Adaptive-Gain Control for Equilibrium Selection in the Logit Dynamics

Rory Gavin, Keith Paarporn, Mengbin Ye, Lorenzo Zino, and Ming Cao

Abstract— **We study the problem of controlling evolutionary game-theoretic dynamics when agents follow sophisticated learning rules, in particular the logit protocol. Much previous work focused on settings where agents are less sophisticated learners following imitative protocols that leads to the well-known replicator dynamic. Here, we consider adaptive control schemes for the logit dynamics with the objective of steering the population to a desired equilibrium by modifying the agents' payoff functions in a 2-action coordination game. Through the analysis of the controlled dynamics, we establish sufficient conditions for global convergence to the desired equilibrium. We find that the conditions to control the logit system have fewer requirements than those to control the replicator equation: Adaptive-gain controllers that are successful in performing their task in the logit system may fail in the replicator system. We then provide numerical simulations to illustrate and compare the amount of control effort needed to achieve the objective in the logit system versus the replicator system.**

I. Introduction

Evolutionary game theory (EGT) provides a powerful set of tools to model population behavior. Though originally developed in the context of biological systems [1], [2], EGT has been applied in many contexts, including traffic networks, control systems, and social behavior during epidemics [3]–[10]. A key element of EGT is the *revision protocol* [11], which governs how the agents dynamically update their strategies over time. Knowing the revision protocol, one can determine the *mean dynamics* induced by that revision protocol, and predict the population's emergent behavior.

As researchers increasingly utilize EGT to model populations of intelligent decision makers, it is also increasingly important to understand how different revision protocols affect the population dynamics. One example is the *pairwise proportional imitation* revision protocol, which assumes that decision makers interact with one another and imitate those with higher payoffs. This revision protocol induces the well-studied *replicator equation* [2], which has found wide-ranging applications [12]–[15]. Despite their widespread use, imitation protocols fail to capture the influence of information access and rationality in the agents' decision-making process. Human decision makers, for example, have access to many different sources of information, such as books, the internet, and news broadcasts, which they combine with reasoning and rational thought to make their decisions. This complexity necessitates a more sophisticated revision protocol that accounts for rational, informed decision making.

A revision protocol that addresses these shortcomings is the *logit learning* revision protocol [16]. Agents adopting this revision protocol rationally appraise all their options based on information about their payoffs. The logit learning revision protocol parametrizes the level of rationality wherewith the agents appraise their choices by using a parameter that captures the bounded rationality behind human decision making. A growing body of literature [17]–[23] demonstrates that the rational decision-making of logit learners results in interesting and sometimes counterintuitive outcomes. For instance, various studies highlight the presence of bifurcations (e.g. [19], [22]). In the case of two-strategy, normal form population games, a bifurcation in the number of logit fixed points occurs while increasing the agents' rationality [23].

While these works assume that the population is left to their own devices, governments and other authorities (to which we shall refer to as *controllers*) often seek to steer the population to a desired equilibrium. This problem is termed *equilibrium selection*. Many studies focus on open-loop control schemes, e.g., introducing a committed minority [24], [25] or permanently changing the structure of the payoff matrix [26]–[28]. However, these methods are not always feasible due to lacking the information needed to design such schemes, limited capacity to act on complex social systems, and ethical concerns. To overcome these limitations, closed-loop control schemes have been proposed [29]–[31]. In particular, a recent study proposed a flexible, closed-loop adaptive-gain control framework for the replicator equation [32], [33], able to solve the equilibrium selection problem with limited a priori information on the game.

Motivated by the success of this approach, we expand it to more sophisticated revision protocols. Specifically, we demonstrate that the same family of adaptive-gain controllers is able to achieve results similar to those for the replicator equation, for the broad class of coordination games, with fewer restrictions upon the design parameters. After rigorously demonstrating its convergence, we perform simulations to compare the adaptive-gain controller's performance on the logit dynamics and replicator equation. In summary, our work contributes to the growing literature on control of EGT, providing novel insights into the impact of different revision

R. Gavin and M. Cao are with the Faculty of Science and Engineering at the University of Groningen ({r.c.gavin, m.cao}@rug.nl). K. Paarporn is with the Department of Computer Science, University of Colorado, Colorado Springs CO, US (kpaarpor@uccs.edu). M. Ye is with the School of Computer and Mathematical Sciences, University of Adelaide, Australia (ben.ye@adelaide.edu.au). L. Zino is with the Department of Electronics and Telecommunications, Politecnico di Torino, Torino, Italy (lorenzo.zino@ polito.it). This work was supported in part by the Netherlands Organization for Scientific Research (NWO-Vici-19902). The work of L. Zino was carried out within the FAIR — Future Artificial Intelligence Research and received funding from the European Union Next-GenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) — MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.3 – D.D. 1555 11/10/2022, PE00000013). This manuscript reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them.

protocols on the controlled system.

This paper is organized as follows. Section II details the preliminaries of population games as well as evolutionary games using the mean dynamics. Section III defines the problem we aim to solve, i.e., find conditions upon the controller parameters such that the controller stabilizes the desired equilibrium. Section IV, specifically Theorem 1, specifies the conditions upon the controller parameters that stabilize the desired equilibrium. Section V compares the effects of certain controllers on the logit dynamics and the replicator equation, and finally Section VI briefly summarizes the key takeaways of this research.

