamics in finite strategic form games.**

I. Introduction

The study of multi-agent strategic interactions both in economics and engineering mainly relies on the concept of Nash equilibria. A key justification for Nash equilibrium is that adaptive learning dynamics that involve less than fully rational behavior converges to Nash equilibria in the long run. Nevertheless, such results have only been established for some special (but restrictive) classes of games, potential games is an example [1]–[3].

Our goal in this paper is to provide a systematic framework for studying dynamics in finite strategic-form games by exploiting their relation to "close" potential games. Our approach relies on using the potential function of a close potential game for the analysis of dynamics in the original game. This enables us to establish convergence of commonly studied update rules to the set of approximate equilibria, where the size of the set is a function of the distance from a close potential game.[1] We note that our results hold for arbitrary strategic form games, however the bounds on the limiting sets are tighter and hence more informative for games that are close to potential games (in terms of payoffs of the players). We therefore focus our investigation to such games in this paper and refer to them as *near-potential games*.

We study the limiting behavior of three specific update rules. We first focus on *better-response* and *best-response* dynamics. It is known that trajectories of these update rules (i.e., the sequence of strategy profiles generated under these update rules) converge to pure Nash equilibria in potential games [3], [4]. However, in near-potential games pure Nash equilibria need not even exist. For this reason we focus on the notion of *approximate equilibria* or *ϵ-equilibria*, and show that in near-potential games, trajectories of these update rules converge to an approximate equilibrium set, where the size of the set depends on the distance of the original game from a potential game. We also show by means of an example that in general the asymptotic behavior in two close games can be very different (see Example 1). This highlights the importance of closeness to potential games in shaping the dynamical properties of update rules.

We then focus on *logit response* update rule. With this update rule, agents, when updating their strategies, choose their best-responses with high probability, but also explore other strategies with nonzero probability. Logit response induces a random walk on the set of strategy profiles. The stationary distribution of the random walk is used to explain the limiting behavior of this update rule [5], [6]. In potential games, the stationary distribution can be expressed in closed form. Additionally, the *stochastically stable strategy profiles*, i.e., the strategy profiles which have nonzero stationary distribution as the exploration probability goes to zero, are those that maximize the potential function [5], [6]. Exploiting their relation to close potential games, we obtain similar results for near-potential games: (i) we obtain a closed-form expression that approximates the stationary distribution, and (ii) we show that the stochastically stable strategy profiles are the strategy profiles that approximately maximize the potential of a close potential game. Our analysis relies on a novel perturbation result for Markov chains (see Theorem 2) which provides bounds on deviations from a stationary distribution when transition probabilities of a Markov chain are *multiplicatively* perturbed, and therefore may be of independent interest.

A summary of our findings on better/best-response dynamics and logit response can be found in Table I.

The framework proposed in this paper provides a systematic approach for studying limiting behavior of adaptive learning dynamics in strategic form games: Given a finite game, we find a close potential game by solving a convex optimization problem [7], [8]. We then characterize the dynamical properties of the original game using the properties

This research was funded in part by National Science Foundation grants CMMI-0545910, ECCS-1027922, and by the AFOSR MURI R6756-G2.

All authors are with the Laboratory of Information and Decision Systems, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Room 32-D608 Cambridge, MA 02139.

emails: {candogan,asuman,parrilo}@mit.edu

[1]Throughout the paper, we use the terms *learning dynamics* and *update rules* interchangeably.

Update Process	Dynamics Result
Best-Response Dynamics	(Theorem 1) Trajectories converge to $\mathcal{X}_{\delta h}$, i.e., the δh-equilibrium set of $\mathcal{G}$.
Better-Response Dynamics	(Theorem 1) Trajectories converge to $\mathcal{X}_{\delta h}$, i.e., the δh-equilibrium set of $\mathcal{G}$.
Logit Response Dynamics (with parameter τ)	(Corollary 2) Stationary distributions are such that $\vert\mu(\mathbf{p}) - \hat{\mu}(\mathbf{p})\vert \leq \frac{\delta(h-1)}{\tau}$, for all $\mathbf{p}$.
Logit Response Dynamics	(Corollary 3) Stochastically stable strategy profiles of $\mathcal{G}$ are (i) contained in $S = \{\mathbf{p} \vert \phi(\mathbf{p}) \geq \max_{\mathbf{q}} \phi(\mathbf{q}) - 4(h-1)\delta\}$, (ii) $\delta(4h-3)$-equilibria of $\mathcal{G}$.

