Tuning Cooperative Behavior in Games With Nonlinear Opinion Dynamics

Shinkyu Park, *Member, IEEE*, Anastasia Bizyaeva, *Graduate Student Member, IEEE*, Mari Kawakatsu, Alessio Franci, and Naomi Ehrich Leonard, *Fellow, IEEE*

***Abstract*—We examine the tuning of cooperative behavior in repeated multi-agent games using an analytically tractable, continuous-time, nonlinear model of opinion dynamics. Each modeled agent updates its real-valued opinion about each available strategy in response to payoffs and other agents' opinions, as observed over a network. We show how the model provides a principled and systematic means to investigate behavior of agents that select strategies using rationality and reciprocity, key features of human decision-making in social dilemmas. For two-strategy games, we use bifurcation analysis to prove conditions for the bistability of two equilibria and conditions for the first (second) equilibrium to reflect all agents favoring the first (second) strategy. We prove how model parameters, e.g., level of attention to opinions of others, network structure, and payoffs, influence dynamics and, notably, the size of the region of attraction to each stable equilibrium. We provide insights by examining the tuning of the bistability of mutual cooperation and mutual defection and their regions of attraction for the repeated prisoner's dilemma and the repeated multi-agent public goods game. Our results generalize to games with more strategies, heterogeneity, and additional feedback dynamics, such as those designed to elicit cooperation or coordination.**

***Index Terms*—Game theory, opinion dynamics, decision making, distributed control, multi-agent systems.**

Manuscript received September 14, 2021; revised November 22, 2021; accepted December 10, 2021. Date of publication December 27, 2021; date of current version December 30, 2021. This work was supported in part by the King Abdullah University of Science and Technology (KAUST); in part by US Office of Naval Research (ONR) under Grant N00014-19-1-2556; in part by the US Army Research Office (ARO) under Grant W911NF-18-1-0325; in part by Direcciòn General de Asuntos del Personal Académico (DGAPA), National Autonomous University of Mexico (UNAM), through the Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT) Research under Grant IN102420; in part by the Conacyt under Grant A1-S-10610; in part by the NSF Graduate Research Fellowship under Grant DGE-2039656; and in part by the Princeton University School of Engineering and Applied Science through the generosity of Lydia and William Addy '82. Recommended by Senior Editor M. Guay.
(Corresponding author: Shinkyu Park.)

Shinkyu Park is with the Electrical and Computer Engineering, King Abdullah University of Science and Technology, Thuwal 23955, Saudi Arabia (e-mail: shinkyu.park@kaust.edu.sa).

Anastasia Bizyaeva and Naomi Ehrich Leonard are with the Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: bizyaeva@princeton.edu; naomi@princeton.edu).

Mari Kawakatsu is with the Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ 08544 USA (e-mail: mari.kawakatsu@princeton.edu).

Alessio Franci is with the Mathematics Department, National Autonomous University of Mexico, Mexico City 04510, Mexico (e-mail: afranci@ciencias.unam.mx).

Digital Object Identifier 10.1109/LCSYS.2021.3138725

I. Introduction

SOCIOLOGISTS, political scientists, and economists have long argued that *reciprocity* is key to promoting cooperation [1]–[3]. Computer simulations have shown that reciprocal strategies can elicit mutual cooperation in repeated games: the winning strategy for the repeated prisoner's dilemma in Axelrod's tournaments was Tit-for-Tat (TFT), where an agent reciprocates the opponent's strategy in the previous round; more generally, successful strategies were *nice*, *forgiving*, *provocable*, and *clear* [2]. Subsequent laboratory studies have revealed that humans in fact employ such reciprocity-based rules in repeated interactions [4]–[6]. However, the observed reciprocity cannot be recapitulated by game-theoretic models of rational, payoff-maximizing agents, which, in contrast to the experiments, predict convergence toward mutual defection, i.e., the Nash equilibrium in a social dilemma.

Here we investigate the tuning of cooperative behavior, including mutual cooperation or coordination, in repeated games among agents that rely on both rationality and reciprocity. Our first key contribution is a new framework for studying multi-agent repeated games using the *nonlinear opinion dynamics model* [7] (see also [8]) in which agents' strategic decisions depend not only on payoffs, as in rationality models [9], [10], but also on *social interactions* that enable agents to observe strategy preferences (opinions) of other agents. We show how the social interaction term, formulated as a saturation function of observed opinions, provides a representation of reciprocity and a means to tune cooperation (or coordination) in social dilemmas.