II. Preliminaries

We denote by $\mathbb{R}$ the set of all real numbers, $\mathbb{R}^+$ the set of all positive real numbers, $\mathbb{R}^-$ the set of all negative real numbers, $\mathbb{R}_0^+$ and $\mathbb{R}_0^-$ denote $\mathbb{R}^+ \cup \{0\}$ and $\mathbb{R}^- \cup \{0\}$, respectively. Given a vector v, $[v]_i$ denotes its ith element. Given a scalar $a \in \mathbb{R}$, we use the notation $[a]_+ = \max\{a, 0\}$.

A. Two-Strategy Population Games

Consider a two-player, two-strategy, symmetric, normal-form population game with a population of unit mass [27]. The set of strategies available to the players is $\mathcal{S} = \{1, 2\}$, and the fraction of agents playing strategy $i \in \mathcal{S}$ is denoted as $x_i \in [0, 1]$. Define the *population state* $x \in [0, 1]$ as $x := x_1$. From this definition it follows that $x_2 = 1 - x$.

The *nominal* or *underlying payoff matrix* is given by

$$\hat{A} = \begin{bmatrix} R & S \\ T & P \end{bmatrix}, \tag{1}$$

with $R, S, T, P \in \mathbb{R}$. In other words, a 1-strategist receives a payoff of R or S if playing against a 1- or 2-strategist, respectively. Similarly, a 2-strategist receives a payoff of T or P if playing against a 1- or 2-strategist, respectively.

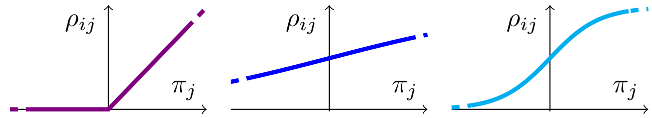
In population games, each agent plays with all the agents in the population and receives a total payoff that coincides with the average payoff obtained [11]. Hence, the total payoff associated with playing strategy i is

$$\hat{\pi}_i(x) = [\hat{A}[x, 1-x]^\top]_i. \tag{2}$$

A central concept in population games is the *Nash equilibrium* (NE), which is a population state at which no agent can improve their payoff by unilaterally switching strategies.

Definition 1. *A population state x^{ne} is a* NE *if for all $i \in \mathcal{S}$ such that $x_i^{ne} > 0$, it holds that $\pi_i(x^{ne}) \geq \pi_j(x^{ne})$, $\forall j \in \mathcal{S}$.*

Let us define the *payoff parameters* $\delta_{RT} := R - T$ and $\delta_{SP} := S - P$. The set of all pairs $(\delta_{SP}, \delta_{RT}) \in \mathbb{R}^2$ is called the *parameter space*, and each pair corresponds to a particular underlying game. Here, we focus on coordination games, where $\delta_{SP} < 0$ and $\delta_{RT} > 0$ (i.e. players obtain higher payoffs by choosing the same strategies), for which there are three NE: $x^{\mathrm{ne}} = 0$, 1, and $\frac{\delta_{SP}}{\delta_{SP} - \delta_{RT}}$. For more details on the parameter space, we refer to [13], [23], [32].



(a) Replicator (b) Logit, $\beta = 1$ (c) Logit, $\beta = 4$

Fig. 1: Illustration of (a) the imitation pairwise comparison protocol and the (b,c) logit protocol for different values of π_j, fixed $\pi_i = 0$ and $x_j = 0.5$. For the logit protocol, we set (b) $\beta = 1$ and (c) $\beta = 4$.

B. Mean Dynamics and Revision Protocol

While the NE concept describes plausible collective behavior, it does not offer any explanation for how agents could arrive at such states from a disequilibrium process. The framework of evolutionary dynamics provides a rich set of tools to address this [11]. At the heart of this framework is the *mean dynamics*, which describes the net rate of change of the fraction of agents playing any given strategy. In the context of two-strategy population games, it takes the form

$$\dot{x} = (1 - x)\rho_{21}(x) - x\rho_{12}(x). \tag{3}$$

Here, the functions $\rho_{ij}(x)$, $i, j \in \mathcal{S}$, are referred to as the *revision protocol*. It specifies the rate at which agents playing strategy i switch to strategy j in the population state x. Put more simply, it describes a learning rule that agents use to update their strategies over time.

Among the most commonly studied revision protocols is the *imitative pairwise comparison* protocol [11]:

$$\rho_{ij}(x) = x_j[\pi_j(x) - \pi_i(x)]_+, \tag{4}$$

illustrated in Fig. 1a. In plain words, an agent currently using strategy i encounters a randomly sampled agent in the population, and considers adopting the same strategy j if the payoff is higher than its current payoff. This induces the well-studied *replicator equation* [2], which is given by

$$\dot{x} = x(1 - x)\Delta\hat{\pi}(x), \tag{5}$$

where $\Delta\hat{\pi}(x) := \hat{\pi}_1(x) - \hat{\pi}_2(x) = (\delta_{RT} - \delta_{SP})x + \delta_{SP}$ is the payoff difference between strategies 1 and 2. We note that any NE of the underlying population game is a fixed point of the replicator equation, due to the fact that $\Delta\hat{\pi}(x^{\mathrm{ne}}) = 0$.