TABLE I: Properties of better/best-response and logit response dynamics in near-potential games. We assume that the dynamical properties of a game $\mathcal{G}$ are studied, and there exists a nearby potential game $\hat{\mathcal{G}}$, with potential function ϕ such that the distance (in terms of the maximum pairwise difference, defined in Section II) between the two games is δ. We use the notation $\mathcal{X}_\epsilon$ to denote the ϵ-equilibrium set of the original game, M and h to denote the number of players and strategy profiles, μ and $\hat{\mu}$ to denote the stationary distributions of logit response dynamics in $\mathcal{G}$ and $\hat{\mathcal{G}}$, respectively.

of this potential game. The characterization is tighter if the "distance" between the games is smaller.

Related Literature: Potential games play an important role in game-theoretic analysis because of existence of pure strategy Nash equilibrium, and the stability (under various learning dynamics such as better-, best-response dynamics, and fictitious play) of pure Nash equilibria in these games [1], [3], [4], [9]–[12]. Because of these properties, potential games found applications in various control and resource allocation problems [3], [13]–[15].

There is no systematic framework for analyzing the limiting behavior of many of the adaptive update rules in general games [1], [2], [16]. However, for potential games there is a long line of literature establishing convergence of natural adaptive dynamics such as better/best-response dynamics [3], [4], fictitious play [9], [10], [17] and logit response dynamics [5], [6], [12].

It was shown in recent work that a close potential game to a given game can be obtained by solving a convex optimization problem [7], [8]. It was also proved that equilibria of a given game can be characterized by first approximating this game with a potential game, and then using the equilibrium properties of close potential games [7], [8]. This paper builds on this line of work to study dynamics in games by exploiting their relation to a close potential game.

Paper Organization: The rest of the paper is organized as follows: We present the game theoretic preliminaries for our work in Section II. We present an analysis of better- and best-response dynamics in near-potential games in Section III. In Section IV, we extend our analysis to logit response, and focus on the stationary distribution and stochastically stable states of logit response. We close in Section V with concluding remarks and future work. Due to space constraints all proofs are omitted and can be found in [18].

II. Preliminaries

In this section, we present the game-theoretic background that is relevant to our work. Additionally, we introduce the closeness measure for games that is used in the rest of the paper.

A. Finite Strategic-Form Games

A (noncooperative) finite game in strategic-form consists of:

- A finite set of players, denoted by $\mathcal{M} = \{1, \ldots, M\}$.
- Strategy spaces: A finite set of strategies (or actions) E^m, for every $m \in \mathcal{M}$.
- Utility functions: $u^m : \prod_{k \in \mathcal{M}} E^k \to \mathbb{R}$, for every $m \in \mathcal{M}$.

A (strategic-form) game instance is accordingly given by the tuple $\langle \mathcal{M}, \{E^m\}_{m\in\mathcal{M}}, \{u^m\}_{m\in\mathcal{M}} \rangle$. The joint strategy space of a game is denoted by $E = \prod_{m\in\mathcal{M}} E^m$. We refer to a collection of strategies of all players as a *strategy profile* and denote it by $\mathbf{p} = (p^1, \ldots, p^M) \in E$. The strategies of all players but the mth one is denoted by $\mathbf{p}^{-m}$.

We use (pure) Nash equilibrium and ϵ-Nash equilibrium solution concepts for games. Formally, a strategy profile $\mathbf{p} \triangleq (p^1, \ldots, p^M)$ is an ϵ-equilibrium ($\epsilon \geq 0$) if

$$u^m(q^m, \mathbf{p}^{-m}) - u^m(p^m, \mathbf{p}^{-m}) \leq \epsilon \qquad (1)$$

for every $q^m \in E^m$ and $m \in \mathcal{M}$. The set of all ϵ-Nash equilibria of a game are denoted by $\mathcal{X}_\epsilon$. If $\mathbf{p}$ satisfies (1) with $\epsilon = 0$, then it is a (pure) Nash equilibrium. Thus, a Nash equilibrium is a strategy profile from which no player can unilaterally deviate and improve its payoff. On the other hand, ϵ-equilibria are strategy profiles that satisfy the unilateral deviation constraints approximately. For this reason, in the rest of the paper we use the terms approximate Nash equilibrium and ϵ-equilibrium interchangeably.

B. Potential Games

We next describe a particular class of games that is central in this paper, the class of potential games [3]:

Definition 2.1 (Potential Game): A potential game is a noncooperative game for which there exists a function $\phi : E \to \mathbb{R}$ satisfying

$$u^m(p^m, \mathbf{p}^{-m}) - u^m(q^m, \mathbf{p}^{-m}) = \phi(p^m, \mathbf{p}^{-m}) - \phi(q^m, \mathbf{p}^{-m}),$$

for every $m \in \mathcal{M}$, $p^m, q^m \in E^m$, $\mathbf{p}^{-m} \in E^{-m}$. The function ϕ is referred to as a *potential* function of the game.

