On solving the closed-loop setpoint regulation problem for the replicator equation via nonlinear MPC

# On solving the closed-loop setpoint regulation problem for the replicator equation via nonlinear MPC

Giulia Brusadin, Michele Pagone, Lorenzo Zino, and Alessandro Rizzo

*Abstract*— **We address the setpoint regulation problem in evolutionary game-theoretic dynamics within population games, aiming to design an algorithm that guides the population toward a desired collective behavior —particularly, an equilibrium point. In detail, we focus on a discrete-time replicator equation, which models the collective behavior of a population engaged in two-player, two-action matrix games with every other member of the population. To tackle the problem, we develop an optimal control strategy that manipulates the payoff matrix in a closed-loop manner by adding a nonnegative gain to one of its entries. This strategy is intended to direct the population behavior toward the desired equilibrium. The control problem is resolved using nonlinear model predictive control, ensuring both closed-loop stability and recursive feasibility through appropriate selection of the terminal ingredients.**

## I. Introduction

Evolutionary game theory has emerged as a popular mathematical framework for studying social dynamics, since it can effectively capture the emergent behavior of a large population of individuals who engage repeatedly in strategic interactions [1]–[6]. In particular, in population games individuals play two-player games against all others and then revise their action according to a decision-making process. This process is captured at the population-level via nonlinear ordinary differential equations (ODEs) that describe the evolution of the fraction of adopters of each action [7], [8].

One of the most utilized options for such ODEs is the replicator equation [8]–[10], which guides the system toward a Nash equilibrium (NE) [8]. However, this NE may not represent the optimal outcome for societal interests. For example, in the prisoner's dilemma, the replicator equation predicts that the entire population would defect [10], even though cooperation yields a superior global outcome [4], [11]. Consequently, the development of techniques to steer the population toward a desired equilibrium —a challenge known as setpoint regulation— is critical [7], [8].

Considerable efforts have been made to tackle this problem [12]–[14]. However, most approaches depend on systematic alterations to individual behavioral mechanisms [15]–[17], which may be impractical, or on open-loop schemes, thus necessitating extensive a priori information about the game and the system's state [18]–[20]. Recently, closed-loop schemes have been employed to address the setpoint regulation problem in population games. These methods involve establishing a feedback loop between the controller and the population, allowing the controller to adjust the payoff functions. This dynamic revision can either be designed explicitly, as in the payoff dynamics model [21]–[23], or implicitly using an adaptive-gain approach [24], [25], which requires minimal a priori information. Nonetheless, while these studies focus on establishing conditions that ensure convergence to the desired equilibrium, there is limited research on designing such intervention policies in a cost-effective manner that minimizes control effort.

In this context, optimal control —among which model predictive control (MPC)— is emerging as a powerful tool to deal with complex interconnected dynamics [2], [26]–[28]. In particular, MPC's success stems from its inherent ability to explicitly handle linear and nonlinear systems and to provide optimal control commands for multi-variable systems in the presence of inputs, outputs, and state constraints [29]. Consequently, MPC applications have extended beyond conventional engineering and industrial applications. In recent years, the increasing complexity of social systems, encompassing areas such as opinion formation, information diffusion, and crowd behavior, has highlighted the potential of MPC as an effective technology for addressing future social challenges. MPC can assist policymakers in developing strategies to guide social behavior toward desired collective outcomes. However, the application of MPC to social dynamics is still a nascent area of research [30]–[32].

Here, we use a nonlinear MPC (NMPC) approach in order to address the setpoint regulation problem for population games in a cost-effective manner. In particular, we consider a replicator equation for a two-action two-player matrix game [13], and incorporate a control action that dynamically revises one entry of the payoff matrix, similar to [25]. For the MPC problem at hand, we provide both closed-loop stability and recursive feasibility certificates by suitably selecting the terminal ingredients of the controller. Specifically, we propose a stabilizing design in which the admissible invariant set of the terminal control is collapsed into a singleton. This approach bears the hallmark of trading the ease of design effort with a reduced domain of attraction of the controller. Finally, validate our methodology through numerical examples under different settings of the game-theoretic dynamics. In particular, we consider the cases of dominant strategy and anti-coordination games and we switch the control gain matrix in order to adapt the entry of the payoff matrix to steer the population toward the desired behavior. Our extensive numerical campaign validates the proposed control over different combinations of uncontrolled system equilibrium, desired equilibrium, and initial conditions.

The authors are with the Department of Electronics and Telecommunications, Politecnico di Torino, Turin, Italy (e-mails: {giulia.brusadin, michele.pagone, lorenzo.zino, alessandro.rizzo}@polito.it).