C. Logit Dynamics

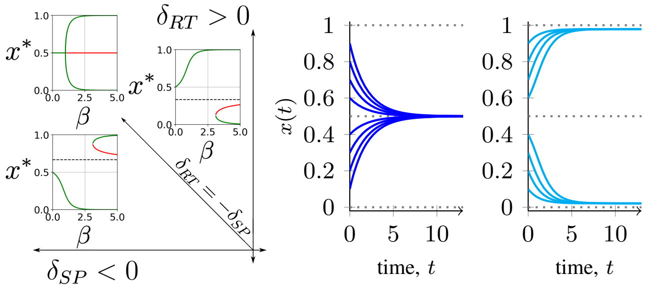
We consider a population of agents equipped with learning protocols that are more sophisticated than imitation. In particular, we study the *logit protocol* [16], that is,

$$\rho_{ij}(x) = \frac{e^{\beta\pi_j(x)}}{\sum_{k \in \mathcal{S}} e^{\beta\pi_k(x)}}, \tag{6}$$

where $\beta \in \mathbb{R}_0^+$ is the *rationality* parameter. The logit protocol reflects agents with a bounded level of rationality, as they select payoff-maximizing strategies with higher probabilities as β becomes larger, as illustrated in Figs. 1b and 1c. This revision protocol induces the *logit dynamics* given by

$$\dot{x} = \frac{1}{1 + e^{-\beta\Delta\hat{\pi}(x)}} - x, \tag{7}$$

which, as shown in a recent study [23], behaves differently than the replicator equation, as summarized in the following.



(a) Equilibria (b) $\beta = 1$ (c) $\beta = 4$

Fig. 2: Illustration of (a) the equilibria of the logit dynamics for different values of β for the games $(\delta_{SP}, \delta_{RT}) = (-2, 2)$, $(-1, 2)$, and $(-2, 1)$. Stable equilibria are in green, unstable ones in red. In (b,c), trajectories of the logit dynamics for a coordination game with $\delta_{SP} = -1$ and $\delta_{RT} = +1$; each curve with a different initial condition. Trajectories shown for (b) $\beta = 1 < \beta_r$ and (c) $\beta = 4 > \beta_r$. The gray dotted lines are the NE.

Proposition 1 (Theorem 4.2 in [23])**.** *Consider* (7) *for a coordination game* $(\delta_{SP} < 0, \delta_{RT} > 0)$*. Then,*

- *For $\beta < \beta_r$, there is one fixed point;*
- *For $\beta = \beta_r$, there are two fixed points; and*
- *For $\beta > \beta_r$, there are three fixed points,*

where β_r is a threshold that depends on the model parameters $(\delta_{SP}, \delta_{RT})$ and is explicitly defined in [23]. Moreover, fixed points have the expression

$$x_i^*(\beta) = \frac{1}{k} W_{r,i}(kr), \tag{8}$$

where $k = -\beta(\delta_{RT} - \delta_{SP})$, $r = e^{\beta \delta_{SP}}$, and $W_{r,i}(z)$ is the ith branch of the r*-Lambert function.*

It is important to note that while the underlying NE (0, 1, and $\frac{\delta_{SP}}{\delta_{SP} - \delta_{RT}}$) are associated with fixed points in the replicator equation, this is not the case in the logit dynamics. Indeed, it is shown in [23] that the set of logit fixed points converges to the set of NE as $\beta \to \infty$ (see Fig. 2a). In particular, based on the theoretical results in [23], we observe that, for a coordination game with $\beta > \beta_r$, two equilibria are stable: a smaller one that approaches the NE $x^* = 0$ as $\beta \to \infty$, and a larger one that approaches the NE $x^* = 1$ as $\beta \to \infty$. The initial condition determines which equilibrium is reached, as can be observed in Fig. 2c.

III. Problem Statement

As the logit dynamics for a coordination game only has multiple fixed points given certain conditions (Proposition 1), we make the following assumption to ensure that we study a meaningful equilibrium selection problem.

Assumption 1. *The game has $\delta_{SP} < 0$, $\delta_{RT} > 0$, and $\beta > \beta_r$, where β_r is defined as in [23].*

Therefore, Proposition 1 and Assumption 1 ensure that there are three equilibria given by (8) with $i = 0, -1, -2$. Specifically, as it is known [34] that $W_{r,0}(z) > W_{r,-1}(z) > W_{r,-2}(z)$, we can order the three equilibria as $0 < x_0^* < x_{-1}^* < x_{-2}^* < 1$. Without any loss in generality, we opt for casting our control problem with the objective of steering the system to the smallest equilibrium point x_0^*.

A recent study demonstrates that an adaptive-gain controller can force the replicator equation to a desired consensus for any coordination game [33]. This work shows that a modification of the same framework can steer the logit dynamics to a desired fixed point. Following [33], we assume that we can apply an additive gain term to the underlying payoff matrix. Thus, the controlled payoff matrix becomes

$$A(t) = \hat{A} + g(t) \cdot G, \tag{9}$$

where $g : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ is the *control gain* and $G \in \mathbb{R}^{2\times 2}$ is a constant *control matrix* which determines how the scalar gain is applied to the nominal payoffs. The payoff associated with playing strategy i is given by $\pi_i(x, g) = [A(t)[x, 1 - x]^\top]_i$. For the sake of simplifying the controller design, in this paper we focus on the scenario of one specific control matrix (see [33] for more details).

