Tuning Rate of Strategy Revision in Population Games

Shinkyu Park

Abstract— We investigate a multi-agent decision problem in population games where each agent in a population makes a decision on strategy selection and revision to engage in repeated games with others. The strategy revision is subject to time delays which represent the time it takes for an agent revising its strategy needs to spend before it can adopt a new strategy and return back to the game. We discuss the effect of the time delays on long-term behavior of the agents' strategy revision. In particular, when the time delays are large, the strategy revision would exhibit oscillation and the agents spend substantial time in "transitioning" between different strategies, which prevents the agents from attaining the Nash equilibrium of the game. As a main contribution of the paper, we propose an algorithm that tunes the rate of the agents' strategy revision and show such tuning approach ensures convergence to the Nash equilibrium. We validate our analytical results using simulations.

I. INTRODUCTION

Consider a multi-agent decision problem where each agent selects a strategy to engage in repeated interactions with other agents. Agents receive payoffs as a function of their *strategy profile* – the distribution of their strategy selection – and revise their strategy selection based on the payoffs. Such problem setting is prevalent in many engineering applications [1]–[10] ranging from demand response in smart grids to task allocation in multi-robot systems research.

We adopt the population game formalism [11] to model the process of the agents' strategy revision and assess long-term behavior of the revision process. Different from conventional population game problems, in this work, we investigate the scenario where the strategy revision is subject to time delays for which every agent revising its strategy needs to spend a fixed amount of time for transitioning to a new strategy. As a case in point, in multi-robot task allocation applications [8], [10], [12] where the strategy revision corresponds to assignment of a new task to a robot, time delays in the strategy revision reflect the requirement for each robot to travel to a distant location to take on a new task.

When the agents' strategy revision is subject to time delays, as we discuss in the paper, their strategy profile would exhibit oscillation implying that a portion of the agent population persistently transitions between different strategies. In the task allocation applications, this means that the robots would spend substantial time on traveling between distant locations which is a disadvantage since during the traveling, they cannot carry out assigned tasks. The authors of [8], [12] concretely discuss such phenomenon observed in

their linear models, and empirical results presented in [10] illustrate the oscillation of the strategy profile in population game models.

As a main contribution, we propose an algorithm that tunes the rate of the agents' strategy revision to eliminate the oscillation of the strategy profile, and prove that the algorithm ensures convergence of the strategy profile to the Nash equilibrium. Of relevance to this work are [13], [14, Section 5.9] and references therein that investigate the scenarios where payoff mechanisms underlying population games are subject to time delays and hence the agents revise their strategy selection based on time-delayed payoffs. Also, [15] examines a relevant problem in which, when revising its strategy, each agent needs to take on a sequence of sub-tasks (sub-strategies) which would potentially cause time delays in the agent's strategy revision.

Our contributions are distinct from existing work in the following aspects: **(i)** unlike [13], [14, Section 5.9] and references therein, our problem formulation considers that the agents' strategy revision is subject to time delays, which is different from the delays in payoff mechanisms, and **(ii)** different from [8], [12], [15], our proposed approach ensures convergence of the strategy profile to the Nash equilibrium, which, in our problem formulation, requires that no agents are in transition or taking on any sub-strategies. Below we summarize the main results of the paper.

- We model and analyze the effect of time delays in the strategy revision using the population game formalism and discuss long-term behavior of the agents' strategy profile. In particular, our analysis explains how the rate of the agents' strategy revision affects the convergence of the strategy profile.
- Based on the analysis, we propose an algorithm for tuning the strategy revision rate. The algorithm judiciously decreases the revision rate and ensures the convergence of the strategy profile to the Nash equilibrium in *contractive* population games. We illustrate our analytical results using simulations.

The paper is organized as follows. In Section II, we introduce the population game framework and strategy revision model we adopt in this study. We also illustrate how the rate of the agents' strategy revision affects long-term behavior of their strategy profile. In Section III, we propose an algorithm that judiciously tunes the strategy revision rate and show that the proposed algorithm ensures the convergence of the strategy profile to the Nash equilibrium in the class of *contractive* population games.[1] We illustrate our analytical results using

---

[1]See Definition 1 for its formal definition.

simulations with a numerical example in Section IV. We conclude the paper with discussions and future plans in Section V.

Notation: We denote by $\mathbb{R}^n$ the set of n-dimensional real vectors and by $\mathbb{R}_+^n$ the set of (element-wise) non-negative n-dimensional real vectors. Throughout the paper, we adopt the Euclidean norm for matrices and vectors.

II. PROBLEM DESCRIPTION

We describe the foundation of the population game formalism [11] and explain how we adopt it in our multi-agent decision problem to study the effect of time delays in strategy revision.

A. Population Games and Evolutionary Dynamics

Consider that a population of decision-making agents are engaged in repeated strategic interactions with one another. Each agent has a finite set $\{1, \cdots, n\}$ of strategies available and can select one strategy at a time. Based on its strategy selection, the agent receives a payoff determined by a payoff mechanism underlying the interactions.

Adopting the conventional notation and formalism, we denote the state of the population by $x = (x_1, \cdots, x_n) \in \mathbb{R}_+^n$, where x_i represents the portion of the population selecting strategy i and x sums up to 1, i.e., $\sum_{i=1}^n x_i = 1$. A payoff vector $p = (p_1, \cdots, p_n) \in \mathbb{R}^n$ is determined by a continuous function $\mathcal{F}$ as $p = \mathcal{F}(x)$, where p_i is the payoff assigned to the agents selecting strategy i. Assuming that there are a large number of agents in the population, we can define the set of viable population states as $\mathbb{X} = \{z \in \mathbb{R}_+^n \mid \sum_{i=1}^n z_i = 1\}$.

Contractive population games are defined as follows.

Definition 1 (Contractive Population Game [16]): A population game $\mathcal{F}$ is called *contractive* if it holds that

$$(w - z)^T (\mathcal{F}(w) - \mathcal{F}(z)) \leq 0, \ \forall w, z \in \mathbb{X}. \quad (1)$$

Assumption 1: We assume that the differential map $D\mathcal{F}$ of $\mathcal{F}$ exists and is bounded: there is a positive constant $B_{D\mathcal{F}}$ satisfying $\|D\mathcal{F}(z)\|_2 \leq B_{D\mathcal{F}}, \ \forall z \in \mathbb{X}$.

Note that when $D\mathcal{F}$ exists, the requirement (1) can be cast as

$$\tilde{z}^T D\mathcal{F}(z)\tilde{z} \leq 0, \ \forall z \in \mathbb{X}, \tilde{z} \in T\mathbb{X}, \quad (2)$$

where $T\mathbb{X}$ is the tangent space of $\mathbb{X}$.

The Nash equilibrium of $\mathcal{F}$ is defined as follows.

Definition 2 (Nash Equilibrium): An element z^{NE} in $\mathbb{X}$ is called the *Nash equilibrium* of a population game $\mathcal{F}$ if it holds that

$$(z^{\text{NE}} - z)^T \mathcal{F}(z^{\text{NE}}) \geq 0, \ \forall z \in \mathbb{X}. \quad (3)$$

A population game $\mathcal{F}$ can have multiple Nash equilibria. We denote by $\mathbb{NE}(\mathcal{F})$ the set of Nash equilibria. We adopt the following example to illustrate our main results.[2]

---

[2] For a simple presentation of the paper, we use the RPS game for the illustration purpose.