## II. Model and Problem Definition

*Notation:* We denote by $\mathbb{R}$, $\mathbb{R}_{\geq 0}$, $\mathbb{R}_{>0}$, $\mathbb{N}_{\geq 0}$, and $\mathbb{N}_{>0}$ the real, nonnegative real, strictly positive real, positive integer, and strictly positive integer numbers, respectively. The all-1 and all-0 vectors and the identity matrix are denoted by $\mathbf{1}$, $\mathbf{0}$, and $I$, respectively, with dimensions omitted when clear from the context. Given a square positive definite matrix $W \succ 0$, we denote $\|x\|_W = \sqrt{x^\top W x}$. A function $\delta : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is a $\mathcal{K}$-function, if it is continuous, strictly increasing and $\delta(0) = 0$. A function $\alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is a $\mathcal{K}_\infty$-function, if it is a $\mathcal{K}$-function and not bounded above.

### A. Population Games and Replicator Equation

We consider a (large) population where each individual engages in a two-player two-action matrix game with the others [13]. In this setting, each player can choose among two possible actions (1 and 2), and the payoffs that the player receives are described using a $2 \times 2$ matrix. Assuming that all players have the same payoff structure, we define

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \tag{1}$$

with $a, b, c, d \in \mathbb{R}_{\geq 0}$, where entry $a_{ij}$ corresponds to the reward that a player receives for choosing action $i$ against an opponent who is playing action $j$, for $i, j \in \{1, 2\}$. Based on 1, we can categorize games into three main classes [33].

**Definition 1.** *A two-player two-action matrix game with payoff matrix in Eq. (1) is said to be*

1) *a* coordination game*, if $d > b$ and $a > c$;*
2) *a* dominant-strategy game*, if $d > b$ and $a < c$, or $d < b$ and $a > c$;*
3) *an* anti-coordination games*, if $d < b$ and $a < c$.*

Since each individual plays the game against all the others, we introduce the state variable $x(t) \in [0, 1]$ that describes the fraction of adopters of 1 at time $t$ and we omit to explicitly write the dependence on $t$, when not necessary. Then, we write the average reward that one gets from all the games played as a function of the state. In particular, we obtain

$$\begin{bmatrix} r_1(x, A) \\ r_2(x, A) \end{bmatrix} = A \begin{bmatrix} x \\ 1 - x \end{bmatrix} = \begin{bmatrix} ax + b(1 - x) \\ cx + d(1 - x) \end{bmatrix}, \tag{2}$$

where $r_1(x, A)$ and $r_2(x, A)$ are the reward received for choosing action 1 and 2, respectively[1]. For more details on the derivation of these functions we refer to [25].

Players dynamically revise their strategy in order to improve their payoff. Assuming a large number of players, the collective behavior of such population in terms of the fraction of adopters of each action can be studied using the replicator equation [9], which captures the tendency to imitate peers who have higher rewards. In its original formulation, the replicator equation consists of the following nonlinear ordinary differential equation:

$$\begin{aligned} \dot{x} &= x(1 - x)(r_1(x, A) - r_2(x, A)) \\ &= x(1 - x)((a + d - b - c)x + b - d), \end{aligned} \tag{3}$$

[1]We assume that each player plays also against themselves, under the understanding that, since the population is large, such impact is negligible [8].

whose behavior is completely determined by the game class, as summarized in the following classical result from [8].

**Proposition 1.** *Consider the replicator equation in Eq. (3) and the classes of games from Definition 1.*

1) *If the game is a coordination game, then it has two stable equilibria, which are the consensus configurations. In particular, $x(t) \to 0$ if $x(0) < x_e$, and $x(t) \to 1$ if $x(0) > x_e$; with the unstable equilibrium*

$$x_e := \frac{d - b}{a + d - b - c}. \tag{4}$$

2) *If the game is a dominant-strategy game, then it has a unique stable equilibrium. In particular,*
   a) *If $d > b$ and $a < c$, then $x(t) \to 0$ for any $x(0) < 1$;*
   b) *$d < b$ and $a > c$, then $x(t) \to 1$ for any $x(0) > 0$;*
3) *If the game is an anti-coordination game, then it has a unique stable equilibrium, and $x(t) \to x_e$ from Eq. (4), for any $x(0) \in (0, 1)$.*