Our second key contribution leverages analytical tractability of the model: we prove conditions for bistability of two equilibria for repeated two-strategy games in which multiple agents observe the opinions of others over a fixed network. We also show conditions under which each equilibrium corresponds to all agents favoring one of the two strategies. Our proof relies on a bifurcation analysis that builds on the results of [7]. We prove how the bistability of equilibria and the regions of attraction depend on the level of attention to observed opinions, network structure, payoffs, and other model parameters. We apply our theory to the two-agent prisoner's dilemma and the multi-agent public goods game to present further insights into how mutual cooperation emerges through social interaction (reciprocity) and how the predicted likelihood of cooperation can be tuned. Our results apply analogously to tuning coordination in games like the Stag Hunt. Our analytical results complement the large literature on reciprocity-based decision-making [2] that evaluates agents' long-term interaction with computer simulations.

Most models of opinion dynamics in the literature use an opinion updating process that relies on a linear weighted average of exchanged opinions, as in the original work of DeGroot [11]. The nonlinear opinion dynamics model of [7] instead applies a saturation function to exchanged opinions, making the updating process fundamentally nonlinear and thus allowing for multistability of equilibria, a key aspect of our project. For a comprehensive review of, and comparison with, other opinion dynamics models see [7]. Our investigation of the means to tune cooperation in social dilemmas is distinguished from works such as [12], [13] that examine opinion dynamics using game-theoretic approaches.

Our approach is also distinguished from the investigations in [7] as evolving opinions, which represent strategy preferences, depend not only on saturated opinion exchange but also on the payoff mechanism of the game. Our results are also new–they explain the emergence of mutual cooperation (or coordination) in social dilemmas as one of two bistable equilibria that arise through a pitchfork bifurcation.

In Section II, we introduce the nonlinear opinion dynamics model and show how it recovers rationality and reciprocity. In Section III, for two-strategy games, we prove the bistability of equilibria and expressions for the tunability of those equilibria and their corresponding regions of attraction in terms of system parameters. We apply the theory to the prisoner's dilemma and public goods game. In Section IV, we use numerical simulations to illustrate the theoretical predictions on the tuning of cooperation. In Section V, we discuss extensions and generalizations.

II. Opinion Dynamics in Games

Consider an N_a-agent decision-making problem where each agent selects a strategy, continuously in time t, from the set $\{1, \dots, N_s\}$ of N_s available strategies. Each agent performs a probabilistic choice of strategy where $x_i(t) \in \mathbb{X}_i$ is the probability distribution for the strategy selection at time t of agent i and $\mathbb{X}_i$ is the probability simplex in $\mathbb{R}^{N_s}$. The j-th element x_{ij} of x_i is the probability that agent i selects strategy j. Following convention in game theory [14], x_i is the *mixed strategy* of agent i and $x = (x_1, \dots, x_{N_a}) \in \mathbb{X}$ is the *mixed strategy profile*, where $\mathbb{X} = \mathbb{X}_1 \times \cdots \times \mathbb{X}_{N_a}$.

The mixed strategy $x_i(t)$ is defined by the logit choice function [10] and depends on agent i's *opinion state* at time t, $\bar{z}_i(t) = (\bar{z}_{i1}, \dots, \bar{z}_{iN_s})(t) \in \mathbb{R}^{N_s}$, as follows:

$$x_{ij} = \sigma_j(\bar{z}_i) = \frac{\exp\left(\eta^{-1}\bar{z}_{ij}\right)}{\sum_{l=1}^{N_s} \exp\left(\eta^{-1}\bar{z}_{il}\right)}, \tag{1}$$

where the positive constant η is called the noise level [15] or rationality parameter [16].[1] Each entry $\bar{z}_{ij}$ of $\bar{z}_i$ represents agent i's preference for the j-th available strategy. The *relative opinion state* z_i, with j-th entry $z_{ij} = \bar{z}_{ij} - \frac{1}{N_s}\sum_{l=1}^{N_s} \bar{z}_{il}$, defines an agent's preferred strategies, i.e., the inequality $z_{ij} > 0$ can be interpreted as the agent favoring strategy j relative to other strategies and the magnitude $|z_{ij}|$ denotes the level of its preference. Under logit choice (1), the higher $\bar{z}_{ij}$ relative to other entries of $\bar{z}_i$, the more likely agent i selects strategy j. Equation (1) can be interpreted as the best response with respect to the opinion state $\bar{z}_i$ subject to a random perturbation [15].