Example 1: Consider the Rock-Paper-Scissors (RPS) game ($n = 3$) whose payoff function is given by

$$\mathcal{F}(x) = \begin{pmatrix} -ax_2 + bx_3 \\ bx_1 - ax_3 \\ -ax_1 + bx_2 \end{pmatrix}, \quad (4)$$

where a, b are positive constants. Note that (4) is a contractive game if $b \geq a$, and the Nash equilibrium of $\mathcal{F}$ is given by $(1/3, 1/3, 1/3)$ for any $a, b > 0$.

In our study, we consider that the agents are repeatedly engaged in a same game and given an opportunity, they can revise their strategy selection based on the Smith revision protocol, originally presented in [9] and defined as follows.[3]

$$\rho_{ij}(p) = \varrho[p_j - p_i]_+ := \begin{cases} \varrho(p_j - p_i) & \text{if } p_j \geq p_i \\ 0 & \text{otherwise} \end{cases}, \quad (5)$$

where ϱ is a positive constant satisfying[4]

$$\sum_{j=1}^n \varrho[p_j - p_i]_+ \leq 1 \text{ and } \varrho[p_j - p_i]_+ \leq 1. \quad (6)$$

According to (5), each agent switches its strategy selection i to j with probability $\varrho[p_j - p_i]_+$; otherwise, it stays with its current strategy selection i with probability $1 - \sum_{j=1}^n \varrho[p_j - p_i]_+$. Hence, higher the payoff associated with strategy j more likely the agent selects it. The condition (6) is required for such probabilistic strategy revision scheme to be well-defined.

The point in time at which the agents' strategy revision takes place is determined by identical and independent Poisson processes with parameter λ. In particular, each agent can revise its strategy selection at each jump time of an associated Poisson process. Since the parameter λ determines the rate of the strategy revision, we refer to it as the *strategy revision rate*.

Let $x(t) = (x_1(t), \cdots, x_n(t))$ be the population state at time instant t driven by the revision protocol (5), the Poisson processes, and payoffs determined by a population game $\mathcal{F}$. By same discussions presented in [11, Chapter 10], as the population size tends to infinity, the solution to the following state equation approximates the population state $x(t)$ with arbitrary accuracy.

$$\begin{aligned} \dot{x}_i(t) &= \mathcal{V}_i(x(t), p(t)) \\ &:= \lambda \big(\sum_{j=1}^n x_j(t)\varrho \left[p_i(t) - p_j(t)\right]_+ \\ &\qquad - x_i(t) \sum_{j=1}^n \varrho \left[p_j(t) - p_i(t)\right]_+ \big) \end{aligned} \quad (7)$$

Adopting the same naming convention as in [17], we refer to (7) as *Evolutionary Dynamics Model (EDM)*.

B. δ-Passivity Tools for Convergence Analysis

A focal research theme in population games is in establishing convergence of the population state $x(t)$, governed by (7), to the Nash equilibrium set $\mathbb{NE}(\mathcal{F})$ of an underlying

---

[3] Although the analysis we present in this paper can be extended to other class of strategy revision protocols. However, for a concise presentation of our main result, we adopt the Smith protocol.

[4] Such constant ϱ exists since $\mathcal{F}$ is a continuous function and $\mathbb{X}$ is a compact subset of $\mathbb{R}^n$.

population game $\mathcal{F}$. An earlier work [18] in economics literature uses Lyapunov stability theorems to establish the convergence for a class of EDMs in contractive population games. Later, the work of [19] introduces the notion of δ-passivity, originated from dynamical system theory [20], in population games and provides general tools for studying convergence properties of EDMs. Since then, there have been refined definitions and applications of δ-passivity proposed for convergence analysis in population games [17], [21], [22]. Aside from δ-passivity tools relevant to our work, there are other important works that present a different notion of passivity as a tool to study convergence in multi-agent games [23], [24].

We adopt the following definition of δ-passivity for (7) from [17].

Definition 3: EDM (7) is δ-*passive* if there is a continuously differentiable function $\mathcal{S} : \mathbb{X} \times \mathbb{R}^n \to \mathbb{R}_+$ for which

$$\mathcal{S}(x(t), p(t)) - \mathcal{S}(x(t_0), p(t_0)) \leq$$
$$\int_{t_0}^t \dot{x}^T(\tau) \dot{p}(\tau) \, \mathrm{d}\tau, \ \forall t \geq t_0 \geq 0 \quad (8)$$

holds for every payoff vector trajectory $p(t)$, $t \geq 0$. We refer to $\mathcal{S}$ as the δ-*storage function*.

According to [17], [19], EDM (7), which is constructed using the Smith revision protocol (5), is δ-passive and its δ-storage function is given by

$$\mathcal{S}(x, p) = \frac{\lambda}{2} \sum_{i=1}^n \sum_{j=1}^n x_i \varrho[p_j - p_i]_+^2. \quad (9)$$

We make the following two observations which are used to establish the main result of the paper: for all i, j in $\{1, \cdots, n\}$, it holds that

$$\lambda^{-1} \mathcal{S}(x, p) \geq \frac{1}{2} x_i \varrho[p_j - p_i]_+^2 \geq \frac{1}{2} x_i^2 \varrho[p_j - p_i]_+^2 \geq 0$$

from which we conclude that

$$\lambda^{-1} \mathcal{S}(x, p) \to 0 \implies x_i \varrho[p_j - p_i]_+ \to 0. \quad (10)$$

Also, by similar arguments used in the proof of Theorem 4.5 in [19], we can verify that

$$\nabla_x^T \mathcal{S}(x, p) \mathcal{V}(x, p) \leq 0, \ \forall x \in \mathbb{X}, p \in \mathbb{R}^n. \quad (11)$$

C. Analyzing the Effect of Time Delays in Strategy Revision

We consider that the agents' strategy revision is subject to time delays: **(i)** an agent revising its strategy is required to spend a constant time period d_{ji} to conclude its strategy revision from j to i, and **(ii)** during the strategy revision, the agent cannot participate in the game. As a case in point, in the task allocation applications, the robots that are in transition to distant locations to take on newly assigned tasks cannot carry out the tasks until they arrive at the new locations. To model this, we interpret $x_i(t)$ as the portion of the population that are engaged in the game with strategy selection i at time t, which *exclude* the agents transitioning to strategy i. Accordingly, the payoff vector $p(t) = \mathcal{F}(x(t))$ depends only on the agents who are participating in the game.

(a) Population state $x(t)$ (b) Population in transition $y(t)$
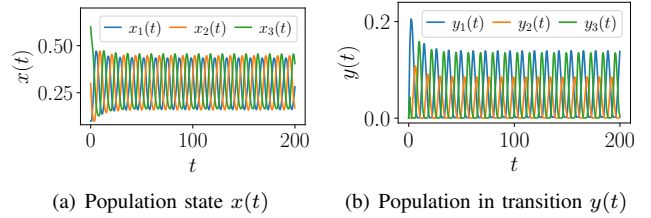
Fig. 1. Graphs depicting the trajectories of the (a) population state $x(t)$ and (b) population in transition $y(t)$ derived by EDM (15) ($d_{ij} = |i - j|, \varrho = 1/4, \lambda = 1$) in the RPS game (4) ($a = 1, b = 2$).

(a) Population state $x(t)$ (b) Population in transition $y(t)$
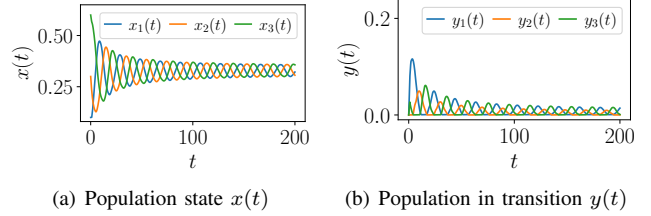
Fig. 2. Graphs depicting the trajectories of the (a) population state $x(t)$ and (b) population in transition $y(t)$ derived by EDM (15) ($d_{ij} = |i - j|, \varrho = 1/4, \lambda = 0.5$) in the RPS game (4) ($a = 1, b = 2$).