### B. Problem Formulation and Controlled Dynamics

In this study, we adopt the viewpoint of a policymaker whose goal is to guide the population toward a desired collective behavior, represented by a target state $x_s$. This desired state may differ from the state to which the system's uncontrolled dynamics would naturally converge, according Proposition 1. To achieve this objective, we build upon the adaptive-gain control scheme introduced in [25]. We propose that the policymaker can influence the system by adjusting the payoff matrix through the introduction of a nonnegative and bounded gain $g(t)$ to one of its elements. Implementing this gain addresses several shortcomings of open-loop control interventions by enhancing robustness and improving overall performance (for further details, refer to [25]). It is important to note that in our model, the gain is assumed to be nonnegative, which means it can represent dynamic incentives provided to the population to encourage the selection of a particular action. Hence, Eq. (1) reads

$$A = \widehat{A} + Gg(t) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} + Gg(t), \tag{5}$$

where $\widehat{A}$ is the nominal payoff matrix of the game from Eq. (1), $g(t) : \mathbb{R}_{\geq 0} \to [0, \bar{g}]$ is the gain, with $\bar{g} \in \mathbb{R}_{>0}$ being its upperbound, and $G$ is the control matrix that determines which entry of the payoff matrix is affected by the gain and is selected among the following possible matrices:

$$\begin{aligned} G^{(1)} &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, & G^{(2)} &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \\ G^{(3)} &= \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, & G^{(4)} &= \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \end{aligned} \tag{6}$$

In other words, matrices $G^{(1)}$ and $G^{(4)}$, through an increase of the diagonal entries, adaptively incentivize coordination on action 1 or 2, respectively; while $G^{(2)}$ and $G^{(3)}$, through an increase of the off-diagonal entry, provide an adaptive advantage for choosing one action against a player who plays the other action. A real-world example can involve a policymaker aiming to encourage the adoption of public transportation (action 1) over private car usage (action 2). In this scenario, the first type of control matrices (in this case, $G^{(1)}$, since the policymaker wants to incentivize action 1) encompasses policies such as offering group-discounted bus or train tickets. Conversely, the second type of matrices (here, specifically, $G^{(2)}$) concerns offering benefits to individuals who use public transportation instead of private cars (e.g., using priority lanes for buses to enhance the speed of public transportation compared to private cars).

The controlled continuous-time replicator equation reads

$$\begin{aligned}\dot{x} =& x(1-x)\big((a+d-b-c)x+b-d \\ &+(g_{11}-g_{21})g(t)x+(g_{12}-g_{22})g(t)(1-x)\big),\end{aligned} \tag{7}$$

where $g_{ij}$ is the $j$th entry of the $i$th row of matrix $G$. In [25], an adaptive-gain control scheme was proposed, through which $g(t)$ is dynamically updated based on the state of the system, ultimately establishing sufficient conditions under which the system converges to the desired equilibrium $x_s$. However, in practical applications, besides guaranteeing convergence to the desired collective behavior, it is key to achieve this goal in a cost-effective manner, i.e., minimizing the additional gain introduced in the system, avoiding unnecessary costs associated with the interventions. To the best of our knowledge, the optimal control of a setpoint regulation problem for a replicator equation is still an unexplored problem, which we address in this paper by means of a nonlinear MPC approach.

To facilitate the application of MPC, Eq. (7) is discretized, with sampling step $h > 0$. We denote by $x_k$ and $g_k$ the state and the gain at the $k$-th time step. Then, the discrete-time controlled replicator equation can be written as follows, using a standard Euler discretization method:

$$\begin{aligned}x_{k+1} =& x_k + hx_k(1-x_k)\big((a+d-b-c)x+b-d \\ &+(g_{11}-g_{21})g_k x_k+(g_{12}-g_{22})g_k(1-x_k)\big).\end{aligned} \tag{8}$$

If the sampling step is sufficiently small —i.e., Assumption 1 holds true— then we can prove that the discrete-time system in Eq. (8) is well-defined, as stated in Proposition 2.

**Assumption 1.** *Let $g_k \in [0,\bar{g}]$ for all $k \in \mathbb{N}_{\geq 0}$. Then, we assume $h \leq \frac{1}{\max\{a+b,c+d\}+\bar{g}}$.*

**Proposition 2.** *If Assumption 1 holds, then the domain $[0,1]$ is positively invariant for Eq. (8).*

*Proof.* We proceed by induction. At $k = 0$, the initial conditions guarantees that $x_k \in [0,1]$. Then, we assume $x_k \in [0,1]$. From Eq. (8), we observe that we can re-write the dynamics as $x_{k+1} \leq x_k + (1-x_k)x_k hM$, with $M = (a+b+\bar{g})$, where the bound is obtained by removing all the negative terms from the right-hand side of the equation and using that $x_k \in [0,1]$ to bound $x_k \leq 1$. Then, we observe that if $x_k hM < 1$, then necessarily $x_{k+1} \leq x_k + (1-x_k) \leq 1$. Hence $h \leq \frac{1}{(a+b+\bar{g})}$ implies $x_{k+1} \leq 1$. Similar, we bound $x_{k+1} \geq x_k[1-h(1-x_k)m]$, where $m = c+d+\bar{g}$. Hence, $hm \leq 1$, then necessarily the right-hand side of the bound is nonnegative, and $x_{k+1} \geq 0$, which yields the claim. □

Finally, under Assumption 1, the behavior of the uncontrolled discrete-time replicator equation in Eq. (8) with $g_k = 0$ for all $k \in \mathbb{N}_{\geq 0}$, coincides with the one of its continuous-time counterpart in discussed in Proposition 1.