Some properties that are specific to potential games are evident from the definition. For instance, it can be seen that unilateral deviations from a strategy profile that maximizes the potential function (weakly) decrease the utility of the deviating player. Hence, this strategy profile corresponds to a Nash equilibrium, and it follows that every potential game has a pure Nash equilibrium.

We next formally define the measure of "closeness" of games, used in the subsequent sections.

Definition 2.2 (Maximum Pairwise Difference): Let $\mathcal{G}$ and $\hat{\mathcal{G}}$ be two games with set of players $\mathcal{M}$, set of strategy profiles E, and collections of utility functions $\{u^m\}_{m\in\mathcal{M}}$ and $\{\hat{u}^m\}_{m\in\mathcal{M}}$ respectively. The *maximum pairwise difference* (MPD) between these games is defined as

$$d(\mathcal{G},\hat{\mathcal{G}}) \triangleq \max_{\mathbf{p}\in E, m\in\mathcal{M}, q^m\in E^m} \left|\left(u^m(q^m,\mathbf{p}^{-m}) - u^m(p^m,\mathbf{p}^{-m})\right) - \left(\hat{u}^m(q^m,\mathbf{p}^{-m}) - \hat{u}^m(p^m,\mathbf{p}^{-m})\right)\right|.$$

We refer to pairs of games with small MPD as *nearby games*, and games that have a small MPD to a potential game as *near-potential games*.

Note that the pairwise difference $u^m(q^m,\mathbf{p}^{-m}) - u^m(p^m,\mathbf{p}^{-m})$ quantifies how much player m can improve its utility by unilaterally deviating from strategy profile $(p^m,\mathbf{p}^{-m})$ to strategy profile $(q^m,\mathbf{p}^{-m})$. The MPD measures the closeness of games in terms of the difference of these unilateral deviations, rather than the difference of their utility functions, i.e., quantities of the form $|(u^m(q^m,\mathbf{p}^{-m}) - u^m(p^m,\mathbf{p}^{-m})) - (\hat{u}^m(q^m,\mathbf{p}^{-m}) - \hat{u}^m(p^m,\mathbf{p}^{-m}))|$ are used to identify close games, rather than quantities of the form $|u^m(p^m,\mathbf{p}^{-m}) - \hat{u}^m(p^m,\mathbf{p}^{-m})|$. This is because difference in unilateral deviations provides a better characterization of the strategic similarities (equilibrium and dynamic properties) between two games than the difference of the utility functions.[2] This can be seen from the following example: Consider two games with utility functions $\{u^m\}$ and $\{u^m+1\}$, i.e., in the second game players receive an additional payoff of 1 at all strategy profiles. It can be seen from the definition of Nash equilibrium that despite the difference of their utility functions, these two games share the same equilibrium set. Intuitively, since the additional payoff is obtained at all strategy profiles, it does not affect any of the strategic considerations in the game. On the other hand, it can be seen that the MPD of these games is equal to zero, thus MPD identifies strategic equivalence between them. Details of strategic equivalence in games and its relation to unilateral deviations can be found in [7].

It can be seen from Definition 2.1 if and only if it satisfies some linear equalities. This suggests that the set of potential games is convex. Hence, the closest (in terms of MPD, or any given norm) potential game to a given game can be found by solving a convex optimization problem (see [7] and [8], for closed form solution obtained using an $L2$-norm). In the rest of the paper, we do not discuss how a nearby potential game to a given game is obtained, but we just assume that a nearby potential game with potential ϕ is known and the MPD between this game and the original game is δ. We obtain results on convergence of dynamics in the original game, using the properties of ϕ, and δ.

[2]MPD can be thought of as the infinity norm of the differences of unilateral deviations in games. Alternative distance measures can be defined using two-norm or one-norm in place of the infinity norm. The closeness notion in Definition 2.2 provides tighter characterization of the limiting behavior (such as the sets trajectories converge to or stationary distribution) of update rules, and hence is preferred in this paper.

III. Better-Response and Best-Response Dynamics

In this section, we consider better-response and best-response dynamics, and study convergence properties of these update rules in near-potential games. Best-response dynamics is an update rule where at each time step a player chooses its best-response to the opponents' current strategy profile. In better-response dynamics, on the other hand, players choose strategies that improve their payoffs, but these strategies need not be their best-responses. Formal descriptions of these update rules are given below.

Definition 3.1 (Better- and Best-Response Dynamics): At each time instant $t \in \{1,2,\dots\}$, a single player is chosen at random for updating its strategy, using a probability distribution with full support over the set of players. Let m be the player chosen at some time t, and let $\mathbf{r}\in E$ denote the strategy profile that is used at time $t-1$.