Note that in this problem setting, the population state $x(t)$ may not sum up to 1, i.e., $\sum_{i=1}^n x(t) \leq 1$. For this reason, we extend the definition of the population state space as

$$\mathbb{X}^{\text{ext}} = \left\{ z \in \mathbb{R}_+^n \mid \sum_{i=1}^n z_i \leq 1 \right\}. \quad (12)$$

Accordingly, we extend the domains of the payoff function $\mathcal{F} : \mathbb{X}^{\text{ext}} \to \mathbb{R}^n$, the vector field $\mathcal{V}_i : \mathbb{X}^{\text{ext}} \times \mathbb{R}^n \to \mathbb{R}$, and the δ-storage function $\mathcal{S} : \mathbb{X}^{\text{ext}} \times \mathbb{R}^n \to \mathbb{R}_+$. Also, we restate the boundedness condition of the map $D\mathcal{F}$ in Assumption 1 as $\max_{z \in \mathbb{X}^{\text{ext}}} \|D\mathcal{F}(z)\|_2 \leq B_{D\mathcal{F}}$.

Denote by $y_{ji}(t)$ the portion of the population that are in transition for the strategy revision from j to i at time t. Given a fixed strategy revision rate λ, the following state equation describes how $y_{ji}(t)$ changes over time.

$$\dot{y}_{ji}(t) = \lambda x_j(t) \varrho [p_i(t) - p_j(t)]_+, \ d_{ji} > t \geq 0 \quad (13a)$$
$$\dot{y}_{ji}(t) = \lambda \big(x_j(t) \varrho [p_i(t) - p_j(t)]_+$$
$$- x_j(t - d_{ji}) \varrho [p_i(t - d_{ji}) - p_j(t - d_{ji})]_+ \big), \ t \geq d_{ji} \quad (13b)$$

with $y_{ji}(0) = 0$. Note that in (13b), the first term $x_j(t) \varrho [p_i(t) - p_j(t)]_+$ denotes the rate at which the agents decide to revise their strategy selection from j to i and initiates transition from j to i at time t. Whereas the second term $x_j(t - d_{ji}) \varrho [p_i(t - d_{ji}) - p_j(t - d_{ji})]_+$ describes the rate at which the agents complete the transition and return back to the game after spending the time period d_{ji} in the transition.

Let us define $y_i(t) = \sum_{j=1}^n y_{ji}(t)$ which denotes the portion of the population transitioning to strategy i. Then, $x_i(t) + y_i(t)$ quantifies the portion of the population revising their strategies to i *including* the agents in transition. Using (7), we can describe the rate of change of $x_i(t) + y_i(t)$ by the following state equation.

$$\dot{x}_i(t) + \dot{y}_i(t) = \mathcal{V}_i(x(t), p(t)), \quad (14)$$

where the vector field $\mathcal{V}_i$ is defined in (7). Hence, from (13) and (14), we can derive the state equation for $x_i(t)$ as

$$\dot{x}_i(t) = \lambda\big(\textstyle\sum_{j=1}^{n} x_j(t-d_{ji})\varrho\,[p_i(t-d_{ji}) - p_j(t-d_{ji})]_+ - x_i(t)\sum_{j=1}^{n}\varrho\,[p_j(t) - p_i(t)]_+\big),\ t \ge d_{\max}, \quad (15)$$

where $d_{\max} = \max_{i,j\in\{1,\cdots,n\}} d_{ji}$. Note that different from EDM (7) in its original form, when the agents' strategy revision is subject to time delays, the revised model (15) depends not only on the current information $x(t), p(t)$ but also on the past information $x(t - d_{ji}), p(t - d_{ji})$ about the population game.

Figs. 1 and 2 depict the population state trajectories of (15) in the RPS game, explained in Example 1, with two different choices of the strategy revision rate ($\lambda = 0.5, 1$). We observe that in both cases, the trajectories oscillate around the Nash equilibrium, but with smaller λ, both the level of the oscillation and the portion of the population in transition decrease. Inspired by the observation, we investigate how the strategy revision rate can be tuned to achieve the convergence of the population state to the Nash equilibrium set.

We summarize the main problem of this paper as follows.

Problem 1: Design an algorithm to tune the strategy revision rate λ of (15) and to achieve the convergence of the population state $x(t)$ to the Nash equilibrium set $\mathbb{NE}(\mathcal{F})$ of an underlying contractive game $\mathcal{F}$:

$$\lim_{t\to\infty}\inf_{z\in\mathbb{NE}(\mathcal{F})} \|x(t) - z\|_2 = 0. \quad (16)$$

Note that according to Definition 2, since every Nash equilibrium z^{NE} in $\mathbb{NE}(\mathcal{F})$ belongs to $\mathbb{X}$, (16) requires the portion of the population in transition to vanish, i.e.,

$$\lim_{t\to\infty}\textstyle\sum_{i=1}^{n} y_i(t) = 1 - \lim_{t\to\infty}\sum_{i=1}^{n} x_i(t) = 0.$$

In this work, we focus on the case where the strategy revision rate λ is homogeneous for all the transition between strategies. Our analysis is rooted in the notion of δ-passivity for the Smith EDM (7). As illustrated in [25], adopting heterogeneous strategy revision rates would not ensure the δ-passivity of (7) and hence our proposed solution may not guarantee the convergence to the Nash equilibrium set. We leave as a future plan to further investigate such heterogeneous rate case.

III. TUNING STRATEGY REVISION RATE

We describe an algorithm that iteratively updates λ and in Section III-A, we explain how a sequence of the revision rates generated by the algorithm ensures the convergence of the population state to the Nash equilibrium set.

To explain the algorithm, we need the following lemma. The proof is given in Appendix of [26].

Lemma 1: Given a fixed strategy revision rate λ, the states $x(t)$ and $y(t)$ satisfy

$$\|\dot{x}(t)\|_2 \le N\lambda, \quad \|\dot{y}(t)\|_2 \le M\lambda^2, \ \forall t \ge d_{\max}, \quad (17)$$

where $d_{\max} = \max_{i,j\in\{1,\cdots,n\}} d_{ji}$ and N, M are positive constants.

The inequalities in (17) imply that as λ becomes smaller, i.e., the agents are less frequently revising their strategies, the rate of change of $y(t)$ decreases faster than that of $x(t)$. Therefore, the implication of Lemma 1 is that λ can be used as a *dial* to reduce the number of agents in transition while allowing them to have flexibility in switching between strategies.