## III. Nonlinear MPC with Stability Guarantees

Consider the model described in Eq. (8) in its compact and general discrete-time nonlinear formulation

$$x_{k+1} = f(x_k, u_k), \tag{9}$$

where, the control matrix $G$ is constant and selected from (6), $x_k \in \mathcal{X} \subseteq [0,1]$ and $u_k \doteq g_k \in \mathcal{U} \subseteq [0,\bar{g}]$ are the state and the input, respectively, at time $k \in \mathbb{N}_{\geq 0}$. Any desired equilibrium pair, to which the system in Eq. (9) is steered, is denoted as $(x_s, u_s)$. Additionally, both the state and the input are constrained within two closed and compact polytopic intervals, namely

$$x_k \in \mathcal{X} \doteq \{x \in \mathbb{R} : x_k \in [0,1],\ \forall k\}, \tag{10}$$

$$u_k \in \mathcal{U} \doteq \{u \in \mathbb{R} : u_k \in [0,\bar{g}],\ \forall k\}, \tag{11}$$

whose interiors are non-empty. Below, we specify some mild assumptions on the system under investigation that are useful to guarantee stability of the MPC design.

**Assumption 2.** *The function $f(\cdot) : \mathbb{R} \times \mathbb{R}^4 \to \mathbb{R}$ enjoys the following properties:*

*(i) $f(\cdot)$ is twice continuously differentiable;*
*(ii) $f(\cdot)$ is fully controllable;*
*(iii) a full information on the system's state is available at each sampling instant $k$;*
*(iv) each equilibrium pair $(x_s, u_s)$ is uniquely determined by $x_s = f(x_s, u_s)$.*
*(v) the equilibrium pair $(x_s, u_s)$ is chosen a-priori to be reachable within a finite number of sampling instants.*

**Remark 1.** *In our setting, the function $f(\cdot)$ from Eq. (8) satisfies the regularity assumptions required for Assumption 2.*

Our control task is to optimally regulate the controlled population game dynamics in Eq. (3) toward an equilibrium $(x_s, u_s)$. To this aim, consider the cost function

$$V_N(x, \mathbf{u}) = \sum\nolimits_{\iota=0}^{N-1} \ell(x_\iota, u_\iota) + V_f(x_N), \tag{12}$$

where $\mathbf{u} = [u_0, u_1, \ldots, u_{N-1}]^\top$ and $N$ is the prediction horizon. We define a quadratic stage cost:

$$\ell(x,u) \doteq Q\|x - x_s\|^2 + R\|u - u_s\|^2, \tag{13}$$

where the constant parameters $Q, R \in \mathbb{R}_{>0}$ establish a tradeoff between the cost associated with deviations from the desired equilibrium and the cost associated with the implementation of the control action. The terminal cost is set to penalize deviations from the desired equilibrium, as

$$V_f(\cdot) \doteq P\|x_N - x_s\|^2, \tag{14}$$

with constant parameter $P \in \mathbb{R}_{>0}$. To minimize the cost in Eq. (12), we use a nonlinear MPC approach that solves following optimization problem, denoted by $\mathcal{P}_N(\hat{x})$:

$$V_N^*(\hat{x}) = \min\nolimits_{\mathbf{u}} V_N(x, \mathbf{u}) \tag{15}$$

$$\text{s.t. } x_0 = \hat{x},$$
$$x_{\iota+1} = f(x_\iota, u_\iota),\ \iota = 1, \ldots, N-1,$$
$$x_\iota \in \mathcal{X}, \quad u_\iota \in \mathcal{U}, \quad x_N \in \mathcal{X}_f.$$

Here, for each $\iota \in \mathbb{N}_{\geq 0}$, $x_\iota = \phi(\iota; x, \mathbf{u})$ is the solution to Eq. (9) at time $\iota$ if the initial state is $x$ at $\iota = 0$ and the control is $\mathbf{u}$. Moreover, we define with $\mathcal{X}_N$ the feasible region of the optimal control problem. In the context of MPC, the implicit optimal control law, denoted as $\kappa_N(\hat{x})$, is typically implemented in a receding horizon manner. Consequently, the control action is defined as $\kappa_N(\hat{x}) = u_0^*$.

### A. Stabilizing the Nonlinear MPC Design

In the context of nonlinear MPC, closed-loop stability and recursive feasibility rely on a suitable choice of the terminal ingredients $V_f(\cdot)$ and $\mathcal{X}_f$. A popular approach in nonlinear MPC sets the terminal ingredients as $\mathcal{X}_f = \{x_s\}$ and $V_f(\cdot) = 0$. To ensure $V_N^*(\hat{x})$ to be a Lyapunov function, the standard additional *weak controllability* assumption is made [29].