1) Better-response dynamics is the update process where player m does not modify its strategy if $u^m(\mathbf{r}) = \max_{q^m} u^m(q^m,\mathbf{r}^{-m})$, and otherwise it updates its strategy to a strategy in $\{q^m | u^m(q^m,\mathbf{r}^{-m}) > u^m(\mathbf{r})\}$, chosen uniformly at random.
2) Best-response dynamics is the update process where player m does not modify its strategy if $u^m(\mathbf{r}) = \max_{q^m} u^m(q^m,\mathbf{r}^{-m})$, and otherwise it updates its strategy to a strategy in $\arg\max_{q^m} u^m(q^m,\mathbf{r}^{-m})$, chosen uniformly at random.

We refer to strategies in $\arg\max_{q^m} u^m(q^m,\mathbf{r}^{-m})$ as *best-responses of player m to* $\mathbf{r}^{-m}$. We denote the strategy profile used at time t by $\mathbf{p}_t$, and we define the trajectory of the dynamics as the sequence of strategy profiles $\{\mathbf{p}_t\}_{t=0}^{\infty}$. In our analysis, we assume that the trajectory is initialized at a strategy profile $\mathbf{p}_0 \in E$ at time 0 and it evolves according to one of the update rules described above.

The following theorem establishes that in finite games, better- and best-response dynamics converge to a set of ϵ-equilibria, where the size of this set is characterized by the MPD to a close potential game.

Theorem 1: Consider a game $\mathcal{G}$ and let $\hat{\mathcal{G}}$ be a nearby potential game, such that $d(\mathcal{G},\hat{\mathcal{G}}) = \delta$. Assume that best-response or better-response dynamics are used in $\mathcal{G}$, and denote the number of strategy profiles in these games by $|E| = h$.

For both update processes, the trajectories are contained in the δh-equilibrium set of $\mathcal{G}$ after finite time with probability 1, i.e., let T be a random variable such that $\mathbf{p}_t \in \mathcal{X}_{\delta h}$, for all $t > T$, then $P(T < \infty) = 1$.

The above theorem also implies that better- and best-response dynamics converge to a Nash equilibrium in potential games, since if $\mathcal{G}$ is a potential game, the nearby potential game $\hat{\mathcal{G}}$ can be chosen such that $d(\mathcal{G},\hat{\mathcal{G}}) = 0$.

Our approach for extending dynamical properties of potential games to nearby games relies on the special structural properties of potential games. Surprisingly, in general, dynamical properties of arbitrary close games can be quite different, as illustrated in Example 1.

Example 1: Consider two games with two players and payoffs given in Figure 1. The entries of these tables indexed by row X and column Y show payoffs of the players when the first player uses strategy X and the second player uses strategy Y. Let $0 < \theta \ll 1$, and assume that players update their strategies according to better-response dynamics. In the game on the left, this update rule converges to strategy profile (C, C). In the game on the right, updates do not converge (it can be shown that strategy profile (A, A) is visited infinitely often), and moreover the trajectory of the updates is not contained in any ϵ-equilibrium set of the game for $\epsilon < 1$. Thus, in arbitrary games, even a small change in the payoffs (2θ in the example), results in significantly different dynamic behavior, unlike near-potential games as established in Theorem 1.

	A	B	C
A	0, 1	1, 0	$\frac{\theta}{2}$, 0
B	1, 0	0, 1	0, 1
C	1, 0	0, $\frac{\theta}{2}$	θ, θ

	A	B	C
A	0, 1	1, 0	$\frac{\theta}{2}$, 0
B	1, 0	0, 1	0, 1
C	1, 0	0, $\frac{\theta}{2}$	-θ, -θ

Fig. 1: Better-response dynamics converge to the Nash equilibrium (C,C) for the game on the left. The game on the right can be obtained by perturbing the payoffs in the first game by 2θ. For this game, trajectories do not converge, and are not contained in any ϵ-equilibrium set for $\epsilon < 1$.

IV. Logit Response Dynamics

In this section we focus on logit response dynamics and characterize the stochastically stable states and stationary distribution of this update rule in near-potential games. In Section IV-A, we provide a formal definition of logit response dynamics, and review some of its properties. In Section IV-B, we show that the stationary distribution of logit response dynamics in a near-potential game can be approximately characterized in terms of the potential function of a close potential game. Moreover, we focus on the stochastically stable states of this update rule and show that the stochastically stable states are contained in the approximate equilibrium sets in near-potential games.

A. Properties of Logit Response

We start by providing a formal definition of logit response dynamics:

Definition 4.1: At each time instant $t \in \{1, 2, \dots\}$, a single player is chosen at random for updating its strategy, using a probability distribution with full support over the set of players. Let m be the player chosen at some time t, and let $\mathbf{r} \in E$ denote the strategy profile that is used at time $t-1$.