In what follows, we discuss tuning of λ. Let us define $\bar{y}(t) = \lambda^{-2}y(t)$, $\bar{\mathcal{V}}(x(t), p(t)) = \lambda^{-1}\mathcal{V}(x(t), p(t))$, and $\bar{\mathcal{S}}(x(t), p(t)) = \lambda^{-1}\mathcal{S}(x(t), p(t))$. Using Lemma 1, we derive the rate of change of $\bar{\mathcal{S}}$ along a solution to (15) as follows.

$$\frac{\mathrm{d}}{\mathrm{d}t}\bar{\mathcal{S}}(x(t), p(t))$$
$$= \lambda\Big(\bar{\mathcal{V}}^T(x(t), p(t))D\mathcal{F}(x(t))\bar{\mathcal{V}}(x(t), p(t))$$
$$+ \nabla_x^T \bar{\mathcal{S}}(x(t), p(t))\bar{\mathcal{V}}(x(t), p(t))$$
$$- \lambda\dot{\bar{y}}^T(t)\big(D\mathcal{F}^T(x(t))\bar{\mathcal{V}}(x(t), p(t))$$
$$+ \nabla_x\bar{\mathcal{S}}(x(t), p(t))\big)\Big)$$
$$\le \lambda\Big(\nabla_x^T \bar{\mathcal{S}}(x(t), p(t))\bar{\mathcal{V}}(x(t), p(t))$$
$$+ \lambda M\big(B_{D\mathcal{F}}\big\|\bar{\mathcal{V}}(x(t), p(t))\big\|_2$$
$$+ \big\|\nabla_x\bar{\mathcal{S}}(x(t), p(t))\big\|_2\big)\Big), \quad (18)$$

where $p(t) = \mathcal{F}(x(t))$ and we use (2), (14), and the fact that $\nabla_p\mathcal{S}(x, p) = \mathcal{V}(x, p)$, derived in [27, Theorem III.3]. For notational convenience, we adopt

$$f(x, p) = M\big(B_{D\mathcal{F}}\big\|\bar{\mathcal{V}}(x, p)\big\|_2 + \big\|\nabla_x\bar{\mathcal{S}}(x, p)\big\|_2\big). \quad (19)$$

Note that $f(x, p)$ does not depend on the revision rate λ.

Given an initial revision rate $\lambda = \lambda_0$ at $t = 0$, the following algorithm is used to tune λ at each discrete time t_k and to generate a sequence $\{\lambda_k\}_{k=1}^{\infty}$.

Algorithm 1: Suppose the strategy revision rate was updated to λ_k at time t_k. At time instant $t \ge t_k + 2d_{\max}$, where $d_{\max} = \max_{i,j\in\{1,\cdots,n\}} d_{ji}$, update the revision rate as

$$\lambda = -\frac{1}{2}\frac{\nabla_x^T \bar{\mathcal{S}}(x(t), p(t))\bar{\mathcal{V}}(x(t), p(t))}{f(x(t), p(t))} \quad (20)$$

if the following two conditions hold:
(C1) $f(x(t), p(t)) > 0$
(C2) $\nabla_x^T \bar{\mathcal{S}}(x(t), p(t))\bar{\mathcal{V}}(x(t), p(t)) + \lambda_k \frac{f(x(t),p(t))}{1-\delta} \ge 0$
where δ is a constant satisfying $0 < \delta < 1/2$. Then, set $t_{k+1} = t$ and $\lambda_{k+1} = \lambda$.

Note that according to ((C1) and ((C2), the update rule (20) finds λ that minimizes the upper bound of $\frac{\mathrm{d}}{\mathrm{d}t}\bar{\mathcal{S}}(x(t), p(t))$ given as in the last inequality of (18).

Remark 1 (Decreasing strategy revision rates): Note that if we select λ according to (20), by ((C2), it satisfies

$$\lambda = -\frac{1}{2}\frac{\nabla_x^T \bar{\mathcal{S}}(x(t), p(t))\bar{\mathcal{V}}(x(t), p(t))}{f(x(t), p(t))} \le \frac{1}{2}\frac{\lambda_k}{1-\delta} < \lambda_k, \quad (21)$$

where the last inequality holds since $0 < \delta < 1/2$. Therefore, Algorithm 1 generates a decreasing sequence $\{\lambda_k\}_{k=1}^\infty$ of the revision rates satisfying $\lim_{k\to\infty} \lambda_k = 0$. ∎

Remark 2 (Comment on the constraint $t \geq t_k + 2d_{max}$): Under Algorithm 1, since λ switches over time, the parameter λ becomes time-dependent and the state equation (13) for $y_{ji}(t)$ can be rewritten as

$$\dot{y}_{ji}(t) = \lambda(t)x_j(t)\varrho\,[p_i(t) - p_j(t)]_+ \\ - \lambda(t-d_{ji})x_j(t-d_{ji})\varrho\,[p_i(t-d_{ji})-p_j(t-d_{ji})]_+ \quad (22)$$

for $t \geq d_{ji}$, where $\lambda(t)$ is the revision rate at time t defined as $\lambda(t) = \lambda_k$ if $t \in [t_k, t_{k+1})$. With the new representation (22), the derivation of similar inequalities as in (17), which are key in establishing our main result, would be technically difficult. To circumvent such difficulty, we require the update time t to satisfy $t \geq t_k + 2d_{\max}$ so that $\lambda(t)$ and $\lambda(t-d_{ji})$ in (22) satisfy $\lambda(t) = \lambda(t-d_{ji}) = \lambda_k$ for $t \in [t_k+2d_{\max}, t_{k+1})$. ∎

Remark 3 (Existence of t satisfying ((C1) and ((C2)): There may be no time instant $t \geq t_k + 2d_{\max}$ satisfying ((C1) and ((C2) and no further update on λ, which we consider as the termination of the algorithm. In what follows, we show that the population state converges to the Nash equilibrium set even if the algorithm terminates after a finite number of λ-updates.

Case I: Suppose $f(x(t), p(t)) = 0$, $\forall t \geq t_k + 2d_{\max}$. Then, from (11) and (18), it holds that

$$\frac{\mathrm{d}}{\mathrm{d}t}\bar{\mathcal{S}}\left(x(t),p(t)\right) \leq \lambda_k \nabla_x^T \bar{\mathcal{S}}(x(t),p(t))\bar{\mathcal{V}}(x(t),p(t)) \leq 0.$$

Noting that

$$\nabla_x^T \bar{\mathcal{S}}(x(t),p(t))\bar{\mathcal{V}}(x(t),p(t)) = 0 \\ \iff \bar{\mathcal{S}}(x(t),p(t)) = 0,$$

we conclude $\lim_{t\to\infty} \bar{\mathcal{S}}(x(t),p(t)) = 0$.

Case II: Suppose the following inequality holds:

$$\nabla_x^T \bar{\mathcal{S}}(x(t),p(t))\bar{\mathcal{V}}(x(t),p(t)) + \lambda_k \frac{f(x(t),p(t))}{1-\delta} < 0, \\ \forall t \geq t_k + 2d_{\max}. \quad (23)$$

Then, from (11) and (18), it holds that

$$\frac{\mathrm{d}}{\mathrm{d}t}\bar{\mathcal{S}}\left(x(t),p(t)\right) \leq \delta\lambda_k \nabla_x^T \bar{\mathcal{S}}(x(t),p(t))\bar{\mathcal{V}}(x(t),p(t)) \leq 0.$$

By the same argument as in (**Case I**), we have $\lim_{t\to\infty} \bar{\mathcal{S}}(x(t),p(t)) = 0$.

