University of Groningen

A survey on the analysis and control of evolutionary matrix games

Riehl, James Robert; Ramazi, Pouria; Cao, Ming

Published in:
Annual Reviews in Control

DOI:
[10.1016/j.arcontrol.2018.04.010](https://doi.org/10.1016/j.arcontrol.2018.04.010)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Final author's version (accepted by publisher, after peer review)

Publication date:
2018

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Riehl, J. R., Ramazi, P., & Cao, M. (2018). A survey on the analysis and control of evolutionary matrix games. *Annual Reviews in Control*, 45, 87-106. <https://doi.org/10.1016/j.arcontrol.2018.04.010>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

A Survey on the Analysis and Control of Evolutionary Matrix Games

James Riehl^{a,*}, Pouria Ramazi^b, Ming Cao^{c,*}

^aDepartment of Electrical and Systems Engineering, Washington University in St. Louis, USA

^bDepartment of Mathematical and Statistical Sciences, University of Alberta, Canada

^cFaculty of Science and Engineering, University of Groningen, Netherlands

Abstract

In support of the growing interest in how to efficiently influence complex systems of interacting self-interested agents, we present this review of fundamental concepts, emerging research, and open problems related to the analysis and control of evolutionary matrix games, with particular emphasis on applications in social, economic, and biological networks.

Keywords: evolutionary games, population dynamics, equilibrium convergence, control strategies

Contents

1	Introduction	2
2	Evolutionary matrix games on networks	3
2.1	Networks, games, and payoffs	4
2.2	Discrete-time evolutionary dynamics	4
2.2.1	Best-response update rule	4
2.2.2	Imitation update rule	6
2.2.3	Proportional imitation update rule	6
2.3	Continuous time evolutionary dynamics	6
2.3.1	Derivation of the replicator dynamics	7
2.3.2	Best-response dynamics	9
3	Equilibrium convergence and stability	10
3.1	Infinite well-mixed populations (continuous-time continuous-state dynamics)	10
3.1.1	Equilibrium states	10
3.1.2	Lyapunov stable states	11
3.1.3	Asymptotically stable states	11
3.1.4	Globally asymptotically stable states	12
3.1.5	Convergence results	13
3.1.6	Folk Theorem	13
3.2	Infinite structured populations (continuous-time continuous-state dynamics)	14
3.3	Finite structured populations (discrete-time discrete-state dynamics)	14
3.3.1	Imitation	14
3.3.2	Best-response	15
4	Control strategies	16
4.1	Infinite well-mixed populations	16
4.2	Finite structured populations	17
4.2.1	Unique convergence for a class of coordinating agents	17
4.2.2	Direct strategy control	18
4.2.3	Incentive-based control	19

*Corresponding author

5	Open Problems	21
5.1	Infinite well-mixed populations	21
5.2	Finite structured populations	22
6	Concluding Remarks	22
Appendix A	2-player matrix games	24
Appendix B	Nash equilibrium	25
Appendix C	Categorization of 2 by 2 matrix games	25
Appendix D	Analyses of the replicator dynamics for 2x2 games	26
Appendix E	Evolutionary stability and neutral stability	26
Appendix F	Lotka-Volterra equation	28
Appendix G	Convergence proofs for asynchronous best-response dynamics	29

1. Introduction

Whether humans in a community, ants in a colony, or neurons in a brain, simple decisions or actions by interacting individuals can lead to complex and unpredictable outcomes in a population. The study of such systems typically presents a choice between micro- and macro-scale analysis. While there exist intricate micro-models of human decision processes, ant behaviors, and single neurons, assembling these high-dimensional components on a large scale most often results in models that are impenetrable to analysis, and therefore unlikely to reveal any useful properties of the collective dynamics. On the other hand, research on these systems at a broader scale, perhaps subject to substantial simplification of the agent-level dynamics, can help to characterize critical properties such as convergence, stability, controllability, robustness, and performance [1]. This helps to explain the recent and remarkable trend towards network-based analysis across various disciplines in engineering and the biological and social sciences, which has led to several important discoveries related to system dynamics on complex networks [2, 3, 4, 5]. For control scientists and engineers, these results facilitate the study of timely and challenging issues related to social, economic, and biological sciences from a control-theoretic perspective.

Evolutionary game theory has emerged as a vital toolset in the investigation of these topics. Originally proposed as a framework to study behaviors such as ritualized fighting in animals [6], it has since been widely adopted in various disciplines outside of biology. The primary innovation of evolutionary game theory is that rather than assuming high levels of rationality in individual choices, perhaps a questionable assumption even for humans, strategies and behaviors propagate through populations via dynamic processes. In the biological world, this propagation is manifested through *survival of the fittest* and reproductive processes, which are widely modeled using *population dynamics* [1, 7, 8]. Systems of first-order differential equations such as *replicator dynamics* (RD) provide an elegant and powerful means to investigate collective behaviors, assuming infinite and well-mixed populations. While these assumptions can lead to reasonable approximations for large, dense populations of organisms, in many other real-world networks, the structure and range of individual interactions plays a major role in the dynamics [9]. Fortunately, it is still possible to study replicator-like dynamics in populations connected by networks [10], and it turns out that certain models of imitation reduce exactly to RD in the limit of large networks [11]. Other seemingly more rational decision models such as best-response dynamics [12, 13, 14, 15, 16] also fit naturally into a network setting, as we will discuss in Section 2.

An extensive literature has emerged in the field of evolutionary games on networks, particularly regarding the question of how cooperation can evolve and persist under various conditions and in various population structures [17, 18, 19]. In this article, rather than survey these works, we will present only some fundamental results in classical evolutionary game dynamics, before discussing some recent developments in the areas of equilibrium convergence and control. Specifically, we set out to achieve three primary goals. First, we aim to introduce the powerful analytical tools of evolutionary game theory to control scientists and engineers not already familiar with the topic. Second, we provide a brief survey of some recent results in the analysis and control of evolutionary matrix games. Third, we discuss some current challenges and open problems in the field for the consideration of interested researchers.

We emphasize that, although game theory and evolutionary game theory are receiving increasing attention as design tools for implementing distributed optimization in industrial and technological systems [20, 21, 22], including water distribution [23, 24], wireless communication [25], optical networks [26], and transportation [27], we focus this survey specifically on the analysis and control of self-organized systems whose constituents are not necessarily subject to design. This stands apart from some ongoing research which aims at engineering the dynamics governing a population of programmable individuals, e.g., robots, in order to drive the state of the system to a desired state. We refer the reader to surveys such as [21, 28] for detailed discussions on this separate but complementary topic. In contrast, typical individuals we model in this paper such as humans, firms, animals, and neurons are clearly not programmable in the same sense, and even if they were, attempting to do so would likely raise ethical concerns. Rather, for the dynamics of these individuals, we take existing models proposed by biologists, sociologists, and economists, and perform convergence and stability analysis, which deepens our understanding of their collective behaviors. As the next step, by exploiting the limited freedom in modifying the dynamics of the individuals, such as providing incentives, we investigate attempts to guide such populations to desired states, beneficial to the overall group of individuals. The results, however, are not necessarily limited to social or biological populations, and could potentially be applied to the design of multi-agent systems to perform a group task in an uncertain, noisy situation that requires decision-making by the agents [29, 30].

2. Evolutionary matrix games on networks

In the context of game theory, a *game* is a simple model of an interaction between two or more players in which individuals' payoffs depend on the actions or strategies of each player. A *matrix game* is a 2-player game in which each player selects from a finite set of strategies and the payoff depends only on the strategies selected by the agents, such that all possible payoff outcomes of the game can be written in matrix form. See Appendix B for the formal definition of a Nash equilibrium and some of its refinements and Appendix A for a brief introduction to 2-player matrix games.

As an illustrative example, we consider one of the most famous games in all of game theory, the *prisoner's dilemma* (*PD*). In the original formulation, two suspects are arrested for a crime, and the police do not have enough evidence to convict either suspect, so they question them separately [31]. Each suspect can either cooperate (*C*) with the other by not answering any questions or defect (*D*) by testifying against the other suspect. The sentences for the two suspects depend on both of their actions, as shown in the following *payoff matrix*, in which one suspect chooses a row and their opponent chooses a column:

$$\pi^{PD} = \begin{matrix} & \begin{matrix} C & D \end{matrix} \\ \begin{matrix} C \\ D \end{matrix} & \begin{pmatrix} -2 & -5 \\ 0 & -4 \end{pmatrix} \end{matrix}.$$

Since this is a *symmetric* game, the payoff matrix is identical from the perspective of both players. If both suspects cooperate, the judge is lenient due to a lack of evidence that either suspect committed the crime and gives both suspects a 2-year sentence. However, if suspect 1 cooperates and suspect 2 defects, then the evidence points to suspect 1 who must serve the full sentence of 5 years while suspect 2 serves no prison time, and vice versa. Finally, if both suspects defect, the judge assumes one of them is guilty but without knowing which one, sentences them both to 4 years in prison.

The dilemma arises from the fact that although mutual cooperation would result in the best combined outcome, there is always a temptation for each suspect to defect to get a shorter sentence, resulting in a state of mutual defection. In other words, defection is the *best response* to both cooperation and defection by the other suspect. A state in which both players are playing best responses to each other is called a Nash equilibrium. See Appendix C for a complete categorization of 2×2 symmetric matrix games, which are the primary focus in this article. However, matrix games in general may have an arbitrary number of strategies, and we will also include some results about this more general case.

Evolutionary game theory extends classical game theory, which deals mostly with static concepts, by relaxing the assumption of perfect rationality and instead assuming that strategies propagate through a population through some dynamic process. Of the many different dynamics proposed for evolutionary games, we focus this article on two of the most widely studied models: imitation and best-response. In *imitation* dynamics, players adopt the strategy of their most successful neighbors, whereas in *best-response* dynamics, players choose strategies that will maximize their own respective payoffs. In summary, an *evolutionary matrix game* models the dynamic interactions of a population of agents, each of which chooses from a finite set of actions, receiving payoffs that depend on these actions. Based on the application and research goals, these dynamics are commonly expressed

either as systems of first-order ordinary differential equations (ODEs), as in RD, described in Section 2.3, or as discrete-time agent-based update rules on networks, defined in Section 2.1.

Analysis of evolutionary matrix games can help to understand and predict behavior of complex interconnected systems to answer questions such as whether each individual will settle on a particular strategy or how many agents will play each strategy on average, which we discuss further in Section 3. These questions motivate the search for engineering solutions to the associated control problems. For example, suppose selfish individuals tend to drive a particular network to undesired outcomes for the group, but the strategies of some agents in a network can be controlled through payoff incentives or other means. How can such a network be driven to a desired equilibrium state using a minimum amount of effort? How can the distribution of agent actions be changed and what are the achievable distributions? These and other issues are discussed further in Section 4.

2.1. Networks, games, and payoffs

Consider an undirected network $\mathbb{G} = (\mathcal{V}, \mathcal{E})$ in which the nodes $\mathcal{V} = \{1, \dots, n\}$ correspond to players, or *agents*, in a population and each edge in the set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ represents a 2-player game between neighboring agents. Each agent $i \in \mathcal{V}$ chooses *pure* strategies from a finite set $\mathcal{S} = \{1, 2, \dots, m\}$, and the payoff matrices associated with each edge $\{i, j\} \in \mathcal{E}$ are given by $\pi^{i,j}, \pi^{j,i} \in \mathbb{R}^{m \times m}$. Let $x := [x_1, \dots, x_n]^\top$ denote the strategy state of the network, where $x_i \in \mathcal{S}$ is the strategy played by agent i . When agent i plays strategy x_i against agent j who plays strategy x_j , agent i receives a payoff of $\pi_{x_i, x_j}^{i,j}$ and agent j receives a payoff of $\pi_{x_j, x_i}^{j,i}$. We denote the total payoff or utility to each agent $i \in \mathcal{V}$, which is accumulated over games with all neighbors, by $u_i : \mathcal{S} \times \mathcal{S}^{n-1} \rightarrow \mathbb{R}$, defined as

$$u_i(x_i, x_{-i}) = \sum_{j \in \mathcal{N}_i} \pi_{x_i, x_j}^{i,j}, \quad (1)$$

where $\mathcal{N}_i := \{j \in \mathcal{V} \mid \{i, j\} \in \mathcal{E}\}$ is the neighbor set of agent i , and x_{-i} denotes the strategies of all agents other than i . We will often consider the case when the payoff matrix for each agent is the same for all neighbors, i.e. $\pi^{i,1} = \pi^{i,2} = \dots = \pi^{i,|\mathcal{N}_i|}$ for each $i \in \mathcal{V}$, in which case we use the simplified notation π^i as in Appendix A.

2.2. Discrete-time evolutionary dynamics

In a model of the decisions or behaviors of a group of individuals, games generally correspond to real interactions occurring at specific times after which finite payoffs are collected. The dynamics therefore take place over a sequence of discrete times $k \in \{0, 1, 2, \dots\}$. At each time step, all agents in the nonempty set $\mathcal{A}^k \subseteq \mathcal{V}$ update their strategies according to the rule

$$x_i(k+1) = f_i(x(k), u(k)),$$

where $u(k) := [u_1(x_1(k), x_{-1}(k)), \dots, u_n(x_n(k), x_{-n}(k))]^\top$ denotes the vector of all agents' utilities at time k , and the functions f_i generally depend on local strategies $x_j(k)$ and/or payoffs $u_j(\cdot, \cdot)$ for each agent $j \in \mathcal{N}_i \cup \{i\}$ in the neighborhood of agent i . It is generally assumed that all games occur and payoffs are collected at every time step, regardless of which agents have updated. Updates can be *synchronous*, when every agent updates at each time step, or *asynchronous*, when only one agent updates at a time. Cases in which multiple agents but not necessarily all agents update at the same time will be referred to as *partially synchronous*. We define the *activation sequence* of the agents as $(\mathcal{A}^k)_{k=0}^\infty$ where $\mathcal{A}^k \subseteq \mathcal{V}$ denotes the set of agents active at time k . For synchronous updates, $\mathcal{A}^k = \mathcal{V}$, and for asynchronous updates, $\mathcal{A}^k$ is a singleton at every time k . We refer to the network, associated payoff matrices, and update rule as a *network game*, denoted by $\Gamma := (\mathbb{G}, \pi, f)$, where $f := [f_1, \dots, f_n]^\top$ and $\pi \in (\mathbb{R}^{m \times m})^{2|\mathcal{E}|}$ such that each $\pi^{i,j}$ is an entry in π .

2.2.1. Best-response update rule

In the *best-response update rule*, an agent who is active at time k updates at time $k+1$ to a strategy that achieves the highest total payoff, i.e. is a best response, against the strategies of its neighbors at time k . Sometimes referred to as *myopic* best-response, this rule tries to optimize payoff at the next time step based on only the state at the current time. Perhaps surprisingly, social experiments have revealed that for some simple game types, humans use myopic best responses as much as 96% of the time [32]. The best response update rule is given by:

$$x_i(k+1) \in \mathcal{B}_i(k), \quad (2)$$

where

$$\mathcal{B}_i := \{X \in \mathcal{S} \mid u_i(X, x_{-i}) \geq u_i(Y, x_{-i}) \quad \forall Y \in \mathcal{S}\}$$

As in the above equation, we sometimes omit the time k as long as there is no ambiguity. In the case that multiple strategies are best responses, i.e. $|\mathcal{B}_i(k)| > 1$, agents may be biased towards a particular strategy or prefer to keep their own strategy, provided that it belongs to the set of best responses.

Two-strategy best-response update. Since evolutionary game theory most often centers around two-strategy games, we give more emphasis to this special case, which, in the context of best-response dynamics, turns out to be equivalent to a threshold-based model. For this case, we have $S = \{A, B\}$ and a standard 2×2 payoff matrix:

$$\begin{array}{cc} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{pmatrix} a_i & b_i \\ c_i & d_i \end{pmatrix}, \quad a_i, b_i, c_i, d_i \in \mathbb{R}. \end{array}$$

Let $n_i^A(k)$ and $n_i^B(k)$ denote the number of neighbors of agent i playing A and B at time k , respectively. Accumulated over all neighbors, the total payoffs to agent i at time k amount to $a_i n_i^A(k) + b_i n_i^B(k)$ when $x_i(k) = A$, or $c_i n_i^A(k) + d_i n_i^B(k)$ when $x_i(k) = B$.