Logit response dynamics with parameter τ is the update process, where player m chooses a strategy $q^m \in E^m$ with probability

$$P_\tau^m(q^m|\mathbf{r}) = \frac{e^{\frac{1}{\tau}u^m(q^m,\mathbf{r}^{-m})}}{\sum_{p^m\in E^m} e^{\frac{1}{\tau}u^m(p^m,\mathbf{r}^{-m})}}.$$

In this definition, $\tau > 0$ is a fixed parameter that determines how often players choose their best-responses. The probability of not choosing a best-response decreases as τ decreases, and as $\tau \to 0$, players choose their best-responses with probability 1. This feature suggests that logit response dynamics can be viewed as a generalization of best-response dynamics, where with small but nonzero probability players use a strategy that is not a best-response.

For $\tau > 0$, this update process can be represented as a finite, aperiodic, irreducible Markov chain [5], [6]. The states of the Markov chain correspond to the strategy profiles in the game. Denoting the probability that player m is chosen for a strategy update by α_m, transition probability from strategy profile $\mathbf{p}$ to $\mathbf{q}$ can be given by (assuming $\mathbf{p} \neq \mathbf{q}$, and denoting the transition from $\mathbf{p}$ to $\mathbf{q}$ by $\mathbf{p} \to \mathbf{q}$):

$$P_\tau(\mathbf{p} \to \mathbf{q}) = \begin{cases} \alpha_m P_\tau^m(q^m|\mathbf{p}), & \text{if } \mathbf{q}^{-m} = \mathbf{p}^{-m} \\ & \text{for some } m \in \mathcal{M} \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

The chain is aperiodic and irreducible because a player updating its strategy can choose any strategy (including the current one) with positive probability.

We denote the stationary distribution of this Markov chain by μ_τ. A strategy profile $\mathbf{q}$ such that $\lim_{\tau\to 0} \mu_\tau(\mathbf{q}) > 0$ is referred to as a *stochastically stable strategy profile*. Intuitively, these strategy profiles are the ones that are used with nonzero probability, as players adopt their best-responses more and more frequently in their strategy updates.

In potential games, the stationary distribution of the logit response dynamics can be written as an explicit function of the potential. If $\mathcal{G}$ is a potential game with potential function ϕ, the stationary distribution of the logit response dynamics is given by the distribution [5], [6]:[3]

$$\mu_\tau(\mathbf{q}) = \frac{e^{\frac{1}{\tau}\phi(\mathbf{q})}}{\sum_{\mathbf{p}\in E} e^{\frac{1}{\tau}\phi(\mathbf{p})}}. \quad (3)$$

It can be seen from (3) that $\lim_{\tau\to 0} \mu_\tau(\mathbf{q}) > 0$ if and only if $\mathbf{q} \in \arg\max_{\mathbf{p}\in E} \phi(\mathbf{p})$. Thus, in potential games the stochastically stable strategy profiles are those that maximize the potential function.

B. Stationary Distribution and Stochastically Stable Strategy Profiles of Logit Response

In this section we show that the stationary distribution of logit response dynamics in near-potential games can be approximated by exploiting the potential function of a nearby potential game. We then use this result to identify stochastically stable strategy profiles in near-potential games.

We start by showing that in games with small MPD, logit response dynamics have similar transition probabilities.

Lemma 1: Consider a game $\mathcal{G}$ and let $\hat{\mathcal{G}}$ be a nearby potential game, such that $d(\mathcal{G}, \hat{\mathcal{G}}) = \delta$. Denote the transition probability matrices of logit response dynamics in $\mathcal{G}$ and $\hat{\mathcal{G}}$

[3] Note that this expression is independent of $\{\alpha_m\}$, i.e., the probability distribution that is used to choose which player updates its strategy has no effect on the stationary distribution of logit response.

by P_τ and $\hat{P}_\tau$ respectively. For all strategy profiles $\mathbf{p}$ and $\mathbf{q}$ that differ in the strategy of at most one player, we have

$$e^{-\frac{2\delta}{\tau}} \leq \hat{P}_\tau(\mathbf{p} \to \mathbf{q})/P_\tau(\mathbf{p} \to \mathbf{q}) \leq e^{\frac{2\delta}{\tau}}.$$

Definition 4.1 suggests that perturbation of utility functions changes the transition probabilities multiplicatively in logit response. The above lemma supports this intuition: if utility gains due to unilateral deviations are modified by δ, the ratio of the transition probabilities can change at most by $e^{\frac{2\delta}{\tau}}$. Thus, if two games are close, then the transition probabilities of logit response in these games should be closely related.