For both (**Case I**) and (**Case II**), by (10), it holds that $\lim_{t\to\infty} x_i(t)\varrho[p_j(t) - p_i(t)]_+ = 0$, $\forall i,j \in \{1,\cdots,n\}$. Then, using Lemma 3 stated in Section III-A, we conclude $\lim_{t\to\infty} \inf_{z\in\mathbb{NE}(\mathcal{F})} \|x(t) - z\|_2 = 0$. ∎

A. Convergence Analysis

Recall that by Remark 3, even if Algorithm 1 terminates after a finite number of λ-updates, the population state still converges to the Nash equilibrium set. In this section, we consider the case where Algorithm 1 generates an infinite

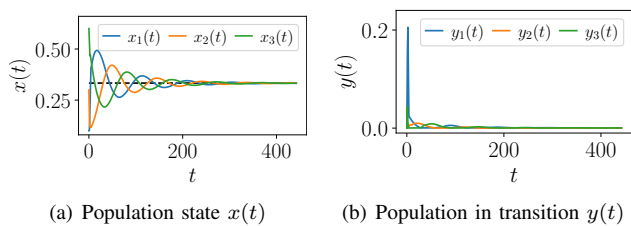(a) Population state $x(t)$ (b) Population in transition $y(t)$

Fig. 3. Graphs illustrating the trajectories of the (a) population state $x(t)$ and (b) population in transition $y(t)$ derived by EDM (15) ($d_{ij} = |i - j|, \varrho = 1/4$) in the RPS game (4) ($a = 1, b = 2$).

sequence $\{\lambda_k\}_{k=1}^\infty$ of strategy revision rates. The following theorem states the main result.

Theorem 1: Let $\{\lambda_k\}_{k=1}^\infty$ be a decreasing sequence of strategy revision rates determined by Algorithm 1. The population state converges to the Nash equilibrium set, i.e., it holds that

$$\lim_{t\to\infty} \inf_{z\in\mathbb{NE}(\mathcal{F})} \|x(t) - z\|_2 = 0. \quad (24)$$

Using the following two lemmas, we prove the theorem by showing that the δ-storage function $\mathcal{S}$ decreases at each time instant t_k of the λ-update (Lemma 2) which implies that the population state converges to the Nash equilibrium set $\mathbb{NE}(\mathcal{F})$ (Lemma 3). We provide the proofs of the theorem and lemmas in Appendix of [26].

Lemma 2: Given $\mathcal{V} = (\mathcal{V}_1, \cdots, \mathcal{V}_n)$ and $\mathcal{S}$ as in (7) and (9), respectively, define $\bar{\mathcal{V}} = \lambda^{-1}\mathcal{V}$ and $\bar{\mathcal{S}} = \lambda^{-1}\mathcal{S}$. Let K be a positive constant satisfying $K > (1 - \delta)^{-1}f(x,p)$ for all x in $\mathbb{X}^{\text{ext}}$ and $p = \mathcal{F}(x)$, where δ and $f(x,p)$ are defined in Algorithm 1. There is a decreasing function $\epsilon : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying

$$\nabla_x^T \bar{\mathcal{S}}(x,p)\bar{\mathcal{V}}(x,p) + \lambda K > 0 \implies \bar{\mathcal{S}}(x,p) < \epsilon(\lambda) \quad (25)$$

and $\lim_{\lambda\to 0} \epsilon(\lambda) = 0$.

Lemma 3: The following statement is true.

$$\lim_{t\to\infty} x_i(t)\varrho[p_j(t) - p_i(t)]_+ = 0, \ \forall i,j \in \{1,\cdots,n\} \\ \implies \lim_{t\to\infty} \inf_{z\in\mathbb{NE}(\mathcal{F})} \|x(t) - z\|_2 = 0. \quad (26)$$

IV. SIMULATIONS

Using Example 1, we illustrate our convergence result (Theorem 1) through simulations. To this end, we adopt the payoff function (4) with $a = 1, b = 2$ as follows.

$$\mathcal{F}(x) = \begin{pmatrix} -x_2 + 2x_3 \\ 2x_1 - x_3 \\ -x_1 + 2x_2 \end{pmatrix} \quad (27)$$

Note that $\mathcal{F}$ is contractive satisfying (2) and the upper bound of $D\mathcal{F}$ can be computed as $B_{D\mathcal{F}} \approx 2.65$.

We use Algorithm 1 to compute the sequence $\{\lambda_k\}_{k=1}^\infty$ for EDM (15) with the initial revision rate $\lambda_0 = 1$. Fig. 3 depicts the trajectories for both $x(t)$ and $y(t)$. Contrast to the fixed revision rate cases illustrated in Figs. 1 and 2, through tuning the revision rate using Algorithm 1, we observe that the population state $x(t)$ converges to the Nash equilibrium $(1/3, 1/3, 1/3)$ of (27). Also the portion $y(t)$ of the population in transition vanishes as the revision rate decreases.

V. Conclusions

We investigate a multi-agent decision problem in population games where the strategy revision of the agents is subject to time delays, which prevent the state of the population from converging to the Nash equilibrium set. We propose an algorithm that judiciously decreases the rate of the agents' strategy revision and prove that by tuning the revision rate, the population state converges to the Nash equilibrium set when the underlying population game is contractive.

As future directions, we plan to improve the proposed algorithm to attain faster convergence to the Nash equilibrium set. The current version of the algorithm is designed to minimize the upper bound of the rate of change of the δ-passive storage function; however this is mainly used to establish the convergence result and would not guarantee fast convergence. Also, validating the proposed approach in engineering applications, where the population would consist of a finite number of decision-making agents, will be another direction we plan to explore.

Appendix

A. Proof of Lemma 1

Using (15), we can derive a bound on $\dot{x}_i(t)$ as follows.

$$|\dot{x}_i(t)| \leq \lambda \Big(\sum_{j=1}^n x_j(t - d_{ji}) \varrho \left[p_i(t - d_{ji}) - p_j(t - d_{ji}) \right]_+$$
$$+ x_i(t) \sum_{j=1}^n \varrho \left[p_j(t) - p_i(t) \right]_+ \Big)$$
$$\leq \lambda(n - 1 + x_i(t)) \leq n\lambda, \tag{28}$$

where to establish the second inequality, we use (6) and the fact that

$$x_j(t - d_{ji}) \varrho \left[p_i(t - d_{ji}) - p_j(t - d_{ji}) \right]_+ \leq 1, \ j \neq i,$$
$$x_j(t - d_{ji}) \varrho \left[p_i(t - d_{ji}) - p_j(t - d_{ji}) \right]_+ = 0, \ j = i.$$

Using (28), we can establish Lipschitz continuity of $x(t)$: For any $d \geq 0$, it holds that

$$|x_j(t) - x_j(t - d)| \leq \max_{\tau \in [t-d, t]} |\dot{x}_j(\tau)| d \leq n d\lambda \tag{30}$$

and

$$\|x(t) - x(t - d)\|_2 \leq n^{1.5} d\lambda. \tag{31}$$

Consequently, using (31), the revision protocol $\varrho[p_i(t) - p_j(t)]_+$ satisfies

$$\left| \varrho \left[p_i(t) - p_j(t) \right]_+ - \varrho \left[p_i(t - d_{ji}) - p_j(t - d_{ji}) \right]_+ \right|$$
$$\leq \varrho \left| p_i(t) - p_j(t) - p_i(t - d_{ji}) + p_j(t - d_{ji}) \right|$$
$$\leq \varrho \left| p_i(t) - p_i(t - d_{ji}) \right| + \varrho \left| p_j(t) - p_j(t - d_{ji}) \right|$$
$$\leq \varrho \left(\|D\mathcal{F}_i\|_2 + \|D\mathcal{F}_j\|_2 \right) \|x(t) - x(t - d_{ji})\|_2$$
$$\leq 2\varrho B_{D\mathcal{F}} n^{1.5} d_{ji} \lambda, \tag{32}$$

where $\mathcal{F}_i$ is the i-th component of $\mathcal{F} = (\mathcal{F}_1, \cdots, \mathcal{F}_n)$, $\|D\mathcal{F}_i\|_2 = \max_{x \in \mathbb{X}^{\text{ext}}} \|D\mathcal{F}_i(x)\|_2 \leq B_{D\mathcal{F}}$, and we use

$$|p_i(t) - p_i(t - d_{ji})| = |\mathcal{F}_i(x(t)) - \mathcal{F}_i(x(t - d_{ji}))|$$
$$\leq \|D\mathcal{F}_i\|_2 \|x(t) - x(t - d_{ji})\|_2$$
$$\leq B_{D\mathcal{F}} n^{1.5} d_{ji} \lambda. \tag{33}$$