The two-strategy best-response update rule can then be simplified as follows:

$$x_i(k+1) = \begin{cases} A, & \text{if } a_i n_i^A(k) + b_i n_i^B(k) > c_i n_i^A(k) + d_i n_i^B(k) \\ B, & \text{if } a_i n_i^A(k) + b_i n_i^B(k) < c_i n_i^A(k) + d_i n_i^B(k) \\ z_i, & \text{if } a_i n_i^A(k) + b_i n_i^B(k) = c_i n_i^A(k) + d_i n_i^B(k) \end{cases}$$

The case in which strategies A and B result in equal payoffs is often either included in the A or B case, or set to $x_i(k)$ to indicate no change in strategy. For the purposes of generality, we allow for all three possibilities using the notation z_i , which may even vary across individual agents. However, to simplify the analysis, we assume that the z_i 's do not change over time.

Equivalence to linear threshold model. Let us now rewrite the above dynamics in terms of the number of neighbors playing each strategy. Let $\deg_i$ denote the total number of neighbors of agent i . We can simplify the conditions above by using the fact that $n_i^B = \deg_i - n_i^A$ and rearranging terms:

$$\begin{aligned} a_i n_i^A + b_i(\deg_i - n_i^A) &> c_i n_i^A + d_i(\deg_i - n_i^A) \\ n_i^A(a_i - c_i + d_i - b_i) &> \deg_i(d_i - b_i) \\ (\delta_i^A + \delta_i^B)n_i^A &> \delta_i^B \deg_i, \end{aligned} \tag{3}$$

where $\delta_i^A := a_i - c_i$ and $\delta_i^B := d_i - b_i$. The cases ' $<$ ' and ' $=$ ' can be handled similarly. First, consider the case when $\delta_i^A + \delta_i^B \neq 0$, and let $\tau_i := \frac{\delta_i^B}{\delta_i^A + \delta_i^B}$ denote a *threshold* for agent i . Depending on the sign of $\delta_i^A + \delta_i^B$, we have two categories of best-response update rules. If $\delta_i^A + \delta_i^B > 0$, the update rule is given by

$$x_i(k+1) = \begin{cases} A & \text{if } n_i^A(k) > \tau_i \deg_i \\ B & \text{if } n_i^A(k) < \tau_i \deg_i \\ z_i & \text{if } n_i^A(k) = \tau_i \deg_i \end{cases} \tag{4}$$

We call agents following such an update rule *coordinating agents*, because they seek to switch to strategy A if a sufficient number of neighbors are using that strategy, and likewise for strategy B . On the other hand, we call agents for which $\delta_i^A + \delta_i^B < 0$ *anti-coordinating agents*, because if a sufficient number of neighbors are playing A , they will switch to B , and vice versa. The anti-coordination update rule is given by

$$x_i(k+1) = \begin{cases} A & \text{if } n_i^A(k) < \tau_i \deg_i \\ B & \text{if } n_i^A(k) > \tau_i \deg_i \\ z_i & \text{if } n_i^A(k) = \tau_i \deg_i \end{cases} \tag{5}$$

In the special case that $\delta_i^A + \delta_i^B = 0$, the result is a stubborn agent who either always plays A or always plays B depending on the sign of δ_i^B and the value of z_i , and this agent can be considered as either coordinating or anti-coordinating with $\tau_i \in \{0, 1\}$, possibly with a different value of z_i .

The dynamics in (4) and (5) are in the form of the standard *linear threshold model*, which is widely used to study dynamics in social [33, 34], economic [35], neural [36], and various other types of networks. An equilibrium state in the threshold model is a state in which the number of A -neighbors of each agent will not lead to a change in strategy. For example, in a network of anti-coordinating agents in which $z_i = B$ for all i , this means that for each agent $i \in \mathcal{V}$, $x_i = A$ implies $n_i^A < \tau_i \deg_i$ and $x_i = B$ implies $n_i^A \geq \tau_i \deg_i$. Note that this notion of equilibrium is equivalent to a pure strategy Nash equilibrium in the corresponding network game.

2.2.2. Imitation update rule

The imitation update rule, though perhaps less rational, is fundamental to the field of evolutionary game theory, in part due to its close relation to RD, the celebrated model of evolutionary population dynamics [37]. Although originally posed as a behavioral model for animals and simple organisms, it is quite relevant to human social behavior as well [38, 39]. It is also growing in popularity as a research topic in the control community [40, 41, 42, 43]. Before analyzing the connection to RD, we first introduce imitation in a network setting.

The deterministic imitation update rule dictates that an agent i who is active at time k updates at time $k + 1$ to the strategy of the agent earning the highest payoff at time k in the self-inclusive neighborhood $\mathcal{N}_i \cup \{i\}$. In case several agents with different strategies earn the highest payoff, it is typically assumed that an updating agent who is currently playing one of these strategies will maintain the same strategy. Otherwise, to preserve the determinism of the dynamics, we assume there exists a priority such that agent i chooses the strategy with the smallest index, namely

$$x_i(k+1) = \begin{cases} x_i(k) & x_i(k) \in \mathcal{M}_i(k) \\ \min \mathcal{M}_i(k) & x_i(k) \notin \mathcal{M}_i(k) \end{cases} \quad (6)$$

where $\mathcal{M}_i \subseteq \mathcal{S}$ is the set of strategies resulting in the maximum payoff in the neighborhood of agent i :

$$\mathcal{M}_i := \left\{ x_j \in \mathcal{S} \mid u_j(x_j, x_{-j}) = \max_{r \in \mathcal{N}_i \cup \{i\}} u_r(x_r, x_{-r}) \right\}.$$

2.2.3. Proportional imitation update rule

It can also be useful to model agents whose decisions are characterized by some randomness. These agents may not always update when it seems beneficial to do so or they might occasionally make mistakes. A particularly noteworthy example of such dynamics is the *proportional imitation* rule, where each active individual chooses a neighbor at random and then, if her payoff is less than her neighbor's, imitates the neighbor's strategy with a probability proportional to the payoff difference [1, 11]. The proportional imitation rule is discussed in more detail in Section 2.3.1.

2.3. Continuous time evolutionary dynamics

The highly nonlinear nature of the discrete imitation and best-response dynamics makes studying their asymptotic behaviors challenging, especially when the population has some irregular structure. This motivates approximating the dynamics by making some simplifying assumptions on the population. The two most common and perhaps convenient assumptions are that the population has a large-enough size and is well-mixed (every agent interacts with every other agent). These two assumptions allow the evolution of the fraction of individuals playing a specific strategy to be described by an ordinary differential equation or inclusion, resulting in the *mean dynamic*.

For each $i \in \mathcal{S}$, let $0 \leq x_i \leq 1$ denote the population ratio of individuals playing strategy i . Although we use the same notation x as in Section 2.2, it differs from the case of finite populations, where each element x_i represents the strategy of an individual in the network. Rather, for infinite populations, $x := \sum_{i \in \mathcal{S}} x_i e^i$ captures the *population state*, where e^i denotes the i th column of the $m \times m$ identity matrix. Since $\sum_{i \in \mathcal{S}} x_i = 1$, we have that x belongs to the simplex Δ defined by

$$\Delta := \left\{ z \in \mathbb{R}^m \mid \sum_{i \in \mathcal{S}} z_i = 1, \quad z_i \geq 0 \quad \forall i \in \mathcal{S} \right\}.$$

The mean dynamic describing the evolution of x_i is

$$\dot{x}_i = \sum_{j \in \mathcal{S}} x_j \rho_{ji} - x_i \sum_{j \in \mathcal{S}} \rho_{ij}, \quad (7)$$

where $\dot{x}_i$ is the time derivative of x_i , and $\rho_{ij} \geq 0$ is the conditional switch rate from strategy i to strategy j of an arbitrary individual, based on the update rule. The mean dynamic provides a deterministic approximation for

stochastic evolutions. See [1] and the references therein for the derivation of the mean dynamic and results on the precision of the approximation, particularly as the population size and time approach infinity.

The mean dynamic of some stochastic versions of the imitation update rule (6), e.g., proportional imitation, lead to the most well-known evolutionary dynamics, the continuous RD, which are defined by a set of differential equations. The mean dynamic of the best-response update rule (2) leads to the continuous Best-Response Dynamics, a set of differential inclusions.

2.3.1. Derivation of the replicator dynamics

Assume the population is homogeneous in that individuals share a common payoff matrix $\pi \in \mathbb{R}^{m \times m}$, i.e., $\pi^i = \pi$ for all $i \in \mathcal{V}$. Representing the fitness of agent i , the utility of agent i is considered to be the average of her accumulated payoff here, and hence, is different from u_i defined in (1). However, since agents share the same payoff matrix, for every single one of them, we can use a common utility function $u : \Delta \times \Delta \rightarrow \mathbb{R}$ defined by $u(x, y) = x^\top \pi y$, where we have omitted the notation for the time dependency of x and y for simplicity. Assuming the population is well mixed, the average payoff of an individual playing strategy $i \in \mathcal{S}$ is $u(e^i, x)$, also known as the *fitness* of the individual, and the average payoff of the population is $u(x, x)$, also known as the *average fitness*. The replicator equation (dynamic) [37] governing the evolution of the population portion $x_i, i \in \mathcal{S}$, is described by

$$\dot{x}_i = [u(e^i, x) - u(x, x)]x_i. \quad (8)$$

From a biological perspective, RD describe an evolutionary process as follows: the reproduction rate of strategy- i players, $\dot{x}_i$, is proportional to the difference between the fitness of those players and the average population fitness. Put simply, the more fit strategy- i players are compared to the average, the more they reproduce.

To give an idea of the versatility and importance of RD, we present here three different approaches to derive it. The first two are based on the mean dynamic corresponding to two different versions of the imitation update rule, and the third is based on a biological interpretation of the payoffs.

First, RD can be derived from the *pairwise proportional imitation* update rule described in Section 2.2.3. In terms of the utility function $u(\cdot, \cdot)$, we have

$$\rho_{ij} = x_j[u(e^i, x) - u(e^j, x)]_+$$

where for $z \in \mathbb{R}$, $[z]_+ = z$ if $z \geq 0$ and $[z]_+ = 0$ otherwise [1]. From (7), we obtain the corresponding mean dynamic as follows

$$\begin{aligned} \dot{x}_i &= \sum_{j \in \mathcal{S}} x_j x_i [u(e^i, x) - u(e^j, x)]_+ - x_i \sum_{j \in \mathcal{S}} x_j [u(e^j, x) - u(e^i, x)]_+ \\ &= x_i \sum_{j \in \mathcal{S}} x_j [u(e^i, x) - u(e^j, x)] \\ &= x_i [u(e^i, x) \sum_{j \in \mathcal{S}} x_j - \sum_{j \in \mathcal{S}} x_j u(e^j, x)] \\ &= x_i [u(e^i, x) - u(x, x)], \end{aligned}$$

which is the replicator equation.

Second, RD can be derived from the *imitation of success* update rule: when an individual is active, she chooses one of her neighbors at random; then with a probability that is linearly increasing in the neighbor's payoff, she imitates the neighbor's strategy. Namely

$$\rho_{ij} = x_j [u(e^i, x) - M]$$

where M is some constant smaller than any feasible payoff so that the conditional switch rate is always positive [1]. From (7), we obtain the corresponding mean dynamic as follows

$$\begin{aligned} \dot{x}_i &= \sum_{j \in \mathcal{S}} x_j x_i [u(e^i, x) - M] - x_i \sum_{j \in \mathcal{S}} x_j [u(e^j, x) - M] \\ &= x_i \sum_{j \in \mathcal{S}} x_j [u(e^i, x) - u(e^j, x)] \\ &= x_i [u(e^i, x) - u(x, x)]. \end{aligned}$$

Last, RD can be derived directly from a biological perspective [7]. Let the utility of each individual represent

the number of offspring she reproduces per time unit. Suppose that each offspring inherits the strategy of her single parent. Then under the assumption of having continuous reproductions, the birthrate of individuals playing strategy i equals $u(e^i, x) + \alpha - \beta$ where α and β denote the individuals' background fitness (regardless of the course of the game) and death rate, respectively. This yields the population dynamics

$$\dot{n}_i = [u(e^i, x) + \alpha - \beta]n_i$$

where n_i is the number of individuals playing strategy i . Therefore, by using the identities $x_i = n_i/n$ and $n = \sum_i n_i$, we obtain

$$\begin{aligned}\dot{x}_i &= \frac{\dot{n}_i}{n} - \frac{\dot{n}}{n}x_i \\ &= [u(e^i, x) + \alpha - \beta]x_i - [u(x, x) + \alpha - \beta]x_i \\ &= [u(e^i, x) - u(x, x)]x_i.\end{aligned}$$

Although RD in the form of (8) do not allow for mutations, it turns out that the stability of the dynamics is implicitly related to notions such as *evolutionary stability* (ES) and *neutral stability* (NS), which are often stated in the context of mutant invasions to a population. See Appendix E for an introduction to evolutionarily stable strategies (states) (ESS) and neutrally stable strategies (states) (NSS).

Since $u(\cdot, \cdot)$ is continuously differentiable in $\mathbb{R}^m \times \mathbb{R}^m$, the dynamical system (8) has a unique solution for any $x(0) \in \Delta$ [7, Theorem 7.1.1]. The solution indeed satisfies the constraints $0 \leq x_i(t) \leq 1$ for all $i \in \mathcal{S}$ and $t \in \mathbb{R}$. It can be verified that for any time $t \in \mathbb{R}$, if $x(t) \in \Delta$, it holds that $\sum_{i \in \mathcal{S}} \dot{x}_i(t) = 0$. Hence, $\sum_{i \in \mathcal{S}} x_i(t) = 1$ is in force for all $t \in \mathbb{R}$, provided that $x(0) \in \Delta$. Therefore, the simplex Δ is invariant under (8), implying that the dynamical system (8) is well defined on Δ .

The dynamics in (8) form an s -dimensional RD with s strategies; however, due to the constraint $\sum_{i \in \mathcal{S}} x_i = 1$, they indeed form an $(s - 1)$ -dimensional dynamical system. Therefore, two-dimensional RD result in a simple flow on a line segment. See “Analyses of the replicator dynamics for 2x2 games” for some examples. Three-dimensional RD exhibit more complex behaviors such as heteroclinic cycles, yet do not admit limit cycles [44]. Based on the number of equilibria in the interior of the simplex, a complete classification of the possible phase portraits in three-dimensional RD is provided in [45] and [46]. The occurrence of strange attractors and chaos can take place in four-dimensional RD. This is mainly known due to the equivalence of RD and the general *Lotka-Volterra Equations* that are capable of such behaviors. See refappendix-LVE for details of this equivalence and known behavior in low-dimensional Lotka-Volterra systems.

One key concept simplifying the analysis of RD, is a *face* of the simplex. Following convention, the *boundary* of a set X , denoted by $\text{bd}(X)$, is the set of points x such that every neighborhood of x includes at least one point in X and one point out of X , and the *interior* of X , denoted by $\text{int}(X)$, is the greatest open subset of X . A face is defined as the convex hull of a non-empty subset $\mathcal{H}$ of the unit vectors e^i , $i \in \mathcal{S}$, and is denoted by $\Delta(\mathcal{H})$. A two-dimensional face is also called an *edge*. When $\mathcal{H}$ is proper, $\Delta(\mathcal{H})$ is called a *boundary face*, and when it includes only two members, $\Delta(\mathcal{H})$ is called an *edge*. It can be shown that each face of the simplex is invariant under RD [7]. It follows that the analysis of RD on Δ can be divided into two parts: (i) the boundary faces forming $\text{bd}(\Delta)$ and (ii) $\text{int}(\Delta)$. Note that each boundary face that is of a dimension greater than one is itself made of some smaller faces. We proceed with the following example, illustrating a three-dimensional RD.

Example 1 ([1, 7]). Consider the Rock-Paper-Scissors (RPS) game, which is a symmetric two-player game with the strategies “Rock”, “Paper” and “Scissors”, where Rock smashes Scissors, Paper covers Rock and Scissors cuts Paper. If we assign the payoff $w > 0$ to a win, and $-l < 0$ to a loss in a matching, and assume nothing is earned when the same strategies are matched, we obtain the payoff matrix

$$\pi = \begin{matrix} & \begin{matrix} \text{Rock} & \text{Paper} & \text{Scissors} \end{matrix} \\ \begin{matrix} \text{Rock} \\ \text{Paper} \\ \text{Scissors} \end{matrix} & \begin{pmatrix} 0 & -l & w \\ w & 0 & -l \\ -l & w & 0 \end{pmatrix} \end{matrix}.$$

The game is known to be good when $w > l$, standard when $w = l$, and bad when $w < l$ [1]. It can be verified that the unique Nash strategy of the game in all three types is $x^* = \left[\frac{1}{3} \quad \frac{1}{3} \quad \frac{1}{3}\right]^T$, which belongs to the interior of the simplex. So if any type includes an NSS or ESS, it has to be x^* . Indeed, we have the following for each case:

Good RPS: $\Delta^{\text{ESS}} = \Delta^{\text{NSS}} = \Delta^{\text{NE}} = \{x^*\}$.