Assumption 2. *The control matrix is*

$$G = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \tag{10}$$

Furthermore, we require that the control gain dynamically governs itself as in [33]. Therefore, the ODE

$$\dot{g}(t) = g(t)\phi(x(t)) \tag{11}$$

governs the control gain $g(t)$, where the *adaptation rate* $\phi : [0, 1] \to \mathbb{R}$ should be designed to ensure convergence to the desired equilibrium.

The system of equations comprising the logit dynamics with payoff matrix (9) and the governing equation for the control gain (11) yields the *controlled logit dynamics*,

$$\dot{x} = \frac{1}{1 + e^{-\beta \Delta\pi(x,g)}} - x, \tag{12a}$$

$$\dot{g} = g\phi(x), \tag{12b}$$

where $\Delta\pi(x, g) = (\delta_{RT} - \delta_{SP})x + \delta_{SP} - (1 - x)g$ is the payoff difference between playing strategies 1 and 2 using control matrix G. In summary, from (9) and (11), the controller must design the adaptation rate ϕ. Hence, we formalize the problem of choosing ϕ that steers a coordination game to a desired fixed point as follows.

Problem 1. *Establish sufficient conditions on ϕ such that the solution $x(t)$ of* (12) *satisfies $\lim_{t\to\infty} x(t) = x_0^*$.*

Remark 1. *In the replicator equation, the smallest equilibrium coincides with the NE $x^* = 0$. For the logit dynamics, only $\lim_{\beta\to\infty} x_0^*(\beta) = 0$ [23] coincides with the NE. Hence, Problem 1 in this paper is the direct analog to establishing conditions to steer a replicator equation to a consensus state [33, Problem 1].*

IV. Main Results

We consider the scenario in Assumption 1, for which we want to solve Problem 1. We assume that we have some knowledge of the system.

Assumption 3. *We know a constant α such that $x_0^* < \alpha < x_{-1}^*$ and $x(0) \in [0, \alpha]$ implies that $x(t) \to x_0^*$ as $t \to \infty$.*

Within this scenario, we derive the following result.

Theorem 1. *For any initial value* $(x(0), g(0)) \in [0,1] \times \mathbb{R}^+$ *and under Assumption 1, the controlled logit dynamics* (12) *with control matrix* (10) *and any adaptation rate* $\phi(x)$ *satisfying: i)* $\phi(x) < 0$ *for* $x \in [0, \alpha)$ *and ii)* $\phi(x) > 0$ *for* $x \in (\alpha, 1]$ *converges to* $\lim_{t\to\infty} x(t) = x_0^*$ *and* $\lim_{t\to\infty} g(t) = 0$.

Proof. Given Assumptions 1 through 3, (10), and ϕ that satisfies the requirements above, all equilibria of (12) are listed below along with their local stability properties:

- $(x, g) = (x_0^*, 0)$ is a stable node;
- $(x, g) = (x_{-1}^*, 0)$ is an unstable node;
- $(x, g) = (x_{-2}^*, 0)$ is a saddle point; and
- $(x, g) = (\alpha, g_c)$ is outside of the state-space[1].

The expressions and computations for these fixed points follow from [23], and the local stability characteristics follow from the eigenvalues of the Jacobian of (12) at these fixed points. Details are omitted due to space limitations.

To prove the global stability of $(x_0^*, 0)$, we first define

$$\begin{aligned} \mathcal{A} &:= [0, x_{-1}^*) \times \mathbb{R}^+, & \mathcal{B} &:= [x_{-1}^*, x_{-2}^*) \times \mathbb{R}^+, \\ \mathcal{C} &:= \{x_{-2}^*\} \times \mathbb{R}^+, & \mathcal{D} &:= (x_{-2}, 1] \times \mathbb{R}^+. \end{aligned} \tag{13}$$

The proof of Theorem 1 is structured in three steps: i) Prove that for any initial condition in $\mathcal{A}$ (12) converges to $(x_0^*, 0)$; ii) Prove that any initial condition in $\mathcal{B}$ (12) enters $\mathcal{A}$ in finite time; and iii) Prove that any initial condition in $\mathcal{C} \cup \mathcal{D}$ (12) enters $\mathcal{B}$ in finite time.

Firstly, consider an initial condition $(x, g) \in \mathcal{A}$. Note that $(x_0^*, 0)$ is the only fixed point in the neighborhood $[0, \alpha) \times \mathbb{R}_0^+$ and that this fixed point is a stable node. For all $x \in [0, \alpha)$, we observe that the controlled logit dynamics in (12) has right-hand side always smaller than the uncontrolled one (being $g \geq 0$). Hence, if the uncontrolled dynamics (7) is such that $x(t) \to x_0^*$, which is the smallest equilibrium, this holds true also for the controlled dynamics. Furthermore, the inequality $\phi(x) < 0$ holds. Thus, $\dot{g} < 0$ holds too. As $x(t) \to x_0^*$ and $g(t)$ is monotonically decreasing, it necessarily approaches its infimum 0. In summary, any pair $(x(0), g(0)) \in [0, \alpha) \times \mathbb{R}_0^+$ converges to $(x_0^*, 0)$.