Therefore, using (30) and (32), we derive a bound on $\dot{y}(t)$ as follows.

$$|\dot{y}_i(t)|$$
$$\leq \lambda \sum_{j=1}^n \Big| x_j(t) \varrho \left[p_i(t) - p_j(t) \right]_+$$
$$\qquad - x_j(t - d_{ji}) \varrho \left[p_i(t - d_{ji}) - p_j(t - d_{ji}) \right]_+ \Big|$$
$$\leq \lambda \sum_{j=1}^n \Big| x_j(t) \varrho \left[p_i(t) - p_j(t) \right]_+$$
$$\qquad - x_j(t - d_{ji}) \varrho \left[p_i(t) - p_j(t) \right]_+ \Big|$$
$$+ \lambda \sum_{j=1}^n \Big| x_j(t - d_{ji}) \varrho \left[p_i(t) - p_j(t) \right]_+$$
$$\qquad - x_j(t - d_{ij}) \varrho \left[p_i(t - d_{ji}) - p_j(t - d_{ji}) \right]_+ \Big|$$
$$\leq \lambda \sum_{j=1}^n n d_{ji} \lambda \varrho \left[p_i(t) - p_j(t) \right]_+ + \lambda \sum_{j=1}^n 2\varrho B_{D\mathcal{F}} n^{1.5} d_{ji} \lambda$$
$$\leq (1 + 2\varrho B_{D\mathcal{F}} n^{0.5}) n d_i \lambda^2 \tag{34}$$

where $d_i = \sum_{j=1}^n d_{ji}$. Hence, from (28) and (34), we conclude that

$$\|\dot{x}(t)\|_2 \leq N\lambda, \quad \|\dot{y}(t)\|_2 \leq M\lambda^2$$

where $N = n^{1.5}$ and $M = (1 + 2\varrho B_{D\mathcal{F}} n^{0.5}) n \sqrt{\sum_{i=1}^n d_i^2}$. This completes the proof. ∎

B. Proof of Lemma 2

We first show that there is K such that $K > (1 - \delta)^{-1} f(x, p)$ holds for all x in $\mathbb{X}^{\text{ext}}$ and $p = \mathcal{F}(x)$, where $f(x, p)$ is defined in (19). Note that using (6), we have

$$\left| \bar{\mathcal{V}}_i(x, p) \right| = \left| \sum_{j=1}^n x_j \varrho \left[p_i - p_j \right]_+ - x_i \sum_{j=1}^n \varrho \left[p_j - p_i \right]_+ \right| \leq 1 \tag{35}$$

and

$$\left| \frac{\partial \bar{\mathcal{S}}}{\partial x_k}(x, p) \right| = \frac{1}{2} \sum_{j=1}^n \varrho [p_j - p_k]_+^2 \leq \frac{1}{2\varrho}. \tag{36}$$

Therefore, using the definition of $f(x, p)$ and (35), (36), we can conclude

$$(1 - \delta)^{-1} f(x, p)$$
$$= (1 - \delta)^{-1} M \left(B_{D\mathcal{F}} \left\| \bar{\mathcal{V}}(x, p) \right\|_2 + \left\| \nabla_x \bar{\mathcal{S}}(x, p) \right\|_2 \right)$$
$$\leq (1 - \delta)^{-1} M \left(B_{D\mathcal{F}} + \frac{1}{2\varrho} \right) n^{0.5} < K \tag{37}$$

if we choose $K > (1 - \delta)^{-1} M \left(B_{D\mathcal{F}} + \frac{1}{2\varrho} \right) n^{0.5}$.

To conclude the proof, we prove the contrapositive of (25):

$$\bar{\mathcal{S}}(x, p) \geq \epsilon(\lambda) \implies \nabla_x^T \bar{\mathcal{S}}(x, p) \bar{\mathcal{V}}(x, p) + K\lambda \leq 0 \tag{38}$$

Suppose $\bar{\mathcal{S}}(x, p) \geq \epsilon$ then according to (9), it holds that

$$\max_{1 \leq i, j \leq n} \frac{\varrho}{2} x_i [p_j - p_i]_+^2 \geq \frac{\epsilon}{n^2}. \tag{39}$$

Otherwise $\bar{\mathcal{S}}(x,p) < \epsilon$ which is a contradiction. Let i^*, j^* be indices for which the following holds.

$$\max_{1 \le i,j \le n} \frac{\varrho}{2} x_i [p_j - p_i]_+^2 = \frac{\varrho}{2} x_{i^*} [p_{j^*} - p_{i^*}]_+^2 \qquad (40)$$

from which we obtain

$$\frac{\epsilon}{n^2} \le \max_{1 \le i,j \le n} \frac{\varrho}{2} x_i [p_j - p_i]_+^2 = \frac{\varrho}{2} x_{i^*} [p_{j^*} - p_{i^*}]_+^2$$
$$\le \frac{1}{2} [p_{j^*} - p_{i^*}]_+, \qquad (41)$$

where we use (6) to establish the last inequality. Therefore, we have that

$$x_{i^*} [p_{j^*} - p_{i^*}]_+^2 \ge \frac{2\epsilon}{\varrho n^2} \qquad (42a)$$

$$[p_{j^*} - p_{i^*}]_+ \ge \frac{2\epsilon}{n^2}. \qquad (42b)$$

Also we can derive

$$\nabla_x^T \bar{\mathcal{S}}(x,p) \bar{\mathcal{V}}(x,p)$$
$$= \sum_{i=1}^n \sum_{j=1}^n \left(x_j \varrho [p_i - p_j]_+ - x_i \varrho [p_j - p_i]_+ \right) \frac{\varrho}{2} \sum_{k=1}^n [p_k - p_i]_+^2$$
$$= \frac{\varrho^2}{2} \sum_{i=1}^n \sum_{j=1}^n \left(x_j [p_i - p_j]_+ \sum_{k=1}^n \left([p_k - p_i]_+^2 - [p_k - p_j]_+^2 \right) \right)$$
$$\le \frac{\varrho^2}{2} \left(x_j [p_i - p_j]_+ \left([p_k - p_i]_+^2 - [p_k - p_j]_+^2 \right) \right) \qquad (43)$$

for any $i, j, k \in \{1, \cdots, n\}$. If we choose $i = k = j^*, j = i^*$, using (42), we can derive

$$\nabla_x^T \bar{\mathcal{S}}(x,p) \bar{\mathcal{V}}(x,p) \le -\frac{\varrho^2}{2} x_{i^*} [p_{j^*} - p_{i^*}]_+ [p_{j^*} - p_{i^*}]_+^2$$
$$\le -\frac{2\epsilon^2 \varrho}{n^4} \qquad (44)$$