Standard RPS: $\Delta^{\text{ESS}} = \emptyset$, $\Delta^{\text{NSS}} = \Delta^{\text{NE}} = \{x^*\}$.

Bad RPS: $\Delta^{ESS} = \Delta^{NSS} = \emptyset$, $\Delta^{NE} = \{x^*\}$.

Now we investigate the RD (8) under these three types of games. For all three, the equilibrium states of the system are

$$e^1, e^2, e^3, x^*$$

. We know that the solution trajectories evolves in the planar simplex Δ whose boundary is made of three edges, $\Delta(e^1, e^2)$, $\Delta(e^2, e^3)$ and $\Delta(e^3, e^1)$. Due to invariance of each edge under the RD, the behavior of the system on each edge can be studied independently from the rest of the simplex. For instance, to study the dynamics on $\Delta(e^1, e^2)$, we need to investigate the reduced two-dimensional RD with the 2×2 payoff matrix

$$\pi_{1,2} = \begin{matrix} & \begin{matrix} \text{Rock} & \text{Paper} \end{matrix} \\ \begin{matrix} \text{Rock} \\ \text{Paper} \end{matrix} & \begin{pmatrix} 0 & -l \\ w & 0 \end{pmatrix} \end{matrix},$$

which corresponds to strategies 1 and 2, i.e., Rock and Paper. The resulting game is a PD, resulting in a simple flow from e^1 to e^2 . Other edges exhibit the same flow.

To describe the behavior of solution trajectories with initial conditions in the interior of the simplex, we use the function $v : \text{int}(\Delta) \rightarrow \mathbb{R}_{\geq 0}$, defined by $v(x) = -\log(x_1 x_2 x_3)$ [47, 7]. Clearly, v is lower bounded by 0 and is unbounded above. Using the identity $x_1 + x_2 + x_3 = 1$, the time derivative of v can be calculated as

$$\begin{aligned} \dot{v}(x) &= -\frac{\dot{x}_1}{x_1} - \frac{\dot{x}_2}{x_2} - \frac{\dot{x}_3}{x_3} \\ &= -u(e^1, x) + u(e^2, x) + u(e^3, x) + 3u(x, x) \\ &= (l - w)(x_1 + x_2 + x_3) - \frac{3}{2}(l - w)((x_1 + x_2 + x_3)^2 - (x_1^2 + x_2^2 + x_3^2)) \\ &= (l - w)(1 - \frac{3}{2}(1 - \|x\|^2)) \\ &= \frac{1}{2}(l - w)(3\|x\|^2 - 1). \end{aligned}$$

Since $\frac{1}{3} \leq \|x\|^2 \leq 1$, it holds that $0 \leq \frac{1}{2}(3\|x\|^2 - 1) \leq 1$. So the sign of $\dot{v}(x)$ depends on the sign of $l - w$. Having this in mind, we investigate each type of the game separately:

Good RPS: We have $l - w < 0$, implying that $\dot{v}(x) \leq 0$ for all $x \in \text{int}(\Delta)$. Hence, v performs as a Lyapunov function for the system. On the other hand, $\dot{v}(x) = 0$ if and only if $x = x^*$. So x^* is globally asymptotically stable (in $\text{int}(\Delta)$; see Figure 1(a)).

Standard RPS: We have $l - w = 0$, implying that $\dot{v}(x) = 0$ for all $x \in \text{int}(\Delta)$. Hence, $x_1 x_2 x_3 = c > 0$, for all $x \in \text{int}(\Delta)$. So each interior trajectory $x(t)$ is a closed orbit satisfying $x_1(t)x_2(t)x_3(t) = x_1(0)x_2(0)x_3(0)$ for all t , except for when $x(t)$ starts from $x(0) = x^*$, when the closed orbit is reduced to the equilibrium state x^* (see Figure 1(b)). Therefore, x^* is Lyapunov stable, but no trajectory converges to x^* .

Bad RPS: We have $l - w > 0$, implying that $\dot{v}(x) \geq 0$ for all $x \in \text{int}(\Delta)$. Hence, if we reverse the direction of the vector field in the RD, i.e., $t \rightarrow -t$, then v will perform as a Lyapunov function. So x^* is a source. On the other hand, $v(x)$ is non-decreasing and grows arbitrarily large as $x \rightarrow \text{bd}(\Delta)$. Moreover, by linearization of the system about the vertices of the simplex, we find each of them a hyperbolic saddle. Therefore, it can be shown that every trajectory $x(t)$ starting from $\text{int}(\Delta) - \{x^*\}$, converges to the heteroclinic cycle on the boundary of the simplex (see Figure 1(c)).

2.3.2. Best-response dynamics

Under the best response update rule (2), the mean dynamics governing the evolution of state x_i can be shown to result in the following differential inclusion, known as the best response dynamic [49, 50, 51]:

$$\dot{x} \in \beta(x) - x, \quad \beta(x) := \{y \in \Delta \mid u(y, x) \geq u(z, x) \forall z \in \Delta\} \quad (9)$$

where $\beta(x)$ denotes the set of mixed strategies that maximize the utility against x . In contrast to the finite-population best-response dynamics (2), in which one agent updates to a best-response strategy at each time step, here there is a continuous evolution toward best-response strategies. Solutions to the best response dynamics exist, yet are not necessarily unique, which partly explains why the dynamics have received less attention compared to RD. We do not include the continuous time best response dynamics in our analysis and refer the reader to [52, 1]

Figure 1: Flow patterns of the Replicator Dynamics for the good, standard and bad Rock-Paper-Scissors (RPS) games. The plots indicate the evolution of the solution trajectory $x(t)$ for some initial conditions on the simplex Δ . Color temperatures are used to show motion speeds, where red corresponds to the fastest and blue to the slowest motion. The circles denote equilibria. For all three types of games, the dynamics admit three boundary equilibrium points, e^1, e^2 and e^3 , which are the vertices of the simplex, and a unique interior equilibrium x^* in the middle of the simplex. (a) good RPS: x^* is globally asymptotically stable in the interior of the simplex. (b) standard RPS: x^* is Lyapunov stable, and all orbits in the interior of the simplex are closed orbits, encircling x^* . (c) bad RPS: x^* is a source, and all orbits in the interior of the simplex converge to the heteroclinic cycle formed by the boundary of the simplex. The figures are produced using the software Dynamo [48].

for more information. In the rest of the paper, by “best response dynamics”, rather than (9), we simply mean the discrete dynamics caused by the best response update rule.

3. Equilibrium convergence and stability

3.1. Infinite well-mixed populations (continuous-time continuous-state dynamics)

General stability and convergence results for continuous dynamics such as the Poincaré-Bendixson theorem can, of course, be used for RD whenever applicable [53, 43]. However, interestingly enough, convergence, and particularly, stability notions such as Lyapunov stability for RD, are tightly linked to game theoretical notions such as Nash equilibrium. The first are notions defined for vector fields and dynamical systems whereas the second are defined in the absence of any dynamic. This enables us to make conclusions on the evolution of the solution trajectories of RD through game-theoretic analyses of the corresponding payoff matrix. In what follows, we review some well-known results that reveal these links. The reader may consult any of the books [54, 7, 1] for more details.

3.1.1. Equilibrium states

By letting the right-hand side of (8) equal to zero, we get the condition $u(e^i, x) = u(x, x)$ or $x_i = 0$ for all $i \in \mathcal{S}$, for x to be an equilibrium. This is closely related to the definition of a Nash strategy, as revealed by the following proposition [7]. Let Δ^o denote the set of equilibrium states of (8).

Proposition 1. *The following holds*

- $\Delta^{NE} \cup \{e^1, \dots, e^m\} \subseteq \Delta^o$;
- $\text{int}(\Delta^o) = \Delta^{NE} \cap \text{int}(\Delta)$;
- $\text{int}(\Delta^o)$ is convex.

In addition to the trivial result that every vertex of the simplex is an equilibrium, the first statement implies that every Nash strategy is also an equilibrium of RD. Namely, if a strategy is the best-response to itself, the corresponding population state is an equilibrium of RD. The reverse, however, does not hold. Yet the second statement clarifies that only a non-interior equilibrium state may not be a Nash strategy. That is, all interior equilibria are a Nash strategy. The last statement postulates the convexity of the set of interior equilibria, implying that it can be a singleton, straight line, plane, etc. in the simplex, but it cannot be two disjoint points, for example.

3.1.2. Lyapunov stable states

The role of Nash strategies is not limited to equilibrium states. The following result states that every stable state in (8) has to be a Nash strategy [55].

Proposition 2. *If $x^* \in \Delta$ is Lyapunov stable in (8), then $x^* \in \Delta^{NE}$.*

As an application of the Proposition, the equilibrium point x^* in Example 1 is Lyapunov stable in the good and standard RPS games, and indeed is neutrally stable in both cases. Intuitively, Proposition 2 implies that only population states whose corresponding strategy vector performs best against itself can be stable under RD. This necessary condition is not sufficient though; a simple example is RD for the 2×2 coordination game where the mixed strategy is a Nash strategy but is unstable (see “Analyses of the replicator dynamics for 2x2 games”). The following result provides a sufficient stability condition [7].

Proposition 3. *Every $x^* \in \Delta^{NSS}$ is Lyapunov stable in (8).*

Again back to Example 1, x^* is neutrally stable in the good and standard RPS games, and is also Lyapunov stable in both cases. According to Proposition 3, if strategy x^* performs at least as well as any group of mutants arising in a population of all x^* -players, provided that the mutant population is small enough, then under RD, the solution trajectory remains in a small neighborhood of the population state x^* , if it starts sufficiently close to x^* . The reverse, however, does not hold as discussed in Example 2.

3.1.3. Asymptotically stable states

Clearly, in view of Proposition 2, every asymptotically stable state has to be a Nash strategy, yet this strong notion of stability further confines the Nash strategy to a perfect Nash strategy [55].

Proposition 4. *If $x^* \in \Delta$ is asymptotically stable in (8), then $x^* \in \Delta^{PNE}$.*

So if every close-by trajectory to the population state x^* , not only remains close, but also converges to x^* , then x^* must be a Nash strategy that is robust against some ‘trembles’ in the player’s strategy. The necessary condition in Proposition 4 is not sufficient though. For instance, x^* in Example 1 is perfect since it belongs to the interior of the simplex, yet it is unstable in the bad RPS game.

Now we proceed to the most well-known result for RD, providing a sufficient condition for asymptotic stability by bridging this notion to that of evolutionary stability [37, 56].

Proposition 5. *Every $x^* \in \Delta^{ESS}$ is asymptotically stable in (8).*

Back to Example 1, x^* is evolutionarily stable in the good RPS game and is also asymptotically stable. Note that according to the proposition, *strict Nash strategies* which are Nash strategies that satisfy (B.1) with strict inequality, are also asymptotically stable since they are an ESS. Intuitively, this condition is not necessary for asymptotic stability: evolutionary stability requires x^* to outperform any sufficiently small group of mutants against the resulting population mixture, yet for asymptotic stability of x^* , it is possible to do worse than some mutant type $y \in \Delta$, provided that there is some other mutant $z \in \Delta$ that outperforms y , but does worse against x^* . This is illustrated in the following example.

Example 2. *Consider a symmetric two-player game with the payoff matrix*

$$\pi = \begin{pmatrix} 0 & 6 & -4 \\ -3 & 0 & 5 \\ -1 & 3 & 0 \end{pmatrix}.$$

The game is often referred to as Zeeman’s game, named after Zeeman [57, 1]. We study the game under the RD. The equilibrium points and their local stability are as follows: five equilibrium points on $\text{bd}(\Delta)$: e^1 : hyperbolic stable, e^2 : hyperbolic unstable, e^3 : hyperbolic saddle, $x^{23} = \left[0 \quad \frac{5}{8} \quad \frac{3}{8}\right]^T$, located on the edge $\Delta(e^2, e^3)$: hyperbolic saddle, and $x^{13} = \left[\frac{4}{5} \quad 0 \quad \frac{1}{5}\right]^T$, located on the edge $\Delta(e^1, e^3)$: hyperbolic saddle, and one equilibrium point in $\text{int}(\Delta)$: $x^ = \left[\frac{1}{3} \quad \frac{1}{3} \quad \frac{1}{3}\right]^T$: hyperbolic stable. So e^1 and x^* are asymptotically stable. Indeed, it can be shown that for every initial condition in $\text{int}(\Delta)$ except for those located on the stable manifold of x^{13} which connects e^2 to x^{13} , the solution trajectory converges to either e^1 or x^* (see Figure 2).*

Figure 2: Flow patterns of the Replicator Dynamics for Zeeman's game. The plot indicates the evolution of the solution trajectory $x(t)$ for some initial conditions on the simplex Δ . Color temperatures are used to show motion speeds where red corresponds to the fastest and blue to the slowest motion. The circles denote equilibria. The dynamics admit five boundary equilibrium points e^1 : hyperbolic stable, e^2 : hyperbolic unstable, e^3 : hyperbolic saddle, x^{23} , located on the edge $\Delta(e^2, e^3)$: hyperbolic saddle, and x^{13} , located on the edge $\Delta(e^1, e^3)$: hyperbolic saddle, and one interior equilibrium x^* : hyperbolic stable. Except for those initial conditions located on the stable manifold of x^{13} which connects e^2 to x^{13} , for every other initial condition in $\text{int}(\Delta)$, the solution trajectory converges to either e^1 or x^* . Independently from the dynamics, analysis of the game reveals that $\Delta^{NE} = \{e^1, x^{13}, x^*\}$ and $\Delta^{ESS} = \{e^1\}$. So although x^* is asymptotically stable, it is not an ESS. The figure is produced using the software Dynamo [48].

The asymptotically stable state x^* , however, is not an evolutionary stable state. For example, if any portion $\epsilon \in (0, 1)$ of mutants playing x^{13} appear in a population of all x^* -players, the incumbent x^* s do worse than the mutants against the resulting population mixture:

$$u(x^*, (1 - \epsilon)x^* + \epsilon x^{13}) - u(x^{13}, (1 - \epsilon)x^* + \epsilon x^{13}) = \frac{-\epsilon}{5} < 0. \quad \forall \epsilon \in (0, 1).$$

This explains the deviation of $x(t)$ from x^* to x^{13} in the area between the two. However, the argument does not imply the instability of x^* . The state x^{13} cannot resist any mutant of e^3 -players. This can be seen from both the flow on the segment between x^{13} and e^3 and the following calculation

$$u(x^{13}, (1 - \epsilon)x^{13} + \epsilon e^3) - u(e^3, (1 - \epsilon)x^{13} + \epsilon e^3) = \frac{16\epsilon}{5} > 0 \quad \forall \epsilon \in (0, 1).$$

On the other hand, x^* outperforms any portion of e^3 -playing mutants:

$$u(x^*, (1 - \epsilon)x^* + \epsilon e^3) - u(e^3, (1 - \epsilon)x^* + \epsilon e^3) = \frac{\epsilon}{3} > 0 \quad \forall \epsilon \in (0, 1).$$

So there is indeed a cyclic-invasion relationship between x^* , x^{13} and e^3 , which is in line with the phase plot.

At the same time, the equilibrium e^1 is both asymptotically and evolutionary stable. Apparently, it is the only ESS of the game, i.e., $\Delta^{ESS} = \{e^1\}$. It is worth mentioning that just by the local stability information provided above, we expect e^1 and x^* to be Nash strategies since e^1 and x^* are Lyapunov stable (hence, in view of Proposition 2 they have to be Nash strategies). This is upheld by the fact that $\Delta^{NE} = \{e^1, x^{13}, x^*\}$.

3.1.4. Globally asymptotically stable states

All results presented so far hold only locally. Particularly, for evolutionarily stable states, their asymptotic stability is not necessarily global. If $x^* \in \Delta^{ESS}$ is globally asymptotically stable, then it can be shown to be the unique ESS of the game. We know that an interior ESS fulfills this condition, but is it necessarily globally asymptotically stable? The following proposition provides a positive answer [54].

Proposition 6. *If $x^* \in \text{int}(\Delta) \cup \Delta^{ESS}$, then $\lim_{t \rightarrow \infty} x(t) = x^*$ if $x(0) \in \text{int}(\Delta)$.*

Back to Example 1, the equilibrium state x^* in the good RPS game is the unique ESS of the game, and is also globally asymptotically stable in $\text{int}(\Delta)$.