If $x(0) \in [\alpha, x_{-1}^*)$, then $g(t) \geq 0$ holds as $\phi(\alpha) = 0$ and $\phi(x) > 0$ for $x \in (\alpha, x_{-1}^*)$. Therefore, $\Delta\pi(x, g) \leq \Delta\pi(x, 0) = \Delta\hat{\pi}(x)$. Substituting this into (12a), one finds

$$\dot{x} \leq \frac{1}{1 + e^{-\beta \Delta\hat{\pi}(x)}} - x. \tag{14}$$

If the right side of this inequality is negative, then so is the left. Note that the right side of this inequality equals the vector field for the uncontrolled logit dynamics (7) which is negative for all $x \in (x_0^*, x_{-1}^*)$. Thus, the state x enters $[0, \alpha]$ at some point in time, after which it converges to $(x_0^*, 0)$. Therefore, all initial conditions in $\mathcal{A}$ converge to $(x_0^*, 0)$.

Secondly, consider an initial condition $(x, g) \in \mathcal{B}$. Define $\mu := \min_{x \in [x_{-1}^*, x_{-2}^*]} \phi(x)$. Note that $\mu > 0$ as $\phi(x) > 0$ for $x \in (\alpha, 1]$. Therefore, we can bound (12b) uniformly with respect to $x \in \mathcal{B}$ as $\dot{g} \geq \mu g$. Using Grönwall's inequality, one acquires the bound $g(t) \geq g(0)e^{\mu t}$ for any $x \in \mathcal{B} \cup \mathcal{C}$. Furthermore, $1 - x(t) \geq 1 - x_{-2}^* + \varepsilon$ for $x \in [x_{-1}^*, x_{-2}^* - \varepsilon]$. Therefore, for any constant $\varepsilon > 0$, we can bound

$$\Delta\pi(x, g) \leq \delta_{RT} - (1 - x_{-2}^* + \varepsilon) g(0) e^{\mu t}, \tag{15}$$

and as $x \geq x_{-1}^*$ one can also bound $\dot{x}$ as follows:

$$\dot{x} \leq \frac{1}{1 + e^{-\beta(\delta_{RT} - (1 - x_{-2}^* + \varepsilon) g(0) e^{\mu t})}} - x_{-1}^* \tag{16}$$

for any $x \in [x_{-1}^*, x_{-2}^* - \varepsilon]$. Therefore, there exists a time $\tau_{\mathcal{B}}$ such that $\dot{x} < 0$ for all $t > \tau_{\mathcal{B}}$, where

$$\tau_{\mathcal{B}} = \frac{1}{\mu} \ln \left(\frac{\beta \delta_{RT} + \ln\left(\frac{1 - x_{-1}^*}{x_{-1}^*}\right)}{\beta\left(\varepsilon - x_{-2}^* + 1\right) g(0)} \right) \tag{17}$$

Consider a generic trajectory starting in $\mathcal{B}$. If one assumes that this trajectory does not converge to $\mathcal{C}$, there necessarily exists a constant $\varepsilon > 0$ such that there exists $t_1 > \tau$ with $x(t_1) < x_{-2}^* - \varepsilon$. Letting $\tau = \tau_{\mathcal{B}}$, $\dot{x}(t_1) < 0$ for all $t_1 > \tau_{\mathcal{B}}$. Thus, for an infinitesimally small time increment dt, $x(t_1 + dt) < x(t_1)$ which also implies that $\dot{x}(t_1 + dt) < 0$. Therefore, all $x \in [x_{-1}^*, x_{-2}^* - \varepsilon]$ enter $\mathcal{A}$ after a certain amount of time. Since $(x_0^*, 0)$ is attractive in $\mathcal{A}$, $(x_0^*, 0)$ is attractive for all $x \in [x_{-1}^*, x_{-2}^* - \varepsilon]$ too.

Using a similar method for $(x(0), g(0)) \in \mathcal{C} \cup \mathcal{D}$ but omitting details due to space limitations, one again finds that there is a finite time past which the state enters $\mathcal{B}$. Combining these results, we conclude that any initial condition in $\mathcal{A} \cup \mathcal{B} \cup \mathcal{C} \cup \mathcal{D} = [0,1] \times \mathbb{R}^+$ ultimately converges to the unique stable node of the dynamics, $(x_0^*, 0)$, yielding the claim. ■

Theorem 1 demonstrates sufficient conditions for an adaptation rate ϕ such that $x(t)$ converges to the desired equilibrium x_0^*. Moreover, the gain eventually vanishes as the system approaches the desired equilibrium. Though similar, the conditions upon ϕ are less stringent for the logit dynamics than for the replicator equation in [32, Theorem 1]. In fact, in addition to the conditions stated in Theorem 1, the adaptation rate must also be larger than δ_{RT} in a neighborhood of the NE $x^* = 1$ for the replicator equation. This means that additional information about the game's payoff structure is needed to control imitative learners using (4). This fits with the interpretation of β as representative of the agents' rationality. Intuitively, it should be easier to control agents who act rationally than agents who blindly imitate others.

Remark 2. *In this paper, we focused on the problem of steering the system to the smallest equilibrium,* x_0^*. *However, with a similar line of reasoning and using the control matrix* $G = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, *one can easily design an adaptive-gain controller to reach the largest equilibrium* x_{-2}^*.

Though Theorem 1 demonstrates that it is easier to design an adaptation rate for the logit dynamics than for the replicator equation, it says nothing about their relative performance, which we study numerically in the next section.