Let us assign $\epsilon(\lambda) = \sqrt{\frac{n^4 K \lambda}{2\varrho}}$, then we can conclude

$$\nabla_x^T \bar{\mathcal{S}}(x,p) \bar{\mathcal{V}}(x,p) + K\lambda \le 0$$

and also we can verify that $\epsilon(\lambda) \to 0$ as $\lambda \to 0$. ∎

C. Proof of Lemma 3

Suppose $y_{ji}(0) = 0$, $\forall i,j \in \{1, \cdots, n\}$, i.e., there is no agent in transition for the strategy revision at the beginning of the game. Let t_2 be the time instant that Algorithm 1 updates the revision rate for the second time. For $t \ge t_2$, a solution $y_{ji}(t)$ to (22) satisfies

$$y_{ji}(t)$$
$$= \int_0^t \left(\lambda(\tau) x_j(\tau) \varrho [p_i(\tau) - p_j(\tau)]_+ \right.$$
$$\left. - \lambda(\tau - d_{ji}) x_j(\tau - d_{ji}) \varrho [p_i(\tau - d_{ji}) - p_j(\tau - d_{ji})]_+ \right) d\tau$$
$$= \int_{t - d_{ji}}^t \lambda(\tau) x_j(\tau) \varrho [p_i(\tau) - p_j(\tau)]_+ \, d\tau$$
$$- \int_{-d_{ji}}^0 \lambda(\tau) x_j(\tau) \varrho [p_i(\tau) - p_j(\tau)]_+ \, d\tau$$
$$\le \lambda_1 \int_{t - d_{ji}}^t x_j(\tau) \varrho [p_i(\tau) - p_j(\tau)]_+ \, d\tau, \qquad (45)$$

where we use the facts that

$$\int_{-d_{ji}}^0 \lambda(\tau) x_j(\tau) \varrho [p_i(\tau) - p_j(\tau)]_+ \, d\tau \ge 0,$$

$\lambda(\tau) = \lambda_k$ if $\tau \in [t_k, t_{k+1})$ with $t_{k+1} \ge t_k + 2d_{\max}$, and $\{\lambda_k\}_{k=1}^\infty$ is a decreasing sequence.

Suppose

$$\lim_{t \to \infty} x_i(t) \varrho [p_j(t) - p_i(t)]_+ = 0, \ \forall i,j \in \{1, \cdots, n\} \quad (46)$$

holds as in the statement of the lemma. Then, from (45), we have that

$$\lim_{t \to \infty} y_{ji}(t) = 0, \ \forall i,j \in \{1, \cdots, n\}. \qquad (47)$$

To complete the proof, using (46) and (47), we establish $\lim_{t \to \infty} \inf_{z \in \mathbb{NE}(\mathcal{F})} \|x(t) - z\|_2 = 0$. By contradiction, suppose that there is a converging sequence $\{x(t_k)\}_{k=1}^\infty$ for which its limit point x^* is not contained in $\mathbb{NE}(\mathcal{F})$. By (46) and (47), it holds that $\sum_{i=1}^n x_i^* = 1$ and

$$x_i^* > 0 \implies \mathcal{F}_i(x^*) = \max_{1 \le j \le n} \mathcal{F}_j(x^*), \ \forall i \in \{1, \cdots, n\}.$$

Hence, for any z in $\mathbb{X}$, we have that

$$(x^* - z)^T \mathcal{F}(x^*) = \max_{1 \le j \le n} \mathcal{F}_j(x^*) - z^T \mathcal{F}(x^*) \ge 0.$$

According to Definition 2, we conclude that x^* is the Nash equilibrium, which is a contradiction. This completes the proof. ∎

D. Proof of Theorem 1

Note that by ((C2) of Algorithm 1, at each update time t_{k+1}, it holds that

$$\nabla_x^T \bar{\mathcal{S}}(x(t_{k+1}), p(t_{k+1})) \bar{\mathcal{V}}(x(t_{k+1}), p(t_{k+1})) + \lambda_k K$$
$$> \nabla_x^T \bar{\mathcal{S}}(x(t_{k+1}), p(t_{k+1})) \bar{\mathcal{V}}(x(t_{k+1}), p(t_{k+1}))$$
$$+ \lambda_k f(x(t_{k+1}), p(t_{k+1})) \ge 0,$$

where K is a constant satisfying $K > (1 - \delta)^{-1} f(x,p) \ge f(x,p)$ as in the statement of Lemma 2. Invoking Lemma 2, there is a decreasing function $\epsilon : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying $\bar{\mathcal{S}}(x(t_{k+1}), p(t_{k+1})) < \epsilon(\lambda_k)$. In what follows, we show that there is a function $\bar{\epsilon} : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying

$$\bar{\mathcal{S}}(x(t), p(t)) < \bar{\epsilon}(\lambda_k), \ \forall t \in [t_{k+1}, t_{k+2}) \qquad (48)$$

for every positive integer k and $\lim_{k \to \infty} \bar{\epsilon}(\lambda_k) = 0$.

For $t \in [t_{k+1}, t_{k+1} + 2d_{\max})$, it holds that

$$\bar{\mathcal{S}}(x(t), p(t)) - \bar{\mathcal{S}}(x(t_{k+1}), p(t_{k+1}))$$
$$\le \max_{x \in \mathbb{X}^{\text{ext}}} \left\| \nabla_x \bar{\mathcal{S}}(x,p)|_{p=\mathcal{F}(x)} \right\|_2 \|x(t) - x(t_{k+1})\|_2$$
$$+ B_{D\mathcal{F}} \max_{x \in \mathbb{X}^{\text{ext}}} \left\| \nabla_p \bar{\mathcal{S}}(x,p)|_{p=\mathcal{F}(x)} \right\|_2 \|x(t) - x(t_{k+1})\|_2$$
$$\le 2d_{\max} L \max_{\tau \in [t_{k+1}, t_{k+1} + 2d_{\max}]} \|\dot{x}(\tau)\|_2, \qquad (49)$$

where L is a constant satisfying[5]

$$L = \max_{x \in \mathbb{X}^{\text{ext}}} \left\| \nabla_x \bar{\mathcal{S}}(x,p)|_{p=\mathcal{F}(x)} \right\|_2$$
$$+ B_{D\mathcal{F}} \max_{x \in \mathbb{X}^{\text{ext}}} \left\| \nabla_p \bar{\mathcal{S}}(x,p)|_{p=\mathcal{F}(x)} \right\|_2.$$

---

[5]Note that such constant L exists since $\mathcal{S}$ is continuously differentiable, $\mathcal{F}$ is continuous, and $\mathbb{X}^{\text{ext}}$ is compact.

Note that when the revision rate is repeatedly updated by Algorithm 1, λ becomes a time-dependent parameter and the state equation (13) for $y_{ji}(t)$ can be rewritten as in (22). Hence, using (14) and (22), it holds that for $t \in [t_{k+1}, t_{k+2})$,

$$
\begin{aligned}
|\dot{x}_i(t)| \\
&\leq \textstyle\sum_{j=1}^{n} \lambda(t-d_{ji}) x_j(t-d_{ji}) \varrho \left[p_i(t-d_{ji}) - p_j(t-d_{ji}) \right]_+ \\
&\quad + \textstyle\sum_{j=1}^{n} \lambda(t) x_i(t) \varrho \left[p_j(t) - p_i(t) \right]_+ \\
&\leq \textstyle\sum_{\substack{j=1 \\ j \neq i}}^{n} \lambda(t-d_{ji}) + \lambda_k \\
&\leq n\lambda_k, \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad (50)
\end{aligned}
$$

where we use (6), and the facts that $\{\lambda_k\}_{k=1}^{\infty}$ is a decreasing sequence and $t_{k+1} \geq t_k + 2d_{\max}$. Consequently, from (49) and (50), we can derive

$$
\bar{\mathcal{S}}(x(t), p(t)) \leq \bar{\mathcal{S}}(x(t_{k+1}), p(t_{k+1})) + 2d_{\max} Ln^{1.5}\lambda_k,
$$
$$
\forall t \in [t_{k+1}, t_{k+1} + 2d_{\max}). \quad (51)
$$