3.1.5. Convergence results

As with most other dynamics, existing stability results on RD are insufficient to describe the global behavior of the trajectories. Local stability analysis illustrates the flow only in the neighborhood of the equilibria, and global ones, such as global stability, rarely take place in these dynamics. One fundamental question on the global behavior of the dynamics is *under what conditions every solution trajectory converges to an equilibrium point?* Symmetry of the payoff matrix turns out to guarantee this. The average fitness in *doubly symmetric games*, those are symmetric games with a symmetric payoff matrix, is non-decreasing over time, i.e., $\dot{u}(x, x) \geq 0$ with equality if and only if x is an equilibrium point. This can be considered as a version of the *Fundamental Theorem of Natural Selection* [58] stated in a more general context at 1930, before the invention of RD: *The rate of increase in fitness of any organism at any time is equal to its genetic variance in fitness at that time.* However, this does not prove equilibrium convergence, and by using theorems such as LaSalle's invariance principle, we can only show convergence to the set of equilibrium states in RD. Yet it was later shown in [59] that symmetry of the payoff matrix is indeed sufficient for equilibrium convergence.

Proposition 7. *If the game is doubly symmetric, i.e., $\pi = \pi^\top$, then every trajectory converges to an equilibrium point.*

Often the payoff matrix is not symmetric, yet is *symmetrizable* by local shifts, that is the addition of a constant to every entry of a column, captured by the transformation $T : \mathbb{R}^{m \times m} \rightarrow \mathbb{R}^{m \times m}$ defined as $T(\pi) = \pi + \mathbf{1}c^\top$ where $\mathbf{1}$ is the all-one and c is a constant vector, both in $\mathbb{R}^{m \times 1}$. The RD are invariant under local shifts [7]. Therefore, Proposition 7 also holds for any game with a symmetrizable payoff matrix. For example, all symmetric 2×2 games admit equilibrium convergence under RD. Some games, however, are neither doubly symmetric nor symmetrizable, yet exhibit equilibrium convergence, such as the good RPS game in Example 1.

Now given the convergence of RD for a particular payoff matrix, the next natural question is to what kind of equilibrium point the trajectory may converge. Apparently if the trajectory starts from some interior point of the simplex, the final stationary state has to be the best response to itself [60].

Proposition 8. *If $x(0) \in \text{int}(\Delta)$ and $\lim_{t \rightarrow \infty} x(t) = x^*$, then $x^* \in \Delta^{NE}$.*

Visiting back Example 2, based on the dynamics, we expect $x^{13} \in \Delta^{NE}$ since at least one interior trajectory converges to x^{13} , which is the one on the stable manifold of x^{13} . The role of a Nash strategy is not limited to Proposition 8 though. If the interior of the simplex is empty of Nash strategies, or equivalently empty of any equilibrium point (see Proposition 1), then every trajectory converges to the boundary of the simplex [61].

Proposition 9. *If $\text{int}(\Delta) \cap \Delta^{NE} = \emptyset$, then $x(t) \rightarrow \text{bd}(\Delta)$ as $t \rightarrow \infty$.*

So it is impossible for the trajectories to meander forever in the interior of an empty-equilibrium simplex.

3.1.6. Folk Theorem

By summarizing some of the results we presented so far on the relation of Nash strategies and RD, we arrive at the following theorem [62, 63].

Theorem 1 (Folk Theorem of Evolutionary Game Theory). *The RD (8) satisfy the following statements:*

1. *A Lyapunov stable equilibrium is a Nash strategy;*
2. *If a solution trajectory starting from the interior of the simplex converges to a point, that point has to be a Nash strategy;*
3. *A strict Nash strategy is locally asymptotically stable.*

The results of the theorem are known to be a paradigm shift towards strategic reasoning: *The Folk Theorem means that biologists can predict the evolutionary outcome of their stable systems by examining behavior of the underlying game* [62]. However, the theorem does not hold for every other type of dynamics such as the monotone selection dynamics [62, 7].

3.2. Infinite structured populations (continuous-time continuous-state dynamics)

Before returning to the case of finite populations, we briefly discuss some results for infinite populations in which the well-mixed assumption is relaxed. That is, individuals of a given type (strategy) may interact with other types of individuals in different proportions. These interactions can be modeled by a graph in which the nodes represent the available strategies and the connections are weighted according to the interaction probabilities between each strategy pair. The corresponding mean dynamics can then be expressed while taking these interaction probabilities into account [64]. Some important properties such as stability of Nash equilibria are preserved under this modification for various dynamics, even when individuals playing particular strategies do not have full information about the utilities associated with other strategies [65].

3.3. Finite structured populations (discrete-time discrete-state dynamics)

We now turn back to the general case of discrete-time dynamics on finite networks, for which both the notion of convergence and the tools needed for its analysis differ substantially from the continuous case. In continuous dynamics, convergence of a solution trajectory $x(t)$ to an equilibrium point x^* implies that the distance from the trajectory to the equilibrium, i.e., $\|x(t) - x^*\|$, becomes arbitrary small, yet never necessarily zero, as time goes to infinity. On the other hand, discrete population dynamics take place over a discrete time sequence $k = 0, 1, 2, \dots$ and the state of system is a discrete vector $x \in \mathcal{S}^n$. Therefore, notions such as ‘arbitrary small yet nonzero’ are undefinable for the distance from the trajectory to an equilibrium, because $\|x(k) - x^*\|$ takes discrete values $0, 1, 2, \dots$. Instead, we have the notion of ‘reaching’ an equilibrium state, that is, when $x(k)$ exactly equals x^* . This highlights a key difference from convergence in continuous dynamics, which is that the state $x(k)$ becomes fixed at x^* at some finite time T and does not change afterwards, whereas convergence in continuous dynamics typically never leads to the state reaching an equilibrium in finite time. Despite the differences, ‘convergence’ is often used for both continuous and discrete dynamics, a convention that we also adopt in this paper. However, it should be clear to the reader that the study of equilibrium convergence for discrete population dynamics can be interpreted as the investigation of whether solution trajectories reach equilibrium states.

As discussed in Section 3.1, continuous best-response and especially imitation dynamics can exhibit non-convergence and chaotic asymptotic behaviors, even under the simplifying assumptions of infinite and well-mixed populations. No less complex outcomes are therefore to be expected for finite structured populations governed by the discrete forms of these dynamics. Indeed, populations of just two (resp. three) agents can lead to non-convergence in discrete best-response (resp. imitation) dynamics. Hence, it becomes a challenging problem to identify conditions under which a network can be expected to converge to an equilibrium under imitation and best-response update rules, which we investigate in this section.

Inspired by the convergence results for the continuous population-dynamics discussed in earlier sections, one may try to find Lyapunov (energy) functions, which are strictly positive and decreasing at each time step, to prove equilibrium convergence in the discrete case. Although sometimes possible, one should note that finding *energy-like functions*, which decrease at some time instances but remain constant at others, may be easier to construct and often the only choice. For example, in the discrete population-dynamics, a common phenomenon is an active agent at time k not switching strategies, which keeps the state $x(k+1)$ the same as its value at the previous time instance $x(k)$. Thus, the value of any candidate energy function remains the same at time $k+1$, excluding it from being an strictly decreasing energy function, yet leaving room for being an energy-like function.

The number of agents updating at the same time, namely, the level of synchrony of the dynamics, turns out to greatly influence the asymptotic behavior of the population. For example, synchronous versions of the update rules (2) and (6) are deterministic: there is exactly one possible state x^2 to which the solution trajectory can transit from a given state x^1 , where x^1 and x^2 can be the same. Hence, the dynamics admit a unique ω -limit set in the form of a cycle of length at least 1, which is an equilibrium, and at most $|\mathcal{S}|^n$ the number of possible states. The challenge then becomes determining the length of the limit cycle. For partial and asynchronous updates; however, the transition to the next state from a given state is usually not unique and depends on the active agent, resulting in non-deterministic dynamics, capable of exhibiting chaotic fluctuations as well as perpetual cycles.

In order to guarantee convergence in the discrete case, we must assume that the activation sequence is *persistent*, that is, every agent $i \in \mathcal{V}$ becomes active infinitely many times as time goes to infinity. Formally, given an agent $i \in \mathcal{V}$ and any time $k \in \{0, 1, \dots\}$, there is some finite future time $k' > k$ at which agent i is active. One can easily check that without this assumption equilibrium convergence may never happen.

3.3.1. Imitation

The best-response strategy in a 2-player matrix game is the one corresponding to the maximum entry in the column of the payoff matrix defined by the opponent’s strategy. It is therefore intuitive that convergence relies on the ordering of payoff values within each column (which determines whether the agents are coordinating or

anti-coordinating [13]). On the other hand, under the imitation rule, the payoffs of the opponent play a greater role. Consequently, under imitation, the ordering of payoff values within each row appears to be the key.

We call an agent $i \in \mathcal{V}$ an *opponent-coordinating* agent if each diagonal entry of her payoff matrix π^i is greater than all off-diagonal elements in the same row, that is

$$\pi_{p,p}^i > \pi_{p,q}^i \quad \forall p, q \in \{1, \dots, m\}, p \neq q. \quad (10)$$

The following theorem establishes convergence when updates are fully asynchronous [66, 43].

Theorem 2. *Every network of opponent-coordinating agents reaches an equilibrium under the asynchronous imitation update rule.*

However, in order to guarantee convergence under synchronous and partially synchronous updates, the agents must satisfy a stronger condition than being opponent-coordinating. In other words, their payoff matrices must satisfy

$$\pi_{p,p}^i + (\deg_i - 1)\pi_{p,p_{\min}}^i > \deg_i \pi_{p,p_{\max}}^i \quad (11)$$

where $\deg_i$ denotes the degree of agent i , $p_{\min}$ denotes the column of the minimum off-diagonal entry of the p th row in π^i and $p_{\max}$ denotes the column of the maximum off-diagonal entry of the p th row in π^i .

Then we can assert the following result, regardless of how many agents update simultaneously [66].

Theorem 3. *Every network of strongly-opponent-coordinating agents reaches an equilibrium under the imitation update rule.*

For the special case of $m = 2$, an opponent-coordinating agent turns out to be also strongly opponent-coordinating, implying that Theorem 2 holds even for partially synchronous dynamics.

3.3.2. Best-response

While there is no guarantee that a network of agents using best response updates (2) will reach an equilibrium state in general, we present here three conditions on 2×2 payoff matrices for which convergence is assured when updates are persistent and asynchronous: (i) when all agents are coordinating, (ii) when all agents are anti-coordinating, and (iii) when all games are symmetric. We refer the reader to Appendix G for proofs of the following results using potential functions.

When all agents in the network are coordinating (see Section 2.2.1), we have the following theorem [13].

Theorem 4. *Every network of coordinating agents who update asynchronously and persistently will reach an equilibrium state.*

The following corollary follows directly from Theorem 4.

Corollary 1. *Every network of coordinating agents admits a pure strategy Nash equilibrium.*

The case of all anti-coordinating agents is perhaps a more surprising result, but convergence is indeed guaranteed here as well [13].

Theorem 5. *Every network of anti-coordinating agents who update asynchronously and persistently will reach an equilibrium state.*

The following corollary follows directly from Theorem 5.

Corollary 2. *Every network of anti-coordinating agents admits a pure strategy Nash equilibrium.*

Although we have mostly considered cases when the payoff matrix π^i of an agent applies to all neighbors, the general case allows for a different payoff matrix $\pi^{i,j}$ for each neighbor j . It turns out that such a network will reach equilibrium provided that all pairwise games are symmetric.

Theorem 6. *Every network of agents in which all games are symmetric, i.e. $\pi^{i,j} = \pi^{j,i}$ for all $\{i, j\} \in \mathcal{E}$, who update asynchronously and persistently will reach an equilibrium state.*

The following corollary follows directly from Theorem 6.

Corollary 3. *Every network of agents in which all games are symmetric admits a pure strategy Nash equilibrium.*

Although these three conditions are sufficient to guarantee convergence in asynchronous networks of agents who update with best-responses, if these conditions do not hold, it is quite easy to construct examples in which the agents will never reach an equilibrium state. One such example is a network in which updates are fully synchronous may never reach an equilibrium state. Suppose that, in a network consisting of only two agents, both agents are anti-coordinating, update synchronously, and start from the strategy state (A, A) . The dynamics are therefore deterministic, and the following state of the agents will evolve as follows:

$$(A, A) \rightarrow (B, B) \rightarrow (A, A),$$

resulting in a cycle of length 2. Now suppose that both agents are coordinating and they start from (A, B) . This will lead to the following deterministic sequence: $(A, B) \rightarrow (B, A) \rightarrow (A, B)$, again resulting in a cycle of length 2.

The above examples prove that equilibrium convergence is no longer guaranteed if the agents update in full synchrony.

Finally, we consider the case of partial synchrony, in which multiple agents, but not always all, update at each time step. These results require the following assumption, since we need some notion of real time to study this case.

Assumption 1. *The inter-activation times for each agent are drawn from mutually independent probability distributions with support on $\mathbb{R}_{\geq 0}$.*

Based on this, the following theorem establishes convergence in a probabilistic sense [13].

Theorem 7. *Every network of all coordinating or all anti-coordinating agents who update with partially synchronous dynamics that satisfy Assumption 1 almost surely reaches an equilibrium state in finite time.*

4. Control strategies

Even when populations of interacting agents converge to an equilibrium, it may be an undesirable one from the perspective of a global cost function or some measure of collective welfare. In these cases, and when a central agency has the ability to influence the actions or perturb the dynamics of some of the agents, it is of great interest to understand how such influence can be efficiently used to improve global outcomes in the network. In this section, we investigate and compare several different approaches for doing this.

4.1. Infinite well-mixed populations

When the system of interest is defined on a sufficiently large and densely connected population, it may be useful to devise control strategies based on mean dynamics such as RD. Suppose the goal is to drive the system modeled by RD to a particular population state $x^* \in \Delta$. One of the first things to consider is what the control input should be. Although there are multiple possible approaches, perhaps the most natural and least invasive control input is to alter the payoff functions for the available strategies. This can mean offering incentives by adding to the utility of desired strategies, or enforcing penalties by subtracting from the utility of undesired strategies. When viewed from the perspective of a regulating agency or government, these two options are sometimes considered as subsidies and taxes [67]. This type of control input $v_i(t)$ can be incorporated into RD as follows:

$$\dot{x}_i = [u(e^i, x) + v_i - u(x, x)]x_i.$$

Let $\mathcal{D} := \{i \in \mathcal{S} : x_i^* > 0\}$ denote the set of desired strategies. It was proposed in [68] that a regulator, having a fixed amount αn of available funds, could offer subsidies to agents using each desired strategy $i \in \mathcal{D}$ in the amount of $v_i = \alpha \frac{x_i^*}{x_i}$, and $v_i = 0$ for each $i \notin \mathcal{D}$. The closed-loop dynamics for $i \in \mathcal{D}$ then become

$$\dot{x}_i = \left[\left(u(e^i, x) + \alpha \frac{x_i^*}{x_i} \right) - (u(x, x) + \alpha) \right] x_i = [u(e^i, x) - u(x, x)]x_i + \alpha(x_i^* - x_i). \quad (12)$$

Notice that when adding $\alpha \frac{x_i^*}{x_i}$ to the utility of each pure strategy e^i , the average utility becomes

$$\sum_{i \in \mathcal{S}} x_i \left(u(e^i, x) + \alpha \frac{x_i^*}{x_i} \right) = \sum_{i \in \mathcal{S}} x_i u(e^i, x) + \sum_{i \in \mathcal{S}} \alpha x_i^* = u(x, x) + \alpha,$$

since both $x, x^* \in \Delta$. Also extended to the case of heterogeneous populations (multi-population games) [69], this formulation has the property that equilibrium points of (12) are also equilibrium points of the original RD (8) for any $\alpha > 0$. This allows for stabilization of existing equilibrium points of a system, but not for stabilization of arbitrary setpoints x^* , which remains an open research problem.

Similar to this approach, but using logit choice dynamics (a stochastic or noisy version of the best-response update rule), [70] proposed pricing schemes to encourage efficient choices on roadway networks. This line of research has been extended to more general economic contexts in [71]. In a continuous-time network setting, [72] characterized the number of control nodes required to stabilize desired equilibrium states of a best-response-like epidemic model on path and star networks.

Although beyond the scope of this article, the shifting of equilibrium points has been considered in the context of stochastically stable equilibria of evolutionary snowdrift games [73]. Also in a systems and control setting, RD has been invoked to study task assignment among team members [74], virus spread control [75], extremum-seeking controllers [76] and direct reciprocity [77].

4.2. Finite structured populations

Some concepts related to the control of populations on networks may differ from those related to control of conventional dynamical systems, including the replicator dynamics. For example, there are multiple ways to affect the dynamics in a way that can be considered a control input, including:

- **direct strategy control** – in some cases, it may be possible to fix the strategies of some of the agents in the network, with the hope that there will be a cascading effect that produces a desired global outcome;
- **incentive-based control** – perhaps a more practical input is to give additional payoff as a reward or subtract payoff as a penalty for using particular strategies.