**Assumption 3.** *There exists a $\mathcal{K}_\infty$ function $\alpha(\cdot)$ such that $V_N^*(\hat{x}) \leq \alpha(|x|)$, $\forall x \in \mathcal{X}_N$.*

Given that $\mathcal{X}_f$ lacks an interior, Assumption 3 ensures that $V_N^*(\hat{x})$ is bounded above by a $\mathcal{K}_\infty$ function, thus qualifying it as a Lyapunov function. In essence, Assumption 3 restricts the initial conditions to the set $\mathcal{X}_N$, where each element can be precisely guided to the reference within $N$ steps. It is noteworthy that Assumption 3 can be articulated under lenient conditions concerning stage cost and system dynamics (refer to, e.g., Proposition 2.38 in [29]).

**Remark 2.** *The proposed stabilizing design is widely recognized for its popularity and ease of implementation. However, this approach comes with the tradeoff of a notable reduction in the controller's domain of attraction. An alternative, albeit more complex, design involves determining the terminal components $V_f(\cdot)$ as a control Lyapunov function and $\mathcal{X}_f$ as the maximal control admissible invariant set. These design procedures necessitate system linearization around the equilibrium and the solution of a set of linear matrix inequalities conducted offline. Examples can be found in [34].*

## IV. Numerical Simulations

To numerically validate the proposed methodology, we present two distinct scenarios: one where the population engages in a dominant strategy game and another where they participate in an anti-coordination game. In both scenarios, the NMPC optimal control problem is solved using the CasAdi library [35]. The code employed for the simulations is available at [36]. In both scenarios, the weighting parameters for the NMPC are set as $Q = 140$ and $R = 1$. This configuration imposes a greater penalty for deviations from the desired equilibrium compared to the penalty applied for the control energy utilized.Since the terminal set $\mathcal{X}_f$ is a singleton, we set $P = 0$. Concerning the constraints set, $\mathcal{X} \doteq \{x_k \in \mathbb{R} : 0 \leq x_k \leq 1, \forall k\}$ and $\mathcal{U} \doteq \{u_k \in \mathbb{R} : 0 \leq u_k \leq 5 \text{ and } |u_k - u_{k+1}| \leq 1,\ \forall k\}$. Note that, limiting the increment/decrement of $u_k$ between two successive sampling steps reflects the impossibility —from the policymaker point of view— to change interventions too abruptly. Finally, the prediction horizon is $N = 10$. We conduct simulations of the system over a total time period of $T = 10$, using varying sampling steps that ensure compliance with Assumption 1. This approach guarantees that the discrete-time controlled replicator equation is properly defined, as per Proposition 2.

### A. Dominant Strategy Game

We consider a dominant-strategy game with different initial conditions and targeting different desired equilibria (namely $x_s = 0.4$ and $x_s = 0.8$). All the four matrices in (6) are considered. We simulate two different scenarios:

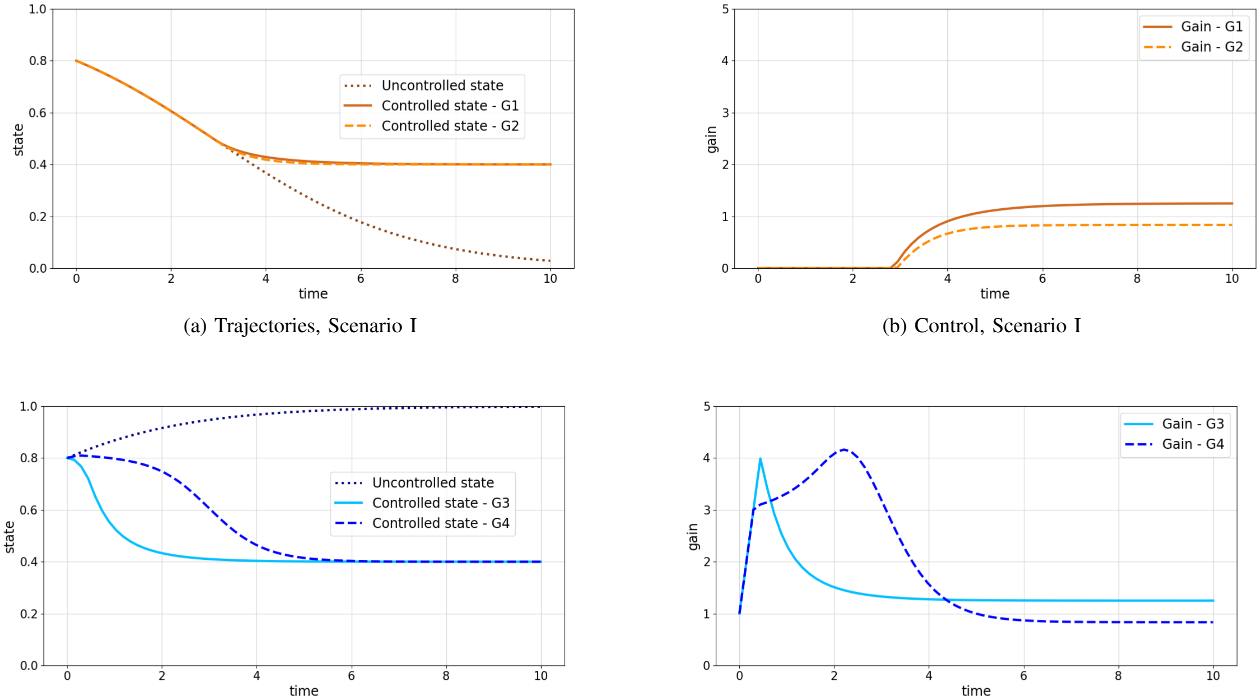
I With $a = 0.5$, $b = 1.5$, $c = 1$, and $d = 2$, where the dominant strategy is action 2, and thus, according to Proposition 1, the uncontrolled dynamics converges to the NE $x_e = 0$ (dotted curve in Fig. 1a). In this case, the controller should promote action 1 by utilizing either control matrix $G^{(1)}$ or $G^{(2)}$.