[1]Since the right-hand side of (12a) is monotonically decreasing in g and $\dot{x} < 0$ for the uncontrolled dynamics ($g = 0$) for $x = \alpha$ by Assumption 3, then it necessarily holds that the zero of (12a) is attained for $g < 0$.

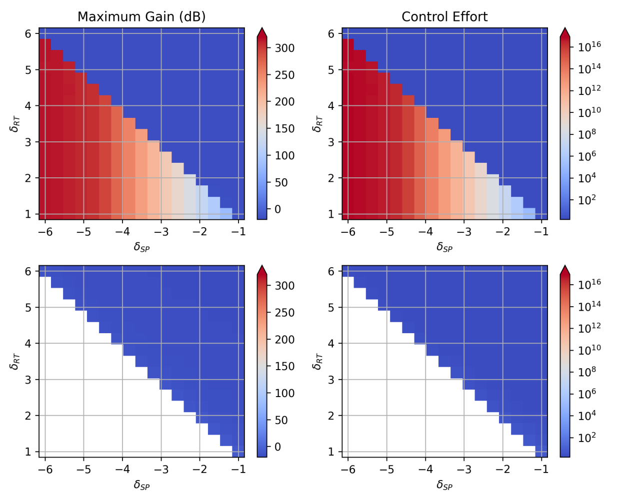


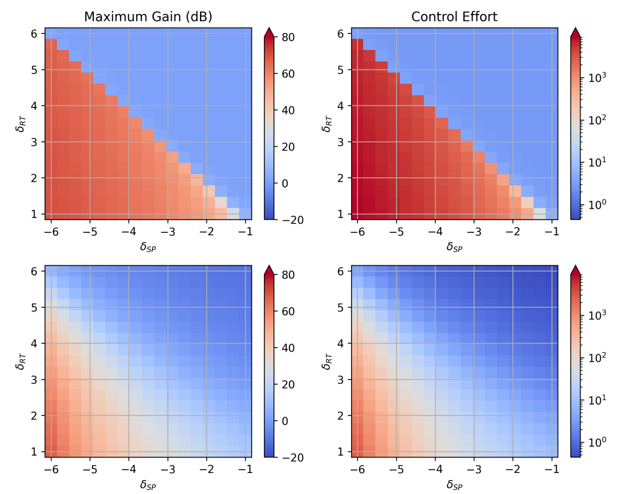
Fig. 3: Maximum gain (left panels) and control effort (right panels) for the logit dynamics (top row) and the replicator equation (bottom row) using $\phi_1(x)$ for $(\delta_{SP}, \delta_{RT}) \in [-6, -1] \times [1, 6]$. Common parameters are $\alpha = 0.1$, $\beta = 8$, $x(0) = 0.5$, and $g(0) = 0.1$. Uncolored regions show where the controller cannot steer the system to the desired equilibrium.

V. Numerical Performance Evaluation

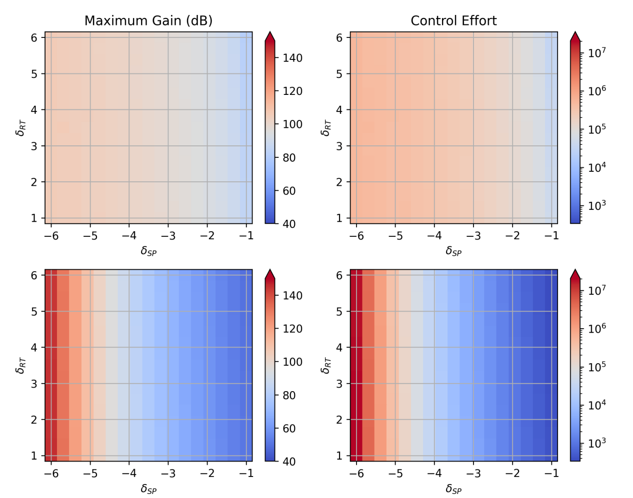
Assume that there is a population whose game and rationality satisfy Assumption 1. Furthermore, assume that the objective is to steer the population state to the system's smallest equilibrium (i.e., $x = x_0^*$ and $x = 0$ for the logit and replicator equation, respectively), while only knowing that the game played by the agents is given by $(\delta_{SP}, \delta_{RT}) \in (-3.5 \pm 2.5, 3.5 \pm 2.5)$ and that the agents' rationality is $\beta = 8$. Despite this uncertainty in the payoff structure, this policymaker knows a constant $\alpha = 0.1$ satisfying Assumption 3. Therefore, following Theorem 1, we set (10) and compare the controller's performance using the adaptation rates $\phi_1(x) := x - \alpha$ and $\phi_2(x) := 10\,(x - \alpha)$.

To quantify the performance of the adaptive controllers on the population of logit learners and imitators, the policymaker defines *maximum gain* as $g_{\max} = \max_{t \in [0,T]} g(t)$, and *control effort* as $J_g = \int_0^T g(t)dt$, where $[0, T]$ is the simulation time window. The policymaker also considers the average time it takes each system to converge to within a distance of 1×10^{-5} of the selected equilibrium over all games $(\delta_{SP}, \delta_{RT}) \in (-3.5 \pm 2.5, 3.5 \pm 2.5)$ when starting from a particular initial condition.