Though not an exhaustive list, since most of the existing literature falls within one of these categories, we will focus our attention on these. The control objective may also take different forms. For example, there may be a desired end state in some cases, but in other cases, this might not be a realistic goal. Rather, the goal could be to guide the population as close as possible toward a desired state.

Since very few mechanisms have been proposed for feedback control of population dynamics, we focus primarily on open-loop methods. A practical and compelling use of feedback in these kinds of systems remains an interesting outstanding problem.

4.2.1. Unique convergence for a class of coordinating agents

Dynamics that are known to converge in the absence of control serve as a natural starting point for the design of control policies that try to steer a network towards a desired equilibrium. A potential obstacle in the design of such policies is that, due to the random sequence of updates under asynchronous dynamics, convergence does not generally result in a unique equilibrium. This makes it much more difficult to predict the outcome of a control action in an open-loop setting. However, there is a very useful property of networks consisting only of coordinating agents, which is that, if some strategies are fixed to A , or if incentives are provided for playing A , not only is convergence guaranteed, but the network will converge to a *unique* equilibrium.

Before introducing this result, we need the following definitions.

Definition 1. We say that agent i is **coordinating in** X if whenever the agent would update to strategy $X \in \mathcal{S}$, it would also do so if some neighbors currently not playing X were instead playing X , that is, if for all $x, x' \in \mathcal{S}^n$,

$$\text{if } x'_j = \begin{cases} x_j & \text{if } j \notin N_i \text{ or } x_j = X \\ x_j \text{ or } X & \text{otherwise} \end{cases},$$

$$\text{then } f_i(x, u_i) = X \implies f_i(x', u_i) = X,$$

where $u_i : \mathcal{S} \times \mathcal{S}^{n-1} \rightarrow \mathbb{R}$ is the utility function for agent i .

For the next two definitions, we consider a situation in which some payoff incentives are offered to the agents in a network which is at equilibrium. Let $(\mathcal{G}, \pi, \mathcal{R})$ denote a network game in which an equilibrium x^* exists, and let $\tilde{\pi}$ denote a modification of the payoff matrix π such that for all $Y, Z \in \mathcal{S}$

$$\tilde{\pi}_{YZ}^i = \begin{cases} \pi_{YZ}^i + r_i & \text{if } Y = X \\ \pi_{YZ}^i & \text{otherwise} \end{cases}$$

for all $i \in \mathcal{V}$, where each $r_i \in \mathbb{R}_{\geq 0}$.

Definition 2. We say that a network game is **monotone in X** if whenever some non-negative value is added to the row corresponding to strategy X in the payoff matrices of some agents, when the network is currently at an equilibrium state, no agent will switch from X to some other strategy. That is, if $x(k) = x^*$, then under the network game $(\mathbb{G}, \tilde{\pi}, \mathcal{R})$,

$$x_i(k) = X \implies x_i(k+t) = X, \quad \text{for all } i \in \mathcal{V} \text{ and } t \geq k.$$

Definition 3. We say that a network game is **uniquely convergent in X** if whenever some non-negative value is added to the row corresponding to strategy X in the payoff matrices of some agents, the network converges to a unique equilibrium state. Namely, when starting from x^* , the network game $(\mathbb{G}, \tilde{\pi}, \mathcal{R})$ converges to a unique equilibrium $\tilde{x}$ regardless of the sequence in which the agents update their strategies.

We are now ready to state an important convergence result for networks of coordinating agents [78].

Theorem 8. Every network that is coordinating in strategy X is both monotone and uniquely convergent in X .

We can relate this result to best-response and imitation dynamics with the following corollaries [78].

Corollary 4. Every network consisting of all coordinating agents under two-strategy best-response dynamics is coordinating in X for all $X \in \mathcal{S}$.

PROOF. According to the best-response rule (2), a coordinating agent will switch to a strategy X if a sufficient number of neighbors are currently playing X . It is therefore trivially true that if more agents in the network are playing X , the agent will still update to X .

The following corollary provides a similar result for opponent-coordinating agents under imitation dynamics [78].

Corollary 5. Every network consisting of all opponent-coordinating agents under imitation dynamics is coordinating in X for all $X \in \mathcal{S}$.

4.2.2. Direct strategy control

Suppose that by some means it is possible to fix the strategies of a subset $\mathcal{L} \subseteq \mathcal{V}$ of the agents in a network:

$$x_i(k+1) = \begin{cases} x_i^*, & i \in \mathcal{L} \\ f_i(\cdot), & \text{otherwise} \end{cases}, \quad (13)$$

where x_i^* denotes the strategy to fix for each control agent $i \in \mathcal{L}$. In this case, it may be useful to know the minimum number of control agents necessary to drive the population to a desired equilibrium state.

Problem 1. Given a network game $(\mathbb{G}, \tilde{\pi}, f)$ and initial state $x(0)$, find the smallest set of control agents $\mathcal{L}^* \subseteq \mathcal{V}$ and corresponding reference state x^* such that the network will converge to the desired equilibrium state.

This problem is computationally complex to solve in general. However, some special network topologies admit exact analytical solutions for certain cases. For example, the results shown in Table 1 hold for two-strategy games on completely connected networks of homogeneous imitating agents when $x(0) = \mathbf{B}$ and $x^* = \mathbf{A}$ [42], which depend on the values of $\delta^B = d - b$ and $\delta^A = a - c$.

Similar results exist for star and ring networks under the same assumptions[42]. For tree networks, it is possible to bound and approximate the solution to this case of Problem 1 using an algorithm that traverses the tree structure from leaves to root [79]. A modification of this algorithm based on minimum spanning trees can also be used to approximate the solution on arbitrary networks for both deterministic and stochastic imitation dynamics [80, 42]. It has also been shown in a simulation study that controlling agents with higher degree results in a higher frequency of cooperation than randomly controlling agents in a prisoner's dilemma game under imitative dynamics on scale-free networks [41].

An alternative framework was presented in [81] for the direct-strategy control of synchronous networked evolutionary games using large-scale logical dynamic networks to model transitions between all possible strategy states for various update rules. Equivalent conditions for reachability and stability are given in this work for a particular set of control nodes. Using this general framework, a method was proposed in [82] for how to guide a

Table 1: Minimum number of direct control agents for to guide a network of completely connected agents who are playing strategy A to all play strategy B .

$\delta^A + \delta^B > 0$	$ \mathcal{L}^* = \begin{cases} \min\left(\lfloor 1 + \frac{\delta^B}{\delta^A + \delta^B} \rfloor, n\right), & \text{if } \delta^B \geq 0 \\ 1, & \text{if } \delta^B < 0 \end{cases}$
$\delta^A + \delta^B = 0$	$ \mathcal{L}^* = \begin{cases} n, & \text{if } \delta^B \geq 0 \\ 1, & \text{if } \delta^B < 0 \end{cases}$
$\delta^A + \delta^B < 0$	$ \mathcal{L}^* = \begin{cases} n, & \text{if } \delta^A \geq 0 \text{ or } \\ & \frac{\delta^B}{\delta^A + \delta^B} < n \text{ and } (\delta^A + \delta^B) \bmod \delta^B = 0 \\ 1, & \text{otherwise} \end{cases}$

network using best-response updates to an optimal state by adding a new player whose strategy can be controlled directly. Similarly in [83], the deterministic network dynamics were controlled by actively assigning the strategies of a group of agents. However, since the underlying logical matrices have dimension m^n , this line of analysis can be intractable for large networks.

4.2.3. Incentive-based control

One potentially effective policy-based control mechanism is to offer incentives to agents for playing a particular strategy. Similar to the approach in Section 4.1, but in a discrete time setting, this is equivalent to adding some positive value to the row in a payoff matrix corresponding to the desired strategy X :

$$\tilde{\pi}_{X,Y}^i := \pi_{X,Y}^i + r \text{ for all } Y \in \mathcal{S}.$$

Although this approach is most intuitive for use with best-response dynamics, it can be applied to networks of imitating agents as well. In the latter case, adding to the payoffs of an agent does not act as an “incentive” in that it does not directly affect the decision of that agent. Rather, neighbors of the agent notice that agents who play the rewarded strategy earn more total payoff and are therefore more likely to imitate them. We consider two cases under this framework: (i) uniform rewards, when the same reward is offered to all agents in the network, and (ii) targeted rewards, when different rewards can be offered to different agents.

Uniform incentive-based control. In this case, the same incentive r_0 is offered to every agent in the network:

$$\tilde{\pi}_{X,Y}^i := \pi_{X,Y}^i + r_0 \text{ for all } i \in \mathcal{V} \text{ and } Y \in \mathcal{S}.$$

Problem 2. Given initial equilibrium state x^* , find the minimum value of r_0 such that the network will converge to the equilibrium in which all agents play X .

It turns out that this problem can be solved exactly and efficiently provided that a network is both monotone and uniquely convergent in the desired strategy, which holds for both best response and imitation dynamics. In both cases, the solution involves computing a set of all candidate rewards and then performing a binary search over this finite set.

For networks of coordinating agents using best-response updates, the candidate set is given by:

$$\mathcal{R}_I := \left\{ r \geq \gamma_{\max} \mid r = \frac{\delta_i(\tilde{n}_i^A - j)}{\deg_i}, i \in \mathcal{V}, j \in \{1, \dots, \tilde{n}_i^A\} \right\}.$$

where

$$\gamma_{\max} = \begin{cases} \max_{i \in \bar{\mathcal{B}}} \delta_i^B & \bar{\mathcal{B}} \neq \emptyset \\ 0 & \bar{\mathcal{B}} = \emptyset \end{cases}, \quad \delta_i := \delta_i^A + \delta_i^B,$$

and $\bar{\mathcal{B}} = \{i \mid \delta_i = 0, x_i(0) = B\}$ denotes a set of *stubborn* agents whose unaltered threshold is such that no number of A -playing neighbors would cause them to switch from B to A [78].

For networks of opponent-coordinating agents using imitation dynamics, since no agent will switch from A to

B , the possible payoffs of an agent i when playing A (resp. B) at any time step are contained in the sets

$$\begin{aligned}\Pi_i^A &:= \{a_i(n_i^A + \omega_i) + b_i(\deg_i - n_i^A - \omega_i) : \omega_i \in \Omega_i\} \\ \Pi_i^B &:= \{c_i(n_i^A + \omega_i) + d_i(\deg_i - n_i^A - \omega_i) : \omega_i \in \Omega_i\},\end{aligned}$$

where $\Omega_i = \{0, 1, \dots, \deg_i - n_i^A\}$.

Now consider an agent s who initially plays B and has a neighbor j whose strategy was either initially A or became A at some other time. In order for agent j to cause agent s to switch to A , the payoff of agent j must be greater than that of each B -playing agent (denoted by i) in the neighborhood of agent s . Therefore, the reward given to agent j must be greater than $\frac{u_i^B - u_j^A}{\deg_j}$ for some $u_i^B \in \Pi_i^B$ and $u_j^A \in \Pi_j^A$. This leads to the following set of all candidate infimum rewards [84].

$$\mathcal{R}_B := \left\{ \frac{u_i^B - u_j^A}{\deg_j} \mid u_i^B \in \Pi_i^B, u_j^A \in \Pi_j^A, j \in \mathcal{N}_s, i \in \mathcal{N}_s \cup \{s\}, x_i(0) = B, s \in \mathcal{V}, x_s(0) = B \right\} \cup \{0\}.$$

For both imitation and best-response, we then sort the elements of $\mathcal{R}_I$ or $\mathcal{R}_B$ in increasing order and denote this vector by $v^{\mathcal{R}}$. Algorithm 1 then performs a binary search over $v^{\mathcal{R}}$ to find the infimum reward such that all agents in the network will eventually play A . Denote by $\mathbf{1}$ the n -dimensional vector containing all ones. In what follows, we also denote by $\bar{x}$ the unique equilibrium resulting from a particular set of incentives being offered to a network of A -coordinating agents starting from x .

```

 $i^- := 1$ 
 $i^+ := |\mathcal{R}|$ 
while  $i^+ - i^- > 1$  do
     $r_0^* := v_j^{\mathcal{R}}$ , where  $j := \lceil \frac{i^- + i^+}{2} \rceil$ 
     $\hat{\Gamma} := (\mathbb{G}, \hat{\pi}, f)$ 
    Simulate  $\hat{\Gamma}$  from  $x(0)$  until equilibrium  $\bar{x}$ 
    if  $\bar{x}_i = A$  for all  $i \in \mathcal{V}$  then
         $i^+ := j$ 
    else
         $i^- := j$ 
    end
end

```

Algorithm 1: Binary search over candidate rewards to find the value of r_0^* that solves Problem 2 for imitative networks of opponent-coordinating agents ($\mathcal{R} := \mathcal{R}_I$) and best-response networks of coordinating agents ($\mathcal{R} := \mathcal{R}_B$)

Exact solutions for cases other than coordination remain an open area for investigation, which is likely to be challenging since the guarantee of unique convergence no longer holds.

Targeted incentive-based control. Now suppose that the policy makers have the ability to offer different incentives to different agents in the network:

$$\bar{\pi}_{X,Y}^i = \pi_{X,Y}^i + r_i \text{ for all } i \in \mathcal{V} \text{ and } Y \in \mathcal{S}.$$

Problem 3. Given initial equilibrium state x^* , find $\mathbf{r} := (r_1, \dots, r_n)^T$ with minimum sum such that the network will converge to the equilibrium in which all agents play X .

Problem 4. Given initial equilibrium state x^* and budget constraint ρ , find $\mathbf{r} := (r_1, \dots, r_n)^T$ such that $\sum_{i=1}^n r_i \leq \rho$ that maximizes the number of agents who converge to X .

Solutions to these problems are more difficult to compute for general networks since the outcome of offering rewards strongly depends on which agents are offered rewards. A useful quantity to know is how much reward is required to cause some agent in the network to switch to the desired strategy. We will denote this quantity by $\tilde{r}_i$, which takes different values depending on the update rule. For networks of opponent-coordinating agents under

imitation dynamics, the infimum reward such that at least one B -playing neighbor will switch to A is

$$\check{r}_i = \max_{j \in \mathcal{N}_i^B} \max_{k \in \mathcal{N}_j^B} u_k(B, \bar{x}) - \bar{u}_i(A, \bar{x}), \quad (14)$$

when the state of the network is $\bar{x}$, where $\mathcal{N}_i^B := \{j \in \mathcal{N}_i \cup \{i\} : \bar{x}_j = B\}$ denotes the self-inclusive set of neighbors of agent i who are playing B [84].

For networks of coordinating agents using best-response updates, the minimum incentive that will cause an agent to switch from B to A depends on the agent payoff matrices and degree in the network. In particular, the following condition must be satisfied for some agent i :

$$\delta_i n_i^A(t_e) = (\gamma_i - \check{r}_i) \deg_i \Rightarrow \check{r}_i = \gamma_i - \frac{\delta_i n_i^A(t_e)}{\deg_i}.$$

Problems 3 and 4 now reduce to determining which agents to offer rewards to and in what order. Some potentially effective heuristics include targeting those agents who cause the most subsequent switches to the desired strategy or those who require the smallest rewards. The following algorithm can be used to iteratively implement any such heuristic.

Initialize $\bar{x}_i = x_i(0)$ and $r_i = 0$ for each $i \in \mathcal{V}$

while $\exists i \in \mathcal{V} : \bar{x}_i \neq A$ and $\sum_{i \in \mathcal{V}} r_i < \rho$ **do**

$\mathcal{A}^B := \{i \in \mathcal{V} : \bar{x}_i = A \wedge [\exists j \in \mathcal{N}_i : \bar{x}_j = B]\} \wedge \check{r}_i \leq \rho - \sum_{i \in \mathcal{V}} r_i\}$

 Choose an agent $j \in \mathcal{A}^B$

 Let $r_j := r_j + \check{r}_j + \epsilon$

$\hat{\Gamma} := (\mathbb{G}, \hat{\pi}, f)$

 Simulate $\hat{\Gamma}$ from $\bar{x}$ until new equilibrium $\bar{x}'$

$\bar{x} := \bar{x}'$

end

Algorithm 2: Generic iterative algorithm for computing a reward vector to approximate the maximum number of agents who will play strategy A . Since r_j represents an infimum, an arbitrarily small additional payoff ϵ is added to ensure that the reward causes an agent to switch.