This suggests using results from perturbation theory of Markov chains to characterize the stationary distribution of logit response in a near-potential game [21], [22]. However, standard perturbation results characterize changes in the stationary distribution of a Markov chain when the transition probabilities are *additively perturbed*. These results, when applied to multiplicative perturbations, yield bounds which are uninformative. We therefore first present a result which characterizes deviations from the stationary distribution of a Markov chain when its transition probabilities are multiplicatively perturbed, and therefore may be of independent interest.[4]

Theorem 2: Let P and $\hat{P}$ denote the probability transition matrices of two finite irreducible Markov chains with the same state space. Denote the stationary distributions of these Markov chains by μ and $\hat{\mu}$ respectively, and let the cardinality of the state space be h. Assume that $\alpha \geq 1$ is a given constant and for any two states $\mathbf{p}$ and $\mathbf{q}$, the following hold

$$\alpha^{-1} P(\mathbf{p} \to \mathbf{q}) \leq \hat{P}(\mathbf{p} \to \mathbf{q}) \leq \alpha P(\mathbf{p} \to \mathbf{q}).$$

Then, for any state $\mathbf{p}$, we have

$$(i) \quad \frac{\alpha^{-(h-1)}\mu(\mathbf{p})}{\alpha^{-(h-1)}\mu(\mathbf{p}) + \alpha^{h-1}(1-\mu(\mathbf{p}))} \leq \hat{\mu}(\mathbf{p}) \leq \frac{\alpha^{h-1}\mu(\mathbf{p})}{\alpha^{h-1}\mu(\mathbf{p}) + \alpha^{-(h-1)}(1-\mu(\mathbf{p}))}$$

and, (ii) $|\mu(\mathbf{p}) - \hat{\mu}(\mathbf{p})| \leq \frac{\alpha^{h-1}-1}{\alpha^{h-1}+1}$.

We next use the above theorem to relate the stationary distributions of logit response dynamics in close games.

Corollary 1: Let $\mathcal{G}$ and $\hat{\mathcal{G}}$ be finite games with number of strategy profiles $|E| = h$, such that $d(\mathcal{G}, \hat{\mathcal{G}}) = \delta$. Denote the stationary distributions of logit response dynamics in these games by μ_τ, and $\hat{\mu}_\tau$ respectively. Then, for any strategy profile $\mathbf{p}$ we have

$$(i) \quad \frac{e^{-\frac{2\delta(h-1)}{\tau}}\mu(\mathbf{p})}{e^{-\frac{2\delta(h-1)}{\tau}}\mu(\mathbf{p}) + e^{\frac{2\delta(h-1)}{\tau}}(1-\mu(\mathbf{p}))} \leq \hat{\mu}(\mathbf{p}) \leq \frac{e^{\frac{2\delta(h-1)}{\tau}}\mu(\mathbf{p})}{e^{\frac{2\delta(h-1)}{\tau}}\mu(\mathbf{p}) + e^{-\frac{2\delta(h-1)}{\tau}}(1-\mu(\mathbf{p}))},$$

and (ii) $|\mu(\mathbf{p}) - \hat{\mu}(\mathbf{p})| \leq \frac{e^{\frac{2\delta(h-1)}{\tau}}-1}{e^{\frac{2\delta(h-1)}{\tau}}+1}$.

The above corollary can be adapted to near-potential games, by exploiting the relation of stationary distribution of logit response and potential function in potential games (see (3)).

Corollary 2: Consider a game $\mathcal{G}$ and let $\hat{\mathcal{G}}$ be a nearby potential game, such that $d(\mathcal{G}, \hat{\mathcal{G}}) = \delta$. Denote the potential function of $\hat{\mathcal{G}}$ by ϕ, and number of strategy profiles in these games by $|E| = h$. Then, the stationary distribution μ_τ of logit response dynamics in $\mathcal{G}$ is such that

$$(i) \quad \frac{e^{\frac{1}{\tau}(\phi(\mathbf{p})-2\delta(h-1))}}{e^{\frac{1}{\tau}(\phi(\mathbf{p})-2\delta(h-1))} + \sum_{\mathbf{q}\neq\mathbf{p}\in E} e^{\frac{1}{\tau}(\phi(\mathbf{q})+2\delta(h-1))}} \leq \mu_\tau(\mathbf{p}) \leq \frac{e^{\frac{1}{\tau}(\phi(\mathbf{p})+2\delta(h-1))}}{e^{\frac{1}{\tau}(\phi(\mathbf{p})+2\delta(h-1))} + \sum_{\mathbf{q}\neq\mathbf{p}\in E} e^{\frac{1}{\tau}(\phi(\mathbf{q})-2\delta(h-1))}},$$

$$\text{and, (ii)} \quad \left|\mu(\mathbf{p}) - \frac{e^{\frac{1}{\tau}\phi(\mathbf{p})}}{\sum_{\mathbf{q}\in E} e^{\frac{1}{\tau}\phi(\mathbf{q})}}\right| \leq \frac{e^{\frac{2\delta(h-1)}{\tau}}-1}{e^{\frac{2\delta(h-1)}{\tau}}+1}.$$