While the adaptation rate $\phi_1(x)$ fulfills the criteria of Theorem 1 for the logit dynamics, it fails to meet the requirements of [33, Theorem 1] for the replicator equation. Consequently, there are no theoretical guarantees that it steers the replicator equation to the desired equilibrium. Our simulations, reported in Fig. 3, confirm these analytical predictions. Using $\phi_1(x)$, the controller can only steer the replicator equation to the desired equilibrium when the initial population state is within the region of attraction of the uncontrolled replicator dynamics' lower fixed point. On the other hand, the use of $\phi_1(x)$ does indeed guarantee the convergence of the logit dynamics to $x = x_0^*$ for all $(x(0), g(0)) \in [0, 1] \times \mathbb{R}^+$. This exhibits that the constraints upon the adaptation rate for the logit dynamics are less stringent than those for the replicator equation. As one would expect, rational, informed agents adapt better to the changing game payoffs than the imitative learners.



(a) $x(0) = 0.5$, $g(0) = 0.1$



(b) $x(0) = 0.95$, $g(0) = 0.1$

Fig. 4: Maximum gain (left panels) and control effort (right panels) for the logit dynamics (top rows) and the replicator equation (bottom rows) using $\phi_2(x)$ for $(\delta_{SP}, \delta_{RT}) \in [-6, -1] \times [1, 6]$. Common parameters are $\alpha = 0.1$ and $\beta = 8$. Initial conditions are reported in the subcaptions.

Next, we perform a deeper analysis of the controller with an adaptation rate applicable to both systems and assess its performance with the maximum gain, control effort, and average convergence time metrics. Fig. 4 illustrates the maximum gains and control efforts of the adaptive controller using $\phi_2(x)$ for different initial conditions. We observe that the adaptive controller steers both the logit dynamics and the replicator dynamics to their respective lowest fixed points.

Interestingly, although the replicator equation does not boast a broad class of applicable adaptation rates, it is easier to control in many instances when using a suitable adaptation rate. The control effort is frequently less for the replicator equation than for the logit dynamics, as illustrated in Figure 4a. As illustrated in Figure 4b, the exceptions to this rule appear to occur when the system starts far from the desired equilibrium ($x(0) = 0.95$ in Fig. 4b) and when δ_{SP} is small. In these cases, the controlled replicator equation requires higher control effort and gain to steer the dynamics away from the NE $x^* = 1$. Furthermore, the average time it takes the controlled logit dynamics to converge is 1.25 and 1.19 times that of the controlled replicator equation for the initial conditions $(x(0), g(0)) = (0.5, 0.1)$ and $(0.95, 0.1)$, respectively. These findings highlight the difficulties in controlling rational decision makers over imitators: they need a compelling reason to adopt one strategy over the other.

In summary, these observations highlight the tradeoff between designing a controller for a population of logit learners (like a population of people) versus imitators (like a population of simple automata). Imitators tend to need less control effort at the cost of requiring more attention in the design of the adaptation rate, where logit learners frequently require more control effort while enjoying simpler constraints upon the adaptation rate.

VI. Conclusion

In this paper, we demonstrated that an adaptive-gain control scheme solves the equilibrium selection problem for coordination games with the logit dynamics. Furthermore, we showed with Theorem 1 that fewer constraints upon the design of the adaptive-gain controller are necessary for the logit dynamics than the replicator equation. However, simulations suggested that the logit learners are not necessarily easier to control, and in many cases more control effort, higher levels of gain, and more time are needed to steer the logit dynamics to the desire equilibrium.

These results pave the way for several future research directions. In particular, future work should focus on extending our approach to broader classes of games (such as dominant strategy games or anti-coordination games) and to different control matrices, going beyond Assumptions 1 and 2. Furthermore, this paper only established sufficient conditions for stability; future work should focus on the extent to which these conditions can be relaxed and the real-world applicability of these methods.

References

[1] P. D. Taylor and L. B. Jonker, "Evolutionary stable strategies and game dynamics," *Math. Biosci.*, vol. 40, no. 1-2, pp. 145–156, 1978.

[2] J. M. Smith, *Evolution and the Theory of Games*. Cambridge University Press, 1982.

[3] M. Smith, "A new dynamic traffic model and the existence and calculation of dynamic user equilibria on congested capacity-constrained road networks," *Transport. Res. B*, vol. 27, no. 1, pp. 49–63, 1993.

[4] 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 Syst.*, vol. 37, no. 1, pp. 70–97, 2017.

[5] J. S. Weitz, C. Eksin, K. Paarporn, S. P. Brown, and W. C. Ratcliff, "An oscillating tragedy of the commons in replicator dynamics with game-environment feedback," *Proc. Natl. Acad. Sci. USA*, vol. 113, no. 47, pp. E7518–E7525, 2016.

[6] K. Frieswijk, L. Zino, A. S. Morse, and M. Cao, "Modeling the co-evolution of climate impact and population behavior: A mean-field analysis," in *21st IFAC World Cong.*, vol. 56, no. 2, 2023, pp. 7381–6.

[7] ——, "A behavioural–environmental model to study the impact of climate change denial on environmental degradation," *Physica D*, vol. 476, p. 134648, 2025.

[8] L. Gong, W. Yao, J. Gao, and M. Cao, "Limit cycles analysis and control of evolutionary game dynamics with environmental feedback," *Automatica*, vol. 145, p. 110536, 2022.