Here we propose a weighted ratio of the two factors mentioned above, which, when implemented in Algorithm 2, we call Iterative Potential-to-Reward Optimization (IPRO). Consider the following simple potential function

$$\Phi(x) = \sum_{i=1}^n n_i^A(x), \quad (15)$$

where $n_i^A(x)$ denotes the number of neighbors of agent i who play strategy A in the state x . The ratio we wish to iteratively maximize is the following:

$$j^* = \arg \max_{j \in \mathcal{A}^B} \frac{\Delta \Phi(\bar{x})^\alpha}{\check{r}_j^\beta}, \quad (16)$$

where $\Delta \Phi(\bar{x}) := \Phi(\bar{x}') - \Phi(\bar{x})$, and $\alpha, \beta \geq 0$ are free design parameters. Although generally suboptimal, simulations have shown that the IPRO algorithm outperforms several other agent-targeting heuristics and performs quite close to optimal for a range of conditions on the network topology and payoff matrices, both for imitation [84] and best-response dynamics [78]. The approach also seems to be effective for networks containing anti-coordinating agents [78], although care should be taken because there is no guarantee of unique convergence in that case.

5. Open Problems

We summarize here some outstanding research problems discussed in this survey as well as some additional ones, categorized by dynamical regime.

5.1. Infinite well-mixed populations

For infinite well-mixed populations, completing the gap between game-theoretical concepts such as evolutionary stability and dynamical concepts such as asymptotic stability is in particularly high demand. With regard to

control, although progress has been made toward stabilizing existing equilibrium points in the replicator dynamics [68], stabilization of arbitrary setpoints of the replicator and other dynamics through payoff augmentation remains a challenging open research problem.

5.2. Finite structured populations

In this survey, we have presented convergence results for networks consisting of all coordinating or all anti-coordinating agents under best-response dynamics and all opponent-coordinating agents under imitation dynamics. These results assume that updates are asynchronous and deterministic and they provide only sufficient conditions. There are very likely to be cases when networks that do not satisfy these conditions will nevertheless reach an equilibrium, such as mixtures of coordinating and anti-coordinating agents or mixtures of imitative and best-response agents, but the conditions for convergence under such conditions remain to be discovered. Moreover, issues of convergence time and uniqueness of equilibria also remain open. Useful notions of stability, which are prevalent in the continuous-time case of infinite populations, remain elusive in the context of discrete-time finite populations.

As a second primary focus, we discussed some approaches for guiding networks of agents who are “coordinating in X ” toward agreement in a desired strategy X . It would be very appealing to broaden the conditions under which optimal solutions could be obtained for Problems 3 and 4, a great challenge due to the complexity involved. Moreover, these results assume agents update asynchronously and deterministically, the network topology is fixed (or at least changes slowly compared to agent updates), and the agents have access to all relevant strategy and payoff information about themselves and their immediate neighbors. Relaxations of these assumptions lead to a wide array of interesting open problems. Controllability and practical notions of feedback control of such networks are other intriguing topics for future study.

6. Concluding Remarks

Evolutionary game theory provides both a compelling framework for studying the dynamics of interacting agents across various scientific domains and a rich set of mathematical problems. For the control engineer, this presents numerous opportunities to apply classical tools or to design new ones towards a greater understanding of the convergence, stability, and control of such complex systems. Significant progress in this field could ultimately help to solve some of society’s most pressing problems, including the control of viral epidemics, preservation of the environment, stabilization of financial markets, and treatment of brain disorders. We hope that curious researchers have gained insight and perhaps even found inspiration in this article, to apply these tools, tackle open problems, and generally further our collective knowledge in this interesting field of study.

References

- [1] W. H. Sandholm, *Population Games and Evolutionary Dynamics*, MIT Press, 2010.
- [2] W. Yu, P. DeLellis, G. Chen, M. Di Bernardo, J. Kurths, Distributed adaptive control of synchronization in complex networks, *IEEE Transactions on Automatic Control* 57 (8) (2012) 2153–2158.
- [3] R. Olfati-Saber, R. M. Murray, Consensus problems in networks of agents with switching topology and time-delays, *IEEE Transactions on automatic control* 49 (9) (2004) 1520–1533.
- [4] D. Madeo, C. Mocenni, Game interactions and dynamics on networked populations, *IEEE Transactions on Automatic Control* 60 (7) (2015) 1801–1810.
- [5] X. Chen, J. Liu, M.-A. Belabbas, Z. Xu, T. Başar, Distributed evaluation and convergence of self-appraisals in social networks, *IEEE Transactions on Automatic Control* 62 (1) (2017) 291–304.
- [6] J. M. Smith, G. R. Price, The logic of animal conflict, *Nature* 246 (5427) (1973) 15.
- [7] J. W. Weibull, *Evolutionary Game Theory*, MIT Press, 1997.
- [8] P. Ramazi, M. Cao, F. J. Weissing, Evolutionary dynamics of homophily and heterophily, *Scientific reports* 6 (2016) 22766.
- [9] M. A. Nowak, C. E. Tarnita, T. Antal, Evolutionary dynamics in structured populations, *Philosophical Transactions of the Royal Society B: Biological Sciences* 365 (1537) (2010) 19–30.
- [10] H. Ohtsuki, M. A. Nowak, The replicator equation on graphs, *Journal of theoretical biology* 243 (1) (2006) 86–97.
- [11] K. H. Schlag, Why imitate, and if so, how?: A boundedly rational approach to multi-armed bandits, *Journal of Economic Theory* 78 (1) (1998) 130–156.
- [12] B. Ghahesifard, B. Touri, T. Başar, J. Shamma, On the convergence of piecewise linear strategic interaction dynamics on networks, *IEEE Transactions on Automatic Control* 61 (6) (2016) 1682–1687.
- [13] P. Ramazi, J. Riehl, M. Cao, Networks of conforming or nonconforming individuals tend to reach satisfactory decisions, *Proceedings of the National Academy of Sciences* 113 (46) (2016) 12985–12990.
- [14] J. S. Shamma, G. Arslan, Unified convergence proofs of continuous-time fictitious play, *IEEE Transactions on Automatic Control* 49 (7) (2004) 1137–1141.
- [15] A. Cortés, S. Martínez, Self-triggered best-response dynamics for continuous games, *IEEE Transactions on Automatic Control* 60 (4) (2015) 1115–1120.
- [16] O. N. Ghahreghshiran, V. Krishnamurthy, G. Yin, Adaptive search algorithms for discrete stochastic optimization: A smooth best-response approach, *IEEE Transactions on Automatic Control* 62 (1) (2017) 161–176.

- [17] R. Axelrod, W. D. Hamilton, The evolution of cooperation, *science* 211 (4489) (1981) 1390–1396.
- [18] M. A. Nowak, Five rules for the evolution of cooperation, *science* 314 (5805) (2006) 1560–1563.
- [19] G. Szabó, G. Fáth, Evolutionary games on graphs, *Physics Reports* 446 (4) (2007) 97–9216.
- [20] N. Li, J. R. Marden, Designing games for distributed optimization, *IEEE Journal of Selected Topics in Signal Processing* 7 (2) (2013) 230–242.
- [21] N. Quijano, C. Ocampo-Martinez, J. Barreiro-Gomez, G. Obando, A. Pantoja, E. Mojica-Nava, The role of population games and evolutionary dynamics in distributed control systems: The advantages of evolutionary game theory, *IEEE Control Systems* 37 (1) (2017) 70–97.
- [22] J. S. Shamma, G. Arslan, Dynamic fictitious play, dynamic gradient play, and distributed convergence to nash equilibria, *IEEE Transactions on Automatic Control* 50 (3) (2005) 312–327.
- [23] E. Ramírez-Llanos, N. Quijano, A population dynamics approach for the water distribution problem, *International Journal of Control* 83 (9) (2010) 1947–1964.
- [24] J. Barreiro-Gomez, C. Ocampo-Martinez, N. Quijano, Dynamical tuning for mpc using population games: A water supply network application, *ISA transactions* 69 (2017) 175–186.
- [25] H. Tembine, E. Altman, R. El-Azouzi, Y. Hayel, Evolutionary games in wireless networks, *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 40 (3) (2010) 634–646.
- [26] L. Pavel, *Game theory for control of optical networks*, Springer Science & Business Media, 2012.
- [27] E. Altman, R. El-Azouzi, Y. Hayel, H. Tembine, The evolution of transport protocols: An evolutionary game perspective, *Computer Networks* 53 (10) (2009) 1751–1759.
- [28] J. R. Marden, J. S. Shamma, Game theory and distributed control, in: *Handbook of game theory with economic applications*, Vol. 4, Elsevier, 2015, pp. 861–899.
- [29] C. Wang, X. Chen, G. Xie, M. Cao, Emergence of leadership in a robotic fish group under diverging individual personality traits, *Royal Society open science* 4 (5) (2017) 161015.
- [30] H. Ackermann, P. Berenbrink, S. Fischer, M. Hoefer, Concurrent imitation dynamics in congestion games, *Distributed Computing* 29 (2) (2016) 105–125.
- [31] W. Poundstone, *Prisoner’s Dilemma/John von Neumann, Game Theory and the Puzzle of the Bomb*, Anchor, 1993.
- [32] M. Mäs, H. H. Nax, A behavioral study of noise in coordination games, *Journal of Economic Theory* 162 (2016) 195–208.
- [33] M. Granovetter, Threshold models of collective behavior, *American Journal of Sociology* (1978) 1420–1443.
- [34] A. Montanari, A. Saberi, The spread of innovations in social networks, *Proceedings of the National Academy of Sciences* 107 (47) (2010) 20196–20201.
- [35] J. Kleinberg, Cascading behavior in networks: Algorithmic and economic issues, *Algorithmic game theory* 24 613–632.
- [36] J. J. Hopfield, Neural networks and physical systems with emergent collective computational abilities, *Proceedings of the national academy of sciences* 79 (8) (1982) 2554–2558.
- [37] P. D. Taylor, L. B. Jonker, Evolutionary stable strategies and game dynamics, *Mathematical Biosciences* 40 (1-2) (1978) 145–156.
- [38] A. Traulsen, D. Semmann, R. D. Sommerfeld, H.-J. Krambeck, M. Milinski, Human strategy updating in evolutionary games, *Proceedings of the National Academy of Sciences* 107 (7) (2010) 2962–2966.
- [39] P. van den Berg, L. Molleman, F. J. Weissing, Focus on the success of others leads to selfish behavior, *Proceedings of the National Academy of Sciences* 112 (9) (2015) 2912–2917.
- [40] L. Zino, G. Como, F. Fagnani, On imitation dynamics in potential population games, in: *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, 2017, pp. 757–762.
- [41] W. Du, H. Zhou, Z. Liu, X. Cao, The effect of pinning control on evolutionary prisoner’s dilemma game, *Modern Physics Letters B* 24 (25) (2010) 2581–2589.
- [42] J. R. Riehl, M. Cao, Towards optimal control of evolutionary games on networks, *IEEE Transactions on Automatic Control* 62 (1) (2017) 458–462.
- [43] P. Ramazi, *Convergence Analysis and Control of Evolutionary Matrix-game Dynamics*, University of Groningen, 2017.
- [44] J. Hofbauer, On the occurrence of limit cycles in the Volterra-Lotka equation, *Nonlinear Analysis: Theory, Methods & Applications* 5 (9) (1981) 1003–1007.
- [45] I. M. Bomze, Lotka-Volterra equation and replicator dynamics: new issues in classification, *Biological Cybernetics* 72 (5) (1995) 447–453.
- [46] I. M. Bomze, Lotka-Volterra equation and replicator dynamics: a two-dimensional classification, *Biological Cybernetics* 48 (3) (1983) 201–211.
- [47] F. J. Weissing, Evolutionary stability and dynamic stability in a class of evolutionary normal form games, in: *Game Equilibrium Models I*, Springer, 1991, pp. 29–97.
- [48] E. D. W. H. Sandholm, F. Franchetti, *Dynamo: Diagrams for evolutionary game dynamics* (2012).
URL <http://www.ssc.wisc.edu/~whs/dynamo>
- [49] I. Gilboa, A. Matsui, Social stability and equilibrium, *Econometrica: Journal of the Econometric Society* (1991) 859–867.
- [50] J. Hofbauer, Stability for the best response dynamics, Unpublished manuscript, University of Vienna.
- [51] A. Matsui, Best response dynamics and socially stable strategies, *Journal of Economic Theory* 57 (2) (1992) 343–362.
- [52] D. Balkenborg, J. Hofbauer, C. Kuzmics, Refined best reply correspondence and dynamics, *Theoretical Economics* 8 (1) (2013) 165–192.
- [53] P. Ramazi, H. Jardón-Kojakhmetov, M. Cao, Limit sets within curves where trajectories converge to, *Applied Mathematics Letters* 68 (2017) 94–100.
- [54] J. Hofbauer, K. Sigmund, *The Theory of Evolution and Dynamical Systems*, Cambridge University Press, Cambridge UK, 1988.
- [55] I. M. Bomze, Non-cooperative two-person games in biology: A classification, *International Journal of Game Theory* 15 (1) (1986) 31–57.
- [56] J. Hofbauer, P. Schuster, K. Sigmund, A note on evolutionary stable strategies and game dynamics, *Journal of Theoretical Biology* 81 (3) (1979) 609–612.
- [57] E. C. Zeeman, *Population dynamics from game theory*, Springer Berlin Heidelberg, Berlin, Heidelberg, 1980, pp. 471–497.
- [58] R. A. Fisher, *The Genetical Theory of Natural Selection: A Complete Variorum Edition*, Oxford University Press, 1930.
- [59] V. Losert, E. Akin, Dynamics of games and genes: Discrete versus continuous time, *Journal of Mathematical Biology* 17 (2) (1983) 241–251.
- [60] J. H. Nachbar, evolutionary selection dynamics in games: Convergence and limit properties, *International Journal of Game Theory* 19 (1) (1990) 59–89.

- [61] J. Hofbauer, A general cooperation theorem for hypercycles, *Monatshefte für Mathematik* 91 (3) (1981) 233–240.
- [62] R. Cressman, Y. Tao, The replicator equation and other game dynamics, *Proceedings of the National Academy of Sciences* 111 (Supplement 3) (2014) 10810–10817.
- [63] J. Hofbauer, K. Sigmund, Evolutionary game dynamics, *Bulletin of the American Mathematical Society* 40 (4) (2003) 479–519.
- [64] J. Barreiro-Gomez, G. Obando, N. Quijano, Distributed population dynamics: Optimization and control applications, *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 47 (2) (2017) 304–314.
- [65] A. Pantoja, N. Quijano, Distributed optimization using population dynamics with a local replicator equation, in: *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*, IEEE, 2012, pp. 3790–3795.
- [66] P. Ramazi, J. Riehl, M. Cao, Imitating successful neighbors can hinder reaching satisfactory decisions for networks of decision-making agents, under review.
- [67] T. Kanazawa, Y. Fukumoto, T. Ushio, T. Misaka, Replicator dynamics with pigovian subsidy and capitation tax, *Nonlinear Analysis: Theory, Methods & Applications* 71 (12) (2009) e818–e826.
- [68] T. Kanazawa, T. Misaka, T. Ushio, Y. Fukumoto, A control method of selfish routing based on replicator dynamics with capitation tax and subsidy, in: *IEEE CCA & ISIC*, 2009, pp. 249–254.
- [69] T. Morimoto, T. Kanazawa, T. Ushio, Subsidy-based control of heterogeneous multiagent systems modeled by replicator dynamics, *IEEE Transactions on Automatic Control* 61 (10) (2016) 3158–3163.
- [70] W. H. Sandholm, Evolutionary implementation and congestion pricing, *The Review of Economic Studies* 69 (3) (2002) 667–689.
- [71] W. H. Sandholm, Pigouvian pricing and stochastic evolutionary implementation, *Journal of Economic Theory* 132 (1) (2007) 367–382.
- [72] A. Khamfer, T. Başar, B. Ghahsifard, Stability properties of infected networks with low curing rates, in: *American Control Conference (ACC)*, 2014, IEEE, 2014, pp. 3579–3584.
- [73] H. Liang, M. Cao, X. Wang, Analysis and shifting of stochastically stable equilibria for evolutionary snowdrift games, *Systems & Control Letters* 85 (2015) 16–22.
- [74] W. Mei, N. E. Friedkin, K. Lewis, F. Bullo, Dynamic models of appraisal networks explaining collective learning, *IEEE Transactions on Automatic Control*.
- [75] E. Ramírez-Llanos, S. Martínez, A distributed dynamics for virus-spread control, *Automatica* 76 (2017) 41–48.
- [76] J. I. Poveda, N. Quijano, Shahshahani gradient-like extremum seeking, *Automatica* 58 (2015) 51–59.
- [77] P. Ramazi, M. Cao, Stability analysis for replicator dynamics of evolutionary snowdrift games, in: *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*, IEEE, 2014, pp. 4515–4520.
- [78] J. Riehl, P. Ramazi, M. Cao, Control of asynchronous imitation dynamics on networks, *arXiv preprint arXiv:1704.04416*.
- [79] J. R. Riehl, M. Cao, Minimal-agent control of evolutionary games on tree networks, in: *The 21st International Symposium on Mathematical Theory of Networks and Systems*, 2014.
- [80] J. R. Riehl, M. Cao, Control of stochastic evolutionary games on networks, *IFAC-PapersOnLine* 48 (22) (2015) 76–81.
- [81] D. Cheng, F. He, H. Qi, T. Xu, Modeling, analysis and control of networked evolutionary games, *IEEE Transactions on Automatic Control* 60 (9) (2015) 2402–2415.
- [82] P. Guo, Y. Wang, H. Li, Algebraic formulation and strategy optimization for a class of evolutionary networked games via semi-tensor product method, *Automatica* 49 (11) (2013) 3384–3389.
- [83] B. Zhu, X. Xia, Z. Wu, Evolutionary game theoretic demand-side management and control for a class of networked smart grid, *Automatica* 70 (2016) 94–100.
- [84] J. Riehl, P. Ramazi, M. Cao, Control of asynchronous imitation dynamics on networks, *arXiv preprint arXiv:1704.04416*.
- [85] M. A. Nowak, *Evolutionary Dynamics*, Harvard University Press, 2006.