With simple manipulations, it can be shown that $(e^x - 1)/(e^x+1) \leq x/2$ for $x \geq 0$. Thus, (ii) in the above corollary implies that $\left|\mu(\mathbf{p}) - \frac{e^{\frac{1}{\tau}\phi(\mathbf{p})}}{\sum_{\mathbf{q}\in E} e^{\frac{1}{\tau}\phi(\mathbf{q})}}\right| \leq \frac{\delta(h-1)}{\tau}$. Therefore, the stationary distribution of logit response in a near-potential game can be approximated by the stationary distribution of this update rule in a potential game close to this game. When τ is fixed and $\delta \to 0$, i.e., when the original game is arbitrarily close to a potential game, the stationary distribution of logit response is arbitrarily close to the stationary distribution of the potential game. On the other hand, for a fixed δ, as $\tau \to 0$, the upper bound in (ii) becomes uninformative. This is the case since $\tau \to 0$ implies that players adopt their best-responses with probability 1, and thus the stationary distribution of the update rule becomes very sensitive to the difference of the game from a potential game.

We conclude this section by characterizing the stochastically stable states of logit response dynamics in near-potential games.

Corollary 3: Consider a game $\mathcal{G}$ and let $\hat{\mathcal{G}}$ be a nearby potential game, such that $d(\mathcal{G}, \hat{\mathcal{G}}) = \delta$. Denote the potential function of $\hat{\mathcal{G}}$ by ϕ, and the number of strategy profiles in these games by $|E| = h$. The stochastically stable strategy profiles of $\mathcal{G}$ are (i) contained in $S = \{\mathbf{p} | \phi(\mathbf{p}) \geq \max_{\mathbf{q}} \phi(\mathbf{q}) - 4(h-1)\delta\}$, (ii) ϵ-equilibria of $\mathcal{G}$, where $\epsilon = (4h-3)\delta$.

This result provides a novel approach for characterizing stochastically stable states of logit response in near-potential games, without explicitly computing the stationary distribution.[5] Moreover, for an arbitrary game, one can check if there is a close potential game to this game, and use the above result to certify whether a given strategy profile is stochastically stable.

[4] A multiplicative perturbation bound similar to ours, can be found in [19]. However, this bound is loose and it does not provide a good characterization of the stationary distribution in our setting. We provide a tighter bound, and obtain stronger predictions on the stationary distribution of logit response.

[5] A tighter version of this result, which shows that the stochastically stable states are contained in $\{\mathbf{p} | \phi(\mathbf{p}) \geq \max_{\mathbf{q}} \phi(\mathbf{q}) - (|E|-1)\delta\}$, can be obtained, using the characterization of the stochastically stable states via resistance tree methods (see [6], [20] for an overview of these methods). We decided to omit this result from the paper due to space constraints.

V. Conclusions

We study the limiting behavior of learning dynamics in strategic form games by exploiting their relation to a nearby potential game. We focus our attention to better/best-response and logit response dynamics. We show that for near-potential games better/best-response dynamics converge to ϵ-equilibrium sets. We study the stochastically stable strategy profiles of logit response dynamics and show that they are contained in the set of strategy profiles that approximately maximize the potential function of a close potential game. Our results suggest that games that are close to a potential game inherit the dynamical properties (such as convergence to approximate equilibrium sets) of potential games. Additionally, since a close potential game to a given game can be found by solving a convex optimization problem [7], [8], this enables us to study dynamical properties of strategic form games by first identifying a nearby potential game to this game, and then studying the dynamical properties of the nearby potential game.

The framework presented in this paper opens up a number of interesting research directions. Among them, we mention the following:

a) Dynamics in "near" zero-sum and supermodular games: Dynamical properties of classes of games such as zero-sum games and supermodular games are also well established [10], [23]. If a game is close to a zero-sum game or a supermodular game, does it still inherit some of the dynamical properties of the original game? A negative answer to this question implies that the results on dynamical properties of these classes of games are fragile, and their implications are not strong. Hence, it would be interesting to understand whether analogous results to the ones in this paper can be established for these classes of games.

b) Different update rules: Potential games admit tractable analysis also for other learning dynamics such as fictitious play. Another interesting research direction is to extend our results on continuity of limiting behavior in near-potential games to other commonly considered updated rules.

c) Guaranteeing desirable limiting behavior: Another promising direction is to use our understanding of simple update rules, such as better/best-response and logit response dynamics to design mechanisms that guarantee desirable limiting behavior, such as low efficiency loss and "fair" outcome. It is well known that equilibria in games can be very different in terms of such properties [24]. Hence, an interesting problem is to develop update rules that converge to desirable equilibria, or to find mechanisms that modify the underlying game in a way that the limiting behavior of dynamics is desirable. Simple pricing mechanisms can ensure convergence to such equilibria in some near-potential games [14]. Extending these results to general games is left for future work.