On the other hand, for $t \in [t_{k+1} + 2d_{\max}, t_{k+2})$, at least one of ((C1) and ((C2) of Algorithm 1 is violated and using (11) and (18), we can derive

$$
\frac{\mathrm{d}}{\mathrm{d}t} \bar{\mathcal{S}}(x(t), p(t)) \leq \lambda_{k+1} \big(\nabla_x^T \bar{\mathcal{S}}(x(t), p(t)) \bar{\mathcal{V}}(x(t), p(t))
$$
$$
+ \lambda_{k+1} f(x(t), p(t)) \big) \leq 0. \quad (52)
$$

Hence, the function $\bar{\mathcal{S}}(x(t), p(t))$ does not increase over $[t_{k+1} + 2d_{\max}, t_{k+2})$. Therefore, we conclude that for all $t \in [t_{k+1}, t_{k+2})$,

$$
\begin{aligned}
\bar{\mathcal{S}}(x(t), p(t)) &\leq \bar{\mathcal{S}}(x(t_{k+1}), p(t_{k+1})) + 2d_{\max} Ln^{1.5}\lambda_k \\
&< \epsilon(\lambda_k) + 2d_{\max} Ln^{1.5}\lambda_k = \bar{\epsilon}(\lambda_k). \quad (53)
\end{aligned}
$$

Since $\lim_{k \to \infty} \lambda_k = 0$, it holds that $\lim_{k \to \infty} \bar{\epsilon}(\lambda_k) = 0$ and we have that $\lim_{t \to \infty} \bar{\mathcal{S}}(x(t), p(t)) = 0$. By (10), it holds that $\lim_{t \to \infty} x_i(t) \varrho [p_j(t) - p_i(t)]_+ = 0, \ \forall i, j \in \{1, \cdots, n\}$, and using Lemma 3, we conclude that $\lim_{t \to \infty} \inf_{z \in \mathbb{NE}(\mathcal{F})} \|x(t) - z\|_2 = 0$. This completes the proof. ■

REFERENCES

[1] D. Liu, A. Hafid, and L. Khoukhi, "Population game based energy and time aware task offloading for large amounts of competing users," in *2018 IEEE Global Comm. Conf. (GLOBECOM)*, 2018, pp. 1–6.

[2] S. Moon, H. Kim, and Y. Yi, "Brute: Energy-efficient user association in cellular networks from population game perspective," *IEEE Trans. Wireless Communications*, vol. 15, no. 1, pp. 663–675, 2016.

[3] D. Lee and D. Kundur, "An evolutionary game approach to predict demand response from real-time pricing," in *2015 IEEE Electrical Power and Energy Conference (EPEC)*, 2015, pp. 197–202.

[4] H. Tembine, E. Altman, R. El-Azouzi, and Y. Hayel, "Evolutionary games in wireless networks," *IEEE Trans. Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 40, no. 3, pp. 634–646, 2010.

[5] P. Srikantha and D. Kundur, "Resilient distributed real-time demand response via population games," *IEEE Transactions on Smart Grid*, vol. 8, no. 6, pp. 2532–2543, 2017.

[6] A. Pashaie, L. Pavel, and C. J. Damaren, "A population game approach for dynamic resource allocation problems," *International Journal of Control*, vol. 90, no. 9, pp. 1957–1972, 2017.

[7] N. Quijano, C. Ocampo-Martinez, J. Barreiro-Gomez, G. Obando, A. Pantoja, and E. Mojica-Nava, "The role of population games and evolutionary dynamics in distributed control systems: The advantages of evolutionary game theory," *IEEE Control Systems Magazine*, vol. 37, no. 1, pp. 70–97, 2017.

[8] S. Berman, A. Halasz, M. A. Hsieh, and V. Kumar, "Optimized stochastic policies for task allocation in swarms of robots," *IEEE Transactions on Robotics*, vol. 25, no. 4, pp. 927–937, 2009.

[9] M. J. Smith, "The stability of a dynamic model of traffic assignment—an application of a method of Lyapunov," *Transportation Science*, vol. 18, no. 3, pp. 245–252, 1984.

[10] S. Park, Y. D. Zhong, and N. E. Leonard, "Multi-robot task allocation games in dynamically changing environments," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 2021, pp. 8678–8684.

[11] W. H. Sandholm, *Population Games and Evolutionary Dynamics*. MIT Press, 2011.

[12] T. W. Mather and M. A. Hsieh, "Macroscopic modeling of stochastic deployment policies with time delays for robot ensembles," *The International Journal of Robotics Research*, vol. 30, no. 5, pp. 590–600, 2011.

[13] S. Park and N. E. Leonard, "KL divergence regularized learning model for multi-agent decision making," in *2021 American Control Conference (ACC)*, 2021, pp. 4509–4514.

[14] M. J. Fox, "Distributed learning in large populations," Ph.D. dissertation, Georgia Institute of Technology, Aug. 2012.

[15] S. Kara, N. C. Martins, and M. Arcak, "Population games with erlang clocks: Convergence to nash equilibria for pairwise comparison dynamics," in *2022 IEEE 61st Conference on Decision and Control (CDC)*, 2022, pp. 7688–7695.

[16] W. H. Sandholm, "Chapter 13: Population games and deterministic evolutionary dynamics," ser. Handbook of Game Theory with Economic Applications, H. P. Young and S. Zamir, Eds. Elsevier, 2015, vol. 4, pp. 703–778.

[17] S. Park, N. C. Martins, and J. S. Shamma, "From population games to payoff dynamics models: A passivity-based approach," in *2019 IEEE 58th Conference on Decision and Control (CDC)*, 2019, pp. 6584–6601.

[18] J. Hofbauer and W. H. Sandholm, "Stable games and their dynamics," *Journal of Economic Theory*, vol. 144, no. 4, pp. 1665–1693.e4, 2009.

[19] M. J. Fox and J. S. Shamma, "Population games, stable games, and passivity," *Games*, vol. 4, pp. 561–583, Oct. 2013.

[20] J. C. Willems, "Dissipative dynamical systems part I: General theory," *Arch. Ration. Mech. Anal.*, vol. 45, no. 5, pp. 321–351, Jan. 1972.

[21] S. Park, N. C. Martins, and J. S. Shamma, "Payoff dynamics model and evolutionary dynamics model: Feedback and convergence to equilibria (arxiv:1903.02018)," arXiv.org, March 2019.

[22] M. Arcak and N. C. Martins, "Dissipativity tools for convergence to Nash equilibria in population games," *IEEE Transactions on Control of Network Systems*, vol. 8, no. 1, pp. 39–50, 2021.

[23] B. Gao and L. Pavel, "On passivity, reinforcement learning and higher-order learning in multi-agent finite games," *IEEE Transactions on Automatic Control*, pp. 1–1, 2020.

[24] L. Pavel, "Dissipativity theory in game theory: On the role of dissipativity and passivity in Nash equilibrium seeking," *IEEE Control Systems Magazine*, vol. 42, no. 3, pp. 150–164, 2022.

[25] S. Kara and N. C. Martins, "Pairwise comparison evolutionary dynamics with strategy-dependent revision rates: Stability and δ-passivity," *IEEE Transactions on Control of Network Systems*, pp. 1–12, 2023.

[26] S. Park, "Tuning rate of strategy revision in population games," 2022. [Online]. Available: https://arxiv.org/abs/2210.05472

[27] S. Park, J. S. Shamma, and N. C. Martins, "Passivity and evolutionary game dynamics," in *2018 IEEE Conference on Decision and Control (CDC)*, 2018, pp. 3553–3560.