Appendix A. 2-player matrix games

A 2-player matrix game comprises two players, usually referred to as *player 1* and *player 2*, two sets of pure strategies $\mathcal{S}_1 = \{1, \dots, m_1\}$ and $\mathcal{S}_2 = \{1, \dots, m_2\}$ from which players 1 and 2 choose strategies, and two payoff matrices $\pi^1, \pi^2 \in \mathbb{R}^{m_1 \times m_2}$ according to which players 1 and 2 earn payoffs, respectively. More specifically, when player 1 plays strategy $i \in \mathcal{S}_1$ and player 2 plays strategy $j \in \mathcal{S}_2$, the payoff of the first and second players will be π_{ij}^1 and π_{ji}^2 , respectively. In this paper, we focus on matrix games that provide the same set of strategies for both players, i.e., $\mathcal{S} = \{1, 2, \dots, m\}$. The game is *symmetric* if $\pi^1 = \pi^2$.

The players may be uncertain on their choices of strategies. For example, a player may choose strategy 1 with probability 0.9 and strategy 2 with probability 0.1. This results in the so called *mixed strategy* which is defined as a vector x belonging to the simplex Δ defined by

$$\Delta := \left\{ x \in \mathbb{R}^m \mid \sum_{i \in \mathcal{S}} x_i = 1, \quad x_i \geq 0 \quad \forall i \in \mathcal{S} \right\}.$$

Namely, if a player plays the mixed strategy x , she will play each *pure strategy* $i \in \mathcal{S}$ with probability $x_i \in [0, 1]$, where $\sum_{i \in \mathcal{S}} x_i = 1$. So if players 1 and 2 play the strategies $x, y \in \Delta$, they will receive the payoffs $u_1(x, y) = x^\top \pi^1 y$ and $u_2(y, x) = y^\top \pi^2 x$, respectively, where $u_i : \Delta \times \Delta \rightarrow \mathbb{R}, i \in \{1, 2\}$ are the *utility functions* of players 1 and 2. To be precise, one should distinguish the case when x equals a unit vector $e^i, i \in \mathcal{S}$, which is the i th column of the $m \times m$ identity matrix, since then x refers to the pure strategy i . For this reason, the unit vector e^i is called a *pure strategy vector*, which, with a minor abuse of notation, may also be referred to as the pure strategy i .

Appendix B. Nash equilibrium

Players' strategies are said to be at a Nash equilibrium if no player would earn a higher payoff by changing their strategy. Particularly, for 2-player games, we have the following definition.

Definition 4. Consider a 2-player game with payoff matrices $\pi^1, \pi^2 \in \mathbb{R}^{2 \times 2}$ and utility functions u_1 and u_2 . A pair of strategies (x, y) , where $x, y \in \Delta$, is a Nash equilibrium if

$$\begin{aligned} u_1(x, y) &\geq u_1(z, y) & \forall z \in \Delta, \\ u_2(y, x) &\geq u_2(z, x) & \forall z \in \Delta. \end{aligned}$$

A *strict Nash equilibrium* is a Nash equilibrium that satisfies both of the above conditions with strict inequality. For *symmetric games*, we further have the definition of a *symmetric Nash equilibrium* where the same strategy is played by both players. We call this same strategy, which is in Nash equilibrium with itself, a *Nash strategy* and formally define it as follows.

Definition 5. Consider a 2-player symmetric game with payoff matrix π and utility function $u = u_1 = u_2$, where $u(x, y) = x^\top \pi y$ for $x, y \in \Delta$. A strategy $x \in \Delta$ is a Nash strategy if

$$u(x, x) \geq u(z, x) \quad \forall z \in \Delta. \quad (\text{B.1})$$

The set of all Nash strategies of the game is denoted by Δ^{NE} .

A *strict Nash strategy* is a Nash strategy that satisfies (B.1) with strict inequality.

Another refined definition of a Nash equilibrium is a *perfect Nash equilibrium* which is a Nash equilibrium that is robust against some 'trembles' in the players' strategies. Particularly, for symmetric games, we have the following definition [7]. Consider a symmetric game G and an *error vector* $\mu \in \Delta$ where $\mu_i < 1$ for $i \in \mathcal{S}$. Define the *perturbed game* $G(\mu)$ as the game G but when each pure strategy $i \in \mathcal{S}$ is played "by mistake" with probability μ_i by both players; namely, the strategies of both players belong to $\Delta(\mu) = \{z \in \Delta \mid z_i \geq \mu_i \forall i \in \mathcal{S}\}$. Denote by $\Delta^{NE}(\mu)$ the set of Nash strategies of the perturbed game $G(\mu)$.

Definition 6. A strategy $x \in \Delta^{NE}$ is a *perfect Nash strategy* if for some sequence $\{G(\mu^t)\}_{\mu^t \rightarrow 0}$ of perturbed games, $\mu^t \in \Delta$, there exist Nash strategies $x^t \in \Delta^{NE}(\mu^t)$ such that $x^t \rightarrow x$. The set of all perfect Nash strategies of the game is denoted by Δ^{PNE} .

It can be shown that every Nash strategy x belonging to the interior of the simplex Δ is perfect.

Appendix C. Categorization of 2×2 matrix games

Symmetric matrix games in which both players have two strategies, typically referred to as *cooperation* (C) and *defection* (D), are widely studied in the context of evolutionary dynamics. These matrix games can be divided into the following four main categories based on the relationship between the payoffs in the matrix

$$\pi = \begin{matrix} & \begin{matrix} C & D \end{matrix} \\ \begin{matrix} C \\ D \end{matrix} & \begin{pmatrix} R & S \\ T & P \end{pmatrix} \end{matrix}, \quad R, S, T, P \in \mathbb{R}.$$

Prisoner's Dilemma. In the prisoner's dilemma game, described in Section 2, the payoffs must satisfy $T > R$ and $P > S$. The state in which both players defect is a unique Nash equilibrium. In addition, it is usually required that $R > P$ to ensure that mutual cooperation is preferable to mutual defection. Otherwise, it is simply a game in which defection is the dominant strategy.

Anti-coordination, Snowdrift, Hawk-Dove, or Chicken. In an anti-coordination game, players benefit when they play the same strategy as their opponents. In this case, the payoffs should satisfy $T > R$ and $S > P$, and there exist two pure strategy Nash Equilibria, which correspond to one player cooperating and the other defecting. Other versions of the game, such as snowdrift, are often considered less generic in the literature by enforcing the additional requirement of $R > S$.

Coordination, Stag-Hunt, or Battle of The Sexes. . In a coordination game, players benefit when their opponent plays the same strategy as themselves. In this case, the payoffs should satisfy $R > T$ and $P > S$, and there exist two Nash Equilibria corresponding to mutual cooperation and mutual defection. Games such as stag-hunt, which are specific cases of the coordination game, also require the condition $T > P$ to be in force.

Harmony. . The harmony game is similar to the prisoner's dilemma game, except that cooperation always yields the higher payoff. That is, the payoffs should satisfy $R > T$ and $S > P$, and the only Nash Equilibrium corresponds to mutual cooperation.

Remark: Versions of these games in which one or more inequalities is substituted by equality have varying interpretations. For example, if $R = T$, then cooperation remains a pure Nash equilibrium strategy, but other properties such as evolutionary stability depend on the ordering of the other payoffs (see Appendix E).

Appendix D. Analyses of the replicator dynamics for 2×2 games

Under the payoff matrix

$$\pi = \begin{pmatrix} R & S \\ T & P \end{pmatrix},$$

the RD for x_1 is

$$\dot{x}_1 = [Rx_1 + Sx_2 - Rx_1^2 - (S + T)x_1x_2 - Px_2^2]x_1,$$

which by using the identity $x_1 + x_2 = 1$, becomes

$$\dot{x}_1 = [((P - S) + (R - T))x_1 - (P - S)]x_1(1 - x_1).$$

The dynamics of x_2 is the opposite of x_1 , i.e., $\dot{x}_2 = -\dot{x}_1$. Now we proceed to each type of game individually.

Prisoner's dilemma game: $(P - S) > 0 > (R - T)$. Hence, $((P - S) + (R - T))x_1 < (P - S)$. So $\lim_{t \rightarrow \infty} x_1(t) = 0$, implying that

$$\lim_{t \rightarrow \infty} x(t) = \begin{cases} e^1 & x(0) = e^1 \\ e^2 & x(0) \in \Delta - \{e^1\} \end{cases}.$$

Therefore, almost always the dynamics converge to the state of full defectors.

Coordination game: $(P - S), (R - T) > 0$. Hence, $((P - S) + (R - T))x_1 > (P - S)$ for $x_1 > x_1^*$, and $((P - S) + (R - T))x_1 < (P - S)$ for $x_1 < x_1^*$ where $\begin{bmatrix} x_1^* & x_2^* \end{bmatrix} = \begin{bmatrix} \frac{P-S}{P-S+R-T} & \frac{R-T}{P-S+R-T} \end{bmatrix}^T$ is the interior equilibrium point of the dynamics. So

$$\lim_{t \rightarrow \infty} x(t) = \begin{cases} e^1 & x(0) \in \{x \in \Delta \mid x_1 \in (x_1^*, 1]\} \\ x^* & x(0) = x^* \\ e^2 & x(0) \in \{x \in \Delta \mid x_1 \in [0, x_1^*)\} \end{cases}.$$

Therefore, almost always the dynamics converge to the states of either full cooperators or full defectors, implying that the individuals will become *coordinated* in the long run.

Anti-coordination game: $(P - S), (R - T) < 0$. Hence, $((P - S) + (R - T))x_1 > (P - S)$ for $x_1 < x_1^*$, and $((P - S) + (R - T))x_1 < (P - S)$ for $x_1 > x_1^*$. So

$$\lim_{t \rightarrow \infty} x(t) = \begin{cases} e^1 & x(0) = e^1 \\ x^* & \Delta - \{e^1, e^2\} \\ e^2 & x(0) = e^2 \end{cases}.$$

Therefore, almost always the dynamics converge to the unique interior equilibrium that is a mixture of cooperation and defection, a *polymorphic population*.

Appendix E. Evolutionary stability and neutral stability

One fundamental notion in evolutionary game theory dealing with invasion of mutants in large populations is *evolutionary stability*. Consider a population of individuals who are repeatedly and randomly drawn to play with each other a symmetric 2-player game with the utility function u which defines their fitnesses. Suppose that initially all individuals play the same strategy $x \in \Delta$, but after a while, suddenly a group of mutants playing

some other strategy $y \in \Delta$ arises in the population. Under what conditions can the incumbent x -players resist the invasion of the y -playing mutants?

To answer this question, denote the ratio of mutants in the whole population by $\epsilon \in (0, 1)$. Then the strategy of the average population is $(1 - \epsilon)x + \epsilon y$. Hence, the fitness of the x -players is $u(x, (1 - \epsilon)x + \epsilon y)$ and that of the mutants is $u(y, (1 - \epsilon)x + \epsilon y)$. Assuming that the evolutionary forces select the fittest players, we acquire the following condition for the incumbent x -players to take over the population:

$$u(x, (1 - \epsilon)x + \epsilon y) > u(y, (1 - \epsilon)x + \epsilon y). \quad (\text{E.1})$$

Now if for any mutant strategy $y \in \Delta$, the x -players can invade the mutant population, provided that their share is sufficiently small, then x is called an evolutionarily stable strategy. More technically, we have the following definition.

Definition 7. A strategy $x \in \Delta$ is an evolutionarily stable strategy (ESS) if for every other strategy $y \in \Delta - \{x\}$, there exists some $\bar{\epsilon}_y \in (0, 1)$ such that inequality (E.1) holds for all $\epsilon \in (0, \bar{\epsilon}_y)$.

Note that by definition, the mutants are not confined to be associated with a pure strategy; mutation to mixed-strategy players is also allowed. Although Definition (7) does not seem to be related to the notion of Nash equilibrium, the following result, while providing an easier test for evolutionary stability, strongly links the two by showing that every ESS is a Nash strategy [7]. See Appendix B for definitions a Nash strategy and its refinements.

Proposition 10. A strategy $x \in \Delta$ is evolutionarily stable if and only if the following conditions hold

$$u(x, x) \geq u(y, x) \quad \forall y \in \Delta, \quad (\text{E.2})$$

$$u(x, x) = u(y, x) \Rightarrow u(x, y) > u(y, y) \quad \forall y \in \Delta - \{x\}. \quad (\text{E.3})$$

This is indeed how evolutionary stability was originally defined [E1, E2], and still sometimes is (Definition 7 was later introduced in [E3], and the version that we used is stated in [7]). According to (E.2), if x is an ESS, no other strategy may perform better against x . In case some strategy y performs equally well, then (E.3) asserts that x must outperform y against y . The first condition is that of a Nash equilibrium, making ESS a refinement of a Nash strategy. Interestingly, when John Maynard Smith came up with this concept, he was not aware of the concept of Nash equilibrium in game theory [85].

It is worth mentioning that the notion of ESS is defined independently from the replicator or any other population dynamics, and implies invasion against mutants. However, as being discussed in the main text, ESS is sufficient for asymptotic stability under the RD and a range of other dynamics, motivating the vast literature dedicated to ESS.

The set of evolutionarily stable strategies is denoted by Δ^{ESS} . As mentioned above, the condition (E.2) implies $\Delta^{ESS} \subseteq \Delta^{NE}$. Moreover, Δ^{ESS} is known to be finite, and in particular, if an ESS belongs to the interior of the simplex, then it is the unique ESS of the game, i.e., Δ^{ESS} is a singleton [7].

In general every vector $x \in \Delta$ can represent both a strategy vector whose entries indicate the probability of playing a specific pure strategy, and a population state whose entries indicate the share of individuals playing a specific pure strategy. For the former case, the evolutionarily stable strategy x may also be referred to as an *evolutionarily stable state*.

In addition to those in Proposition 10, every ESS has two other unique properties. The first is having a *uniform invasion barrier*, i.e., there exists some $\bar{\epsilon} \in (0, 1)$ such that inequality (E.1) holds for all $y \in \Delta - \{x\}$ and all $\epsilon \in (0, \bar{\epsilon})$ [E4, E5]:

Proposition 11. A strategy $x \in \Delta$ is an ESS if and only if x has a uniform invasion barrier.

In other words, for the incumbent strategy x to be an ESS, there must be some population share, for which every invasion of mutants that occupies a smaller share than this is resisted by x . Namely, $\bar{\epsilon}_y$ can be the same for all mutant strategies y . The second unique property is being *locally superior*, that is to earn more against all nearby mutants than what they earn against themselves [E5, 7]. Mathematically, a strategy $x \in \Delta$ is locally superior if it has a neighborhood $\mathcal{U} \subseteq \Delta$ such that $u(x, y) > u(y, y)$ for all $y \in \mathcal{U} - \{x\}$.

Proposition 12. A strategy $x \in \Delta$ is an ESS if and only if x is locally superior.

This local superiority becomes global when the ESS belongs to the interior of the simplex and is, hence, unique. The notion of evolutionary stability requires a strong assumption on the incumbent strategy, that is to

earn more than any group of mutants, provided that their population share is sufficiently small. If this is relaxed to earning no less than the mutants, we obtain a weaker notion of evolutionary stability, *neutrally stability* [E6], defined as follows.