II With $a = 1$, $b = 2$, $c = 0.5$, and $d = 1.5$, where the dominant strategy is action 1, and Proposition 1 prescribes that the uncontrolled dynamics converges to $x_e = 1$ (dotted curve in Fig. 1c). In this case, the controller should promote action 2 by utilizing either control matrix $G^{(3)}$ or $G^{(4)}$.

In both scenarios, consistent with Assumption 1, we set the sampling step to be the slowest permissible value $h = 0.143$.

Figure 1 illustrates how the NMPC scheme is able to steer the system to the desired equilibrium.In particular, in 1a, we report the results for Scenario I, where 2 is the dominant action, with $x(0) = 0.8$ and $x_s = 0.4$. The figure shows that both $G^{(1)}$ and $G^{(2)}$ are able to ensure convergence to the desired equilibrium. From the temporal evolution of the control gains in Fig. 1b, we observe that the NMPC scheme is able to keep the control input limited by setting a nonzero gain only once the natural decreasing drift of the dynamics has led the trajectory close to the desired equilibrium, avoiding unnecessary costs for the policymaker. Comparing the two curves, we observe that inserting the gain in the diagonal term (i.e., using $G^{(1)}$) seems to slightly outperform controlling the off-diagonal term (i.e., $G^{(2)}$). Other scenarios are available at [36].

In 1c, we present the outcomes for Scenario II, where action 1 is predominant, maintaining the same initial conditions and desired equilibrium. In this scenario, both control matrices $G^{(3)}$ and $G^{(4)}$ successfully achieve the objective, with $G^{(3)}$ surpassing $G^{(4)}$ in terms of convergence speed and peak gain, as shown in Fig. 1d. However, $G^{(3)}$ ultimately stabilizes at a higher steady-state gain value. This observation implies that when policymakers aim to encourage behavior adopted by a minority, which might be individually disadvantageous (as seen in many social dilemmas), it may be more effective to provide relative advantages to those adopting the desired action (for instance, implementing priority lanes for public transportation to enhance its speed over private car usage).