[9] S. Leonardos, J. Sakos, C. Courcoubetis, and G. Piliouras, "Catastrophe by design in population games: A mechanism to destabilize inefficient locked-in technologies," *ACM Trans. Econ. Comput.*, vol. 11, no. 1-2, p. 1–36, 2023.

[10] A. Satapathi, N. K. Dhar, A. R. Hota, and V. Srivastava, "Coupled evolutionary behavioral and disease dynamics under reinfection risk," *IEEE Trans. Control Netw. Syst.*, vol. 11, no. 2, pp. 795–807, 2023.

[11] W. H. Sandholm, *Population games and evolutionary dynamics*. MIT Press, 2010.

[12] R. Cressman and Y. Tao, "The replicator equation and other game dynamics," *Proc. Natl. Acad. Sci. USA*, vol. 111, no. supplement_3, pp. 10 810–10 817, 2014.

[13] L. Zino, G. Como, and F. Fagnani, "On imitation dynamics in potential population games," in *56th IEEE Conf. Decis. Control*, 2017, p. 757–762.

[14] G. Como, F. Fagnani, and L. Zino, "Imitation dynamics in population games on community networks," *IEEE Trans. Control Netw. Syst.*, vol. 8, no. 1, pp. 65–76, 2021.

[15] L. Stella, W. Baar, and D. Bauso, "Lower network degrees promote cooperation in the prisoner's dilemma with environmental feedback," *IEEE Control Syst. Lett.*, vol. 6, pp. 2725–2730, 2022.

[16] L. E. Blume, "The statistical mechanics of strategic interaction," *Games Econ. Behav.*, vol. 5, no. 3, p. 387–424, Jul. 1993.

[17] J. R. Marden and J. S. Shamma, "Revisiting log-linear learning: Asynchrony, completeness and payoff-based implementation," *Games Econ. Behav.*, vol. 75, no. 2, p. 788–808, Jul. 2012.

[18] K. Tuyls, K. Verbeeck, and T. Lenaerts, "A selection-mutation model for q-learning in multi-agent systems," in *2nd Int. Jt. Conf. Auton. Agents Multiagent Syst.*, 2003, pp. 693–700.

[19] C. H. Hommes and M. I. Ochea, "Multiple equilibria and limit cycles in evolutionary games with logit dynamics," *Games Econ. Behav.*, vol. 74, no. 1, pp. 434–441, 2012.

[20] D. Bloembergen, K. Tuyls, D. Hennes, and M. Kaisers, "Evolutionary dynamics of multi-agent learning: A survey," *J. Artif. Intell. Res.*, vol. 53, pp. 659–697, 2015.

[21] L. Cianfanelli and G. Como, "On the stability of the logit dynamics in population games," *IEEE Trans. Autom. Control*, 2025.

[22] K. Paarporn, "The madness of people: rational learning in feedback-evolving games," in *2024 Eur. Control Conf.*, 2024, pp. 311–316.

[23] R. Gavin, M. Cao, and K. Paarporn, "An analysis of logit learning with the r-lambert function," in *63rd IEEE Conf. Decis. Control*, 2024, pp. 6774–6779.

[24] S. Morris, "Contagion," *Rev. Econ. Stud.*, vol. 67, pp. 57–78, 2000.

[25] G. Como, S. Durand, and F. Fagnani, "Optimal targeting in super-modular games," *IEEE Trans. Autom. Control*, vol. 67, no. 12, pp. 6366–6380, 2022.

[26] J. Riehl, P. Ramazi, and M. Cao, "Incentive-based control of asynchronous best-response dynamics on binary decision networks," *IEEE Trans. Control Netw. Syst.*, vol. 6, no. 2, pp. 727–736, 2018.

[27] ——, "A survey on the analysis and control of evolutionary matrix games," *Annu. Rev. Contr.*, vol. 45, pp. 87–106, 2018.

[28] Y. Zhu, Z. Zhang, C. Xia, and Z. Chen, "Equilibrium analysis and incentive-based control of the anticoordinating networked game dynamics," *Automatica*, vol. 147, p. 110707, 2023.

[29] S. Park, N. C. Martins, and J. S. Shamma, "From population games to payoff dynamics models: A passivity-based approach," in *58th IEEE Conf. Decis. Control*, 2019.

[30] M. Arcak and N. M. L. C. Martins, "Dissipativity tools for convergence to nash equilibria in population games," *IEEE Trans. Control Netw. Syst.*, vol. 8, no. 1, pp. 39–50, 2021.

[31] J. Martinez-Piazuelo, N. Quijano, and C. Ocampo-Martinez, "A payoff dynamics model for generalized nash equilibrium seeking in population games," *Automatica*, vol. 140, p. 110227, 2022.

[32] L. Zino, M. Ye, A. Rizzo, and G. C. Calafiore, "On adaptive-gain control of replicator dynamics in population games," in *62nd IEEE Conf. Decis. Control*, 2023, pp. 485–490.

[33] L. Zino, M. Ye, G. C. Calafiore, and A. Rizzo, "Equilibrium selection in replicator equations using adaptive-gain control," *IEEE Trans. Autom. Control*, p. 1–16, 2025.

[34] I. Mező and Á. Baricz, "On the generalization of the Lambert W function," *Trans. Am. Math. Soc.*, vol. 369, no. 11, pp. 7917–34, 2017.