Definition 8. A strategy $x \in \Delta$ is a *neutrally stable strategy (NSS)* if for every other strategy $y \in \Delta - \{x\}$, there exists some $\bar{\epsilon}_y \in (0, 1)$ such that the inequality

$$u(x, (1 - \epsilon)x + \epsilon y) \geq u(y, (1 - \epsilon)x + \epsilon y) \quad (\text{E.4})$$

holds for all $\epsilon \in (0, \bar{\epsilon}_y)$.

Similar to the case with ESSs, the following proposition nicely bridges neutrally stable and Nash strategies.

Proposition 13. A strategy $x \in \Delta$ is neutrally stable if and only if the following conditions hold:

$$u(x, x) \geq u(y, x) \quad \forall y \in \Delta, \quad (\text{E.5})$$

$$u(x, x) = u(y, x) \Rightarrow u(x, y) \geq u(y, y) \quad \forall y \in \Delta - \{x\}. \quad (\text{E.6})$$

The set of NSS is denoted by Δ^{NSS} . Based on what has been discussed so far,

$$\Delta^{ESS} \subseteq \Delta^{NSS} \subseteq \Delta^{NE}.$$

Now we investigate these notions in 2×2 symmetric games.

Prisoner's dilemma game: We know $\Delta^{NE} = \{e^2\}$, so $\Delta^{ESS} \subseteq \{e^2\}$. On the other hand, e^2 is a strict Nash equilibrium. Hence, $\Delta^{ESS} = \{e^2\}$. Therefore, in view of $\Delta^{NE} = \{e^2\}$, it also holds that $\Delta^{NSS} = \{e^2\}$. So defection is the only neutrally stable strategy, which is also evolutionarily stable.

Coordination game: We know $\Delta^{NE} = \{e^1, x^*, e^2\}$. Satisfying (E.2) with strict inequality, $e^1, e^2 \in \Delta^{ESS}$. Hence, e^1 and e^2 also belong to Δ^{NSS} . However, x^* incurs equality in (E.5), and does not satisfy (E.6) for $y = e^1$ as an example. So $x^* \notin \Delta^{NSS}$. Hence, $\Delta^{NSS} = \Delta^{ESS} = \{e^1, e^2\}$. So both cooperation and defection are evolutionarily stable.

Anti-Coordination game: We know $\Delta^{NE} = \{x^*\}$. It can be shown that x^* satisfies (E.3). Hence, $\Delta^{NSS} = \Delta^{ESS} = \{x^*\}$. So only a mixture of cooperation and defection is evolutionarily stable.

References

- [E1] J. M. Smith and G. R. Price. "The logic of animal conflict." *Nature*, vol. 246, pp. 15-18, 1973.
- [E2] J. M. Smith. "The theory of games and the evolution of animal conflicts." *Journal of Theoretical Biology*, vol. 47, no. 1, pp. 209-221, 1974.
- [E3] P. D. Taylor and L. B. Jonker. "Evolutionary stable strategies and game dynamics." *Mathematical Biosciences*, vol. 40, no. (1-2), pp. 145-156, 1978.
- [E4] J. Hofbauer, P. Schuster, and K. Sigmund. "A note on evolutionary stable strategies and game dynamics." *Journal of Theoretical Biology*, vol. 81, no. 3, pp. 609-612, 1979.
- [E5] G. T. Vickers, and C. Cannings. "Patterns of ESS's I." *Journal of Theoretical Biology*, vol 132, no. 4, pp. 387-408, 1988.
- [E6] J. M. Smith. *Evolution and The Theory of Games*. Cambridge University Press, 1982.

Appendix F. Lotka-Volterra equation

The Lotka-Volterra Equation, describing the interaction of m species, is given by

$$\dot{y}_i = y_i \left[r_i + \sum_{j=1}^n b_{ij} y_j \right], \quad i = 1, \dots, m$$

where $y_i \geq 0$ denotes the population size of species i , $r_i \in \mathbb{R}$ denotes the growth rate of species i and $b_{ij} \in \mathbb{R}$ describes the extent to which species j competes for resources used by species i [F1, F2]. Apparently, the $(m - 1)$ -dimensional Lotka-Volterra equations are equivalent to the m -dimensional replicator dynamics (8) when $r_i = \pi_{im} - \pi_{mm}$ and $b_{ij} = \pi_{ij} - \pi_{mj}$ for all $i, j \in \{1, \dots, m\}$. This can be shown by applying the transformation $x_i \mapsto \frac{y_i}{1 + \sum_{j=1}^{m-1} y_j}$ for $i = 1, \dots, m$ [F1].

Because of their application in ecology and biology [F3, F4], a huge body of literature is dedicated to the stability analysis of the Lotka-Volterra equations. In particular, the equations are known to exhibit multiple limit cycles [F5, F6], strange attractors [F7] and chaotic behaviors, even for $m = 3$ [F8]. In view of the equivalence explained above, this implies that the same phenomena can happen in 4-dimensional replicator dynamics.

References

- [F1] M. A. Nowak. *Evolutionary dynamics*. Harvard University Press, 2006.
- [F2] J. A. Vano, J. C. Wildenberg, M. B. Anderson, J. K. Noel, and J. C. Sprott. "Chaos in low-dimensional LotkaVolterra models of competition." *Nonlinearity*, vol. 19, no. 10, pp. 2391-2404, 2006.
- [F3] I. M. Bomze. "Lotka-Volterra equation and replicator dynamics: a two-dimensional classification." *Biological cybernetics*, vol. 48, no. 3, pp. 201-211, 1983.
- [F4] Y. Kurihara, K. Watanabe, and H. Tanaka. "A sleep model and an observer using the LotkaVolsterra equation for real-time estimation of sleep." *Artificial Life and Robotics*, vol. 21, no. 1, pp. 132-139, 2016.
- [F5] J. Hofbauer and J. W.-H. So. "Multiple limit cycles for three dimensional Lotka-Volterra equations." *Applied Mathematics Letters*, vol. 7, no. 6, pp. 65-70, 1994.
- [F6] D. Xiao and W. Li. "Limit cycles for the competitive three dimensional LotkaVolterra system." *Journal of Differential Equations*, vol. 164, no. 1, pp. 1-15, 2000.
- [F7] A. Arneodo, P. Coullet, and C. Tresser. "Occurence of strange attractors in three-dimensional Volterra equations." *Physics Letters A*, vol. 79, no. 4, pp. 259-263, 1980.
- [F8] L. Gardini, R. Lupini, C. Mammanna, and M. G. Messina. "Bifurcations and transitions to chaos in the three-dimensional Lotka-Volterra map." *SIAM Journal on Applied Mathematics*, vol. 47, no. 3, pp. 455-482, 1987.

Appendix G. Convergence proofs for asynchronous best-response dynamics

It is convenient for the analysis here to consider a strategy set $\mathcal{S} := \{+1, -1\}$ with corresponding payoff matrix:

$$\pi_{ij} := \begin{matrix} & + & - \\ \begin{matrix} + \\ - \end{matrix} & \begin{pmatrix} a_{ij} & b_{ij} \\ c_{ij} & d_{ij} \end{pmatrix} \end{matrix}, \quad a_{ij}, b_{ij}, c_{ij}, d_{ij} \in \mathbb{R}.$$

The total payoffs can then be expressed in a compact form:

$$u_i(x_i, x_{-i}) = \begin{cases} \frac{1}{2} \sum_{j \in \mathcal{N}_i} a_{ij}(1 + x_j) + b_{ij}(1 - x_j), & \text{if } x_i = +1 \\ \frac{1}{2} \sum_{j \in \mathcal{N}_i} c_{ij}(1 + x_j) + d_{ij}(1 - x_j), & \text{if } x_i = -1 \end{cases}.$$

The best-response update (2) can be expressed as follows:

$$x_i(k+1) = \text{sign} \left(\sum_{j \in \mathcal{N}_i} (\delta_{ij}^+ - \delta_{ij}^-) + x_j(k)(\delta_{ij}^+ + \delta_{ij}^-) \right), \quad (\text{G.1})$$

where $\delta_{ij}^+ = a_{ij} - c_{ij}$ and $\delta_{ij}^- = d_{ij} - b_{ij}$. Here, we have implicitly assumed that the case when both strategies result in equal payoffs results in an update to the +1 strategy, which is equivalent to setting $z_i = A$ for all agents in the update rule (6). By symmetry of the problem setup, every case in which $z_i \in \{A, B\}$ can be transformed into the case where $z_i = A$ for all $i \in \mathcal{V}$ by rearranging payoff matrices of some of the agents. It is therefore sufficient to prove convergence of the update rule (G.1).

Proof of Theorem 4

PROOF. Let $h_i = \frac{\sum_{j \in \mathcal{N}_i} \delta_{ij}^+ - \delta_{ij}^-}{\delta_{ij}^+ + \delta_{ij}^-}$, and define the following potential function:

$$\Phi(x(k)) = \sum_{i=1}^n \phi_i(x(k)), \quad (\text{G.2})$$

where

$$\phi_i(x(k)) = -x_i(k) \left(\frac{1}{2} \sum_{j \in \mathcal{N}_i} x_j(k) + h_i \right).$$

From (G.1), it holds that an agent who activates at time k only changes strategies if

$$\begin{aligned} & x_i(k) \operatorname{sign} \left(\sum_{j \in \mathcal{N}_i} (\delta_i^+ - \delta_i^-) + x_j(k)(\delta_i^+ + \delta_i^-) \right) < 0 \\ \iff & x_i(k) \left(\sum_{j \in \mathcal{N}_i} (\delta_i^+ - \delta_i^-) + x_j(k)(\delta_i^+ + \delta_i^-) \right) < 0 \end{aligned}$$

Since $\delta_i^+ + \delta_i^- > 0$, we can divide by $\delta_i^+ + \delta_i^-$ without changing the sign. The criteria for an agent to switch strategies then becomes

$$x_i(k) \left(\sum_{j \in \mathcal{N}_i} x_j(k) + h_i \right) < 0 \quad (\text{G.3})$$

At this time, the change in potential is given by $\Delta\Phi(x(k))$

$$\begin{aligned} &= \Phi(x(k+1)) - \Phi(x(k)) \\ &= \phi_i(x(k+1)) - \phi_i(x(k)) \\ &+ \sum_{j \in \mathcal{N}_i} \phi_j(x(k+1)) - \phi_j(x(k)) \\ &= -\frac{1}{2} \sum_{j \in \mathcal{N}_i} x_i(k+1)x_j(k+1) - h_i x_i(k+1) \\ &+ \frac{1}{2} \sum_{j \in \mathcal{N}_i} x_i(k)x_j(k) + h_i x_i(k) \\ &- \frac{1}{2} \sum_{j \in \mathcal{N}_i} x_i(k+1)x_j(k+1) - h_j x_j(k+1) \\ &+ \frac{1}{2} \sum_{j \in \mathcal{N}_i} x_i(k)x_j(k) + h_j x_j(k) \end{aligned}$$

Since only the active agent changes strategies, we have $x_i(k+1) = -x_i(k)$ and $x_j(k+1) = x_j(k)$ for all $j \neq i$. Therefore, we can simplify the above expression as follows:

$$\begin{aligned} \Delta\Phi(x(k)) &= 2 \sum_{j \in \mathcal{N}_i} x_i(k)x_j(k) + 2h_i x_i(k) \\ &= 2x_i(k) \left(\sum_{j \in \mathcal{N}_i} x_j(k) + h_i \right). \end{aligned}$$

Due to (G.3), we have $\Delta\Phi(x(k)) < 0$ whenever an agent switches strategies. For a given network, the potential function $\Delta\Phi(x(k))$ is bounded from above and below. Since there are only a finite number of possible values of $\Delta\Phi(x(k))$ for a given network, $|\Delta\Phi(x(k))|$ is lower bounded. It follows that the network will reach an equilibrium state after a finite number of strategy switches. Finally, by the assumption of persistent activations, each agent will activate an infinite number of times as time goes to infinity and thus an equilibrium state will be reached at some finite time.

Proof of Theorem 5

PROOF. We again use the potential function (G.2). From (G.1), it holds that an agent who activates at time k only changes strategies if

$$\begin{aligned} & x_i(k) \operatorname{sign} \left(\sum_{j \in \mathcal{N}_i} (\delta_i^+ - \delta_i^-) + x_j(k)(\delta_i^+ + \delta_i^-) \right) < 0 \\ \iff & x_i(k) \left(\sum_{j \in \mathcal{N}_i} (\delta_i^+ - \delta_i^-) + x_j(k)(\delta_i^+ + \delta_i^-) \right) < 0 \end{aligned}$$

Since $\delta_i^+ + \delta_i^- < 0$, dividing by $\delta_i^+ + \delta_i^-$ will always negate the sign. The criteria for an agent to switch strategies then becomes

$$x_i(k) \left(\sum_{j \in \mathcal{N}_i} x_j(k) + h_i \right) > 0 \quad (\text{G.4})$$

As shown in the proof of Theorem 4, the change in potential at such a time is given by

$$\begin{aligned} \Delta\Phi(x(k)) &= 2 \sum_{j \in \mathcal{N}_i} x_i(k)x_j(k) + 2h_i x_i(k) \\ &= 2x_i(k) \left(\sum_{j \in \mathcal{N}_i} x_j(k) + h_i \right). \end{aligned}$$

Due to (G.4), we have $\Delta\Phi(x(k)) > 0$ whenever an agent switches strategies. The rest of the proof follows exactly as the proof of Theorem 4.

Proof of Theorem 6

PROOF. Let $k_{ij} = \delta_{ij}^+ + \delta_{ij}^-$ and $\tau_i = \sum_{j \in \mathcal{N}_i} \delta_{ij}^+ - \delta_{ij}^-$, and define the following potential function:

$$\Phi(x(k)) = \sum_{i=1}^n \phi_i(x(k)),$$

where

$$\phi_i(x(k)) = -x_i(k) \left(\frac{1}{2} \sum_{j \in \mathcal{N}_i} k_{ij} x_j(k) + \tau_i \right).$$

From (G.1), it holds that an agent who activates at time k only changes strategies if

$$\begin{aligned} &x_i(k) \operatorname{sign} \left(\sum_{j \in \mathcal{N}_i} k_{ij} x_j(k) + \tau_i \right) < 0 \\ \iff &x_i(k) \left(\sum_{j \in \mathcal{N}_i} k_{ij} x_j(k) + \tau_i \right) < 0 \end{aligned} \quad (\text{G.5})$$

At this time, the change in potential is given by $\Delta\Phi(x(k))$

$$\begin{aligned} &= \Phi(x(k+1)) - \Phi(x(k)) \\ &= \phi_i(x(k+1)) - \phi_i(x(k)) \\ &\quad + \sum_{j \in \mathcal{N}_i} \phi_j(x(k+1)) - \phi_j(x(k)) \\ &= -\frac{1}{2} \sum_{j \in \mathcal{N}_i} k_{ij} x_i(k+1) x_j(k+1) - \tau_i x_i(k+1) \\ &\quad + \frac{1}{2} \sum_{j \in \mathcal{N}_i} k_{ij} x_i(k) x_j(k) + \tau_i x_i(k) \\ &\quad - \frac{1}{2} \sum_{j \in \mathcal{N}_i} k_{ji} x_i(k+1) x_j(k+1) - \tau_j x_j(k+1) \\ &\quad + \frac{1}{2} \sum_{j \in \mathcal{N}_i} k_{ji} x_i(k) x_j(k) + \tau_j x_j(k) \end{aligned}$$

Since only the active agent changes strategies, we have $x_i(k+1) = -x_i(k)$ and $x_j(k+1) = x_j(k)$ for all $j \neq i$. Also,

since all games are symmetric, $k_{ij} = k_{ji}$. Therefore, we can simplify the above expression as follows:

$$\begin{aligned}\Delta\Phi(x(k)) &= 2 \sum_{j \in \mathcal{N}_i} k_{ij} x_i(k) x_j(k) + 2\tau_i x_i(k) \\ &= 2x_i(k) \left(\sum_{j \in \mathcal{N}_i} k_{ij} x_j(k) + \tau_i \right).\end{aligned}$$

Due to (G.5), we have $\Delta\Phi(x(k)) < 0$ whenever an agent switches strategies. For a given network, the potential function $\Delta\Phi(x(k))$ is bounded from above and below. Since there are only a finite number of possible values of $\Delta\Phi(x(k))$ for a given network, $|\Delta\Phi(x(k))|$ is lower bounded. It follows that the network will reach an equilibrium state after a finite number of strategy switches. Finally, since the activation sequence is persistent, each agent will activate an infinite number of times as time goes to infinity and thus the equilibrium state will be reached at some finite time.