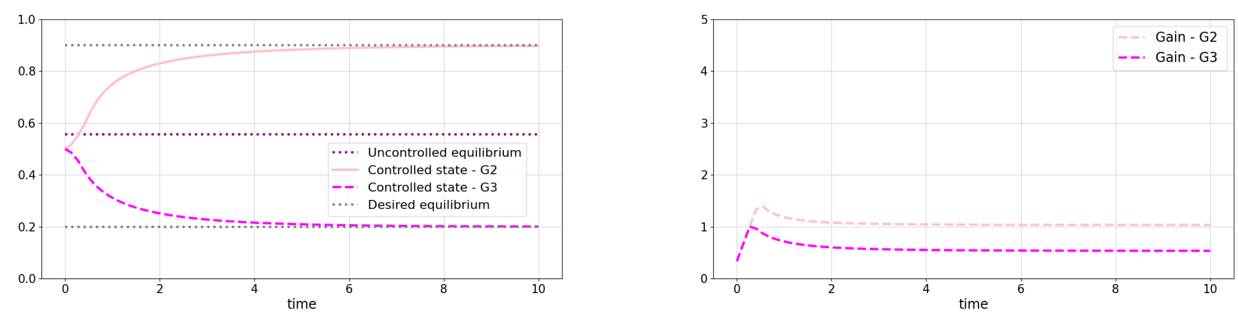

(a) Trajectories, Scenario I

(b) Control, Scenario I

(c) Trajectories, Scenario II

(d) Control, Scenario II

Fig. 1: Simulation results for a dominant strategy game. On panels (a,c), we report the uncontrolled state trajectory (dotted), and the controlled ones with two different control matrices. In (b,d), we report the corresponding control inputs. Common parameters are $x(0) = 0.8$, $x_s = 0.4$, and $h = 0.143$. In Scenario I: $a = 0.5$, $b = 1.5$, $c = 1$, and $d = 2$. In Scenario II: $a = 1$, $b = 2$, $c = 0.5$, and $d = 1.5$.




(a) Trajectories

(b) Control

Fig. 2: Simulation results for an anti-coordination game. On panels (a), we report the uncontrolled state trajectory (dotted), and the controlled ones with two different control matrices (depending on whether $x_s > x_e$ or the opposite inequality holds), highlighting the corresponding desired equilibria using a gray dotted line. In (b), we report the corresponding control inputs. Common parameters are $a = 1.1$, $b = c = 1.5$, $d = 1$, $x(0) = 0.5$, and $h = 0.132$.

### *B. Anti-coordination Game*

For anti-coordination games, the uncontrolled dynamics has a (mixed-strategy) NE $x_e$, as given in Equation Eq. (4), which is asymptotically stable and (almost) globally attractive, according to Proposition 1. As expected and discussed in [25], the choice of control matrix depends on whether the NE $x_e$ is greater or lesser than the desired one $x_s$. If $x_s < x_e$, action 2 should be promoted using $G^{(3)}$ or $G^{(4)}$; otherwise, action 1 should be promoted using $G^{(1)}$ or $G^{(2)}$. In both scenarios, we explore various initial conditions, including cases with $x_e < x_0 < x_s$ or $x_s < x_0 < x_e$.

In Fig. 2, we present outcomes from selected simulations, with additional results available in online [36]. Specifically, we examine two scenarios of the same game (parameters: $a = 1.1$, $b = c = 1.5$, and $d = 1$, yielding $x_e \approx 0.556$) using the same initial condition $x_0 = 0.5$, and target two distinct equilibria: $x_s = 0.2$ and $x_s = 0.9$. In the first scenario with $x_s < x_e$, control matrix $G^{(3)}$ is applied; in the second scenario where $x_s > x_e$, control matrix $G^{(2)}$ is utilized. The results confirm that, in both scenarios, the proposed control scheme guides the system to the desired equilibrium with minimal cost. From Fig. 2b, the controller's speed appears to correlate primarily with the distance between the initial state and the target equilibrium. The scenario with $x_s > x_e$, employing $G^{(2)}$, demonstrates slightly greater speed and, generally, lower gain, both in terms of peak and mean value.

## V. Conclusion

We developed a nonlinear MPC with stability guarantees to address the setpoint regulation problem of the replicator equation, focusing on dominant strategy and anti-coordination games. The structure of the control is designed along the lines of [24], [25], where convergence were established, but without considering performance. Here, we focused on enhancing the performance of the control scheme by formulating an optimal control problem to approach a desired equilibrium while minimizing the costs associated with the control actions. The experiments carried out validate the proposed control, offering insight on its performance and applicability to different operational conditions.

The preliminary results discussed in this paper lay the groundwork for several research directions. First, in the current methodology, the policymaker pre-selects the control matrix, while the MPC controller adaptively determines the gain. Future research should focus on enabling the control scheme to autonomously switch between different control matrices using tools from event-triggered control theory and switched systems. Second, our findings have been developed in the context of a uniform and well-mixed population and should be extended to heterogeneous and networked populations, and multi-action games. Third, game-theoretic dynamics are often employed to capture human behavior in complex cyber-physical-human systems [37]. The seamless integration of the proposed MPC approach within control schemes for these systems is a critical future direction.

## References


[1] J. R. Marden, G. Arslan, and J. S. Shamma, "Cooperative control and potential games," *IEEE Trans. Syst. Man Cybern. B*, vol. 39, no. 6, pp. 1393–1407, 2009.

[2] J. Barreiro-Gomez, G. Obando, and N. Quijano, "Distributed population dynamics: Optimization and control applications," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 47, no. 2, pp. 304–314, 2017.

[3] 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.

[4] 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.

[5] P. Scarabaggio, R. Carli, and M. Dotoli, "Noncooperative equilibrium-seeking in distributed energy systems under ac power flow nonlinear constraints," *IEEE Trans. Control Netw. Syst.*, vol. 9, no. 4, pp. 1731–1742, 2022.

[6] N. C. Martins, J. Certório, and R. J. La, "Epidemic population games and evolutionary dynamics," *Automatica*, vol. 153, p. 111016, 2023.