References

[1] D. Fudenberg and D. Levine, *The theory of learning in games*. MIT press, 1998.

[2] J. Jordan, "Three problems in learning mixed-strategy Nash equilibria," *Games and Economic Behavior*, vol. 5, no. 3, pp. 368–386, 1993.

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

[4] H. Young, *Strategic learning and its limits*. Oxford University Press, USA, 2004.

[5] C. Alós-Ferrer and N. Netzer, "The logit-response dynamics," *Games and Economic Behavior*, vol. 68, no. 2, pp. 413–427, 2010.

[6] J. Marden and J. Shamma, "Revisiting log-linear learning: Asynchrony, completeness and a payoff-based implementation," *Games and Economic Behavior*, submitted, 2008.

[7] O. Candogan, I. Menache, A. Ozdaglar, and P. A. Parrilo, "Flows and Decompositions of Games: Harmonic and Potential Games," *Arxiv preprint arXiv:1005.2405*, 2010.

[8] O. Candogan, A. Ozdaglar, and P. A. Parrilo, "A projection framework for near-potential games," in *Decision and Control (CDC), 2010 49th IEEE Conference on*, 2010, pp. 244 –249.

[9] D. Monderer and L. Shapley, "Fictitious play property for games with identical interests," *Journal of Economic Theory*, vol. 68, no. 1, pp. 258–265, 1996.

[10] J. Shamma and G. Arslan, "Unified convergence proofs of continuous-time fictitious play," *IEEE Transactions on Automatic Control*, vol. 49, no. 7, pp. 1137–1141, 2004.

[11] J. Hofbauer and W. Sandholm, "On the global convergence of stochastic fictitious play," *Econometrica*, pp. 2265–2294, 2002.

[12] L. Blume, "Population Games," *The economy as an evolving complex system II*, p. 425, 1997.

[13] J. Marden, G. Arslan, and J. Shamma, "Cooperative Control and Potential Games," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 39, no. 6, pp. 1393–1407, 2009.

[14] O. Candogan, I. Menache, A. Ozdaglar, and P. A. Parrilo, "Near-Optimal Power Control in Wireless Networks: A Potential Game Approach," *INFOCOM 2010. The 29th Conference on Computer Communications. IEEE*, 2010.

[15] G. Arslan, J. Marden, and J. Shamma, "Autonomous vehicle-target assignment: A game-theoretical formulation," *Journal of Dynamic Systems, Measurement, and Control*, vol. 129, p. 584, 2007.

[16] L. Shapley *et al.*, "Some topics in two-person games," *Advances in game theory*, vol. 52, pp. 1–29, 1964.

[17] J. Marden, G. Arslan, and J. Shamma, "Joint strategy fictitious play with inertia for potential games," *IEEE Transactions on Automatic Control*, vol. 54, no. 2, pp. 208–220, 2009.

[18] O. Candogan, A. Ozdaglar, and P. Parrilo, "Learning in near-potential games," LIDS, MIT, 77 Massachusetts Avenue, Room 32-D608, Tech. Rep. 2867, April 2011.

[19] M. Freidlin and A. Wentzell, *Random perturbations of dynamical systems*. Springer Verlag, 1998.

[20] H. Young, "The evolution of conventions," *Econometrica: Journal of the Econometric Society*, vol. 61, no. 1, pp. 57–84, 1993.

[21] M. Haviv and L. Van der Heyden, "Perturbation bounds for the stationary probabilities of a finite Markov chain," *Advances in Applied Probability*, vol. 16, no. 4, pp. 804–818, 1984.

[22] G. Cho and C. Meyer, "Comparison of perturbation bounds for the stationary distribution of a Markov chain," *Linear Algebra and its Applications*, vol. 335, no. 1-3, pp. 137–150, 2001.

[23] P. Milgrom and J. Roberts, "Rationalizability, learning, and equilibrium in games with strategic complementarities," *Econometrica: Journal of the Econometric Society*, vol. 58, no. 6, pp. 1255–1277, 1990.

[24] T. Roughgarden, *Selfish Routing and the Price of Anarchy*. MIT Press, 2005.