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

[8] W. H. Sandholm, *Population Games and Evolutionary Dynamics*. Cambridge MA, US: MIT Press, 2010.

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

[10] 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.

[11] R. M. May, "The evolution of cooperation," *Nature*, vol. 292, no. 5821, pp. 291–292, 1981.

[12] N. Quijano *et al.*, "The role of population games and evolutionary dynamics in distributed control systems: The advantages of evolutionary game theory," *IEEE Control Syst.*, vol. 37, pp. 70–97, 2017.

[13] J. Riehl, P. Ramazi, and M. Cao, "A survey on the analysis and control of evolutionary matrix games," *Annu. Rev. Contr.*, vol. 45, pp. 87–106, 2018.

[14] S. Grammatico, "Dynamic control of agents playing aggregative games with coupling constraints," *IEEE Trans. Autom. Control*, vol. 62, no. 9, pp. 4537–4548, 2017.

[15] L. Zino, M. Ye, and M. Cao, "On modeling social diffusion under the impact of dynamic norms," in *Proc. 60th IEEE Conf. Decis. Control*, 2021.

[16] S. Park, A. Bizyaeva, M. Kawakatsu, A. Franci, and N. E. Leonard, "Tuning cooperative behavior in games with nonlinear opinion dynamics," *IEEE Control Syst. Lett.*, vol. 6, pp. 2030–2035, 2022.

[17] F. Giardini, D. Vilone, A. Sánchez, and A. Antonioni, "Gossip and competitive altruism support cooperation in a public good game," *Philos. Trans. R. Soc. B*, vol. 376, no. 1838, 2021.

[18] 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.

[19] C. Eksin and K. Paarporn, "Control of learning in anticoordination network games," *IEEE Trans. Control Netw. Syst.*, vol. 7, no. 4, pp. 1823–1835, 2020.

[20] 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.

[21] S. Park, N. C. Martins, and J. S. Shamma, "From population games to payoff dynamics models: A passivity-based approach," in *Proc. 50th IEEE Conf. Decis. Control*, 2019.

[22] 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.

[23] 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.

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

[25] L. Zino, M. Ye, G. C. Calafiore, and A. Rizzo, "Equilibrium selection in replicator equations using adaptive-gain control," *IEEE Trans. Autom. Control*, vol. 70, no. 10, p. 6799–6814, 2025.

[26] A. Ferramosca, D. Limon, I. Alvarado, and E. F. Camacho, "Cooperative distributed MPC for tracking," *Automatica*, vol. 49, pp. 906 – 914, 2013.

[27] G. Gentil and A. Bhaya, "Opinion dynamic games under one step ahead optimal control," *IEEE Trans. Comput. Soc. Syst.*, vol. 11, no. 3, pp. 4202–4213, 2024.

[28] R. Raineri, M. Pagone, L. Zino, and A. Rizzo, "Mitigating vector-borne diseases on networks: A feasible and stable nonlinear mpc approach," *Nonlinear Anal.: Hybrid Syst.*, vol. 61, p. 101719, 2026.

[29] J. Rawlings, D. Mayne, and M. Diehl, *Model Predictive Control: Theory, Computation, and Design*. Nob Hill Publishing, 2017.

[30] B. Sprenger, G. De Pasquale, R. Soloperto, J. Lygeros, and F. Dörfler, "Control strategies for recommendation systems in social networks," *IEEE Control Syst. Lett.*, vol. 8, pp. 634–639, 2024.

[31] E. Villa, V. Breschi, and M. Tanelli, "Fair-MPC: A framework for just decision-making," *IEEE Trans. Autom. Control*, vol. 70, no. 9, pp. 6005–6020, 2025.

[32] L. Piccinin, V. Breschi, C. Ravazzi, F. Dabbene, and M. Tanelli, "Innovation diffusion dynamics toward long-term behavioral shifts," *IEEE Control Syst. Lett.*, vol. 9, pp. 661–666, 2025.

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

[34] M. Lazar and M. Tetteroo, "Computation of terminal costs and sets for discrete-time nonlinear MPC," in *Proc. 6th IFAC Conf. NMPC*, vol. 51, no. 20, 2018, pp. 141–146.

[35] J. A. E. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl, "CasADi – A software framework for nonlinear optimization and optimal control," *Math. Program. Comput.*, vol. 11, pp. 1–36, 2019.

[36] G. Brusadin, M. Pagone, L. Zino, and A. Rizzo. [Online]. Available: https://github.com/GiuliaBrusadin/MPC_set_point_regulation

[37] M. Cao, M. Ye, and L. Zino, "Control of networked cyber–physical–human systems," *Nat. Rev. Electr. Eng.*, vol. 3, no. 1, p. 32–45, 2025.