Given mixed strategy profile $x \in \mathbb{X}$, we let $U_i(x) = (U_{i1}(x), \dots, U_{iN_s}(x)) \in \mathbb{R}^{N_s}$ be the *payoff function* for agent i. Entry $U_{ij}(x)$ defines agent i's payoff associated with strategy j. The following are examples of multi-agent games.

Example 1 (Prisoner's Dilemma): Consider two agents, each with two available strategies: cooperate (strategy 1) and defect (strategy 2). When both agents cooperate or defect, they receive payoff p_{CC} or p_{DD}, respectively. If one defects while the other cooperates, the former receives payoff p_{DC} and the latter receives p_{CD}. The payoff function U_i is

$$U_i(x) = \begin{pmatrix} U_{i1}(x) \\ U_{i2}(x) \end{pmatrix} = \begin{pmatrix} p_{CC} & p_{CD} \\ p_{DC} & p_{DD} \end{pmatrix} x_{-i}, \quad i \in \{1, 2\} \tag{2}$$

where, as shorthand notation, we let $x_{-1} = x_2$ and $x_{-2} = x_1$. The parameters $p_{CC}, p_{CD}, p_{DC}, p_{DD}$ satisfy $p_{DC} > p_{CC} > p_{DD} > p_{CD}$, which means that the agents have individual incentives to defect and receive p_{DD}, even though they would receive the higher payoff p_{CC} by cooperating.

Example 2 (Public Goods Game): There are N_a agents and N_s strategies. Each agent has a total wealth of $a(N_s - 1)$ and selects a strategy j in $\{1, \dots, N_s\}$ that corresponds to contributing $a(N_s - j)$ to a public pool. The total contribution is multiplied by a factor ρ and distributed equally among all agents. The payoff function U_i is

$$U_{ij}(x) = a(j-1) + \frac{\rho}{N_a}\sum_{\substack{k\neq i\\ k=1}}^{N_a}\sum_{l=1}^{N_s} a(N_s - l)\, x_{kl} + \frac{\rho}{N_a} a(N_s - j),$$
$$i \in \{1, \dots, N_a\},\ j \in \{1, \dots, N_s\}, \tag{3}$$

where $a > 0$ and $N_a > \rho > 1$. According to (3), regardless of the others' contributions, each agent receives the highest payoff when it makes no contribution to the pool. Hence, the rational agent contributes nothing, i.e., chooses $j = N_s$.

We define rate-of-change $\dot{\bar{z}}_i = \mathrm{d}\bar{z}_i/\mathrm{d}t$ of agent i's opinion state $\bar{z}_i$ in response to payoffs and social interactions, with the continuous-time nonlinear opinion dynamics model [7][2]:

$$\dot{\bar{z}}_{ij} = -d_i\left(\bar{z}_{ij} - u_i \sum_{k=1}^{N_a} 2R\left(A_{ik}^j z_{kj}\right) - U_{ij}(x)\right), \tag{4}$$

with $\bar{z}_i(0) \in \mathbb{R}^{N_s}$. $A_{ik}^j \in \mathbb{R}$ is the weight agent i places in its evaluation of strategy j on its observation of agent k's opinion of strategy j. The constant *resistance parameter* $d_i > 0$ reflects the speed with which agent i's opinions change; the *attention parameter* $u_i > 0$ reflects the weight placed on incentives derived from social interactions, where $R : \mathbb{R} \to [0, 1]$. Thus, the state $\bar{z}_i$ of agent i, and hence its strategy selection, evolves according to the accumulation over time, with the discount factor d_i, of the payoffs $U_{ij}(x)$ and social incentives $R(A_{ik}^j z_{kj})$.

We define R as the saturating function

$$R(A_{ik}^j z_{kj}) = \mathbf{P}\left(A_{ik}^j z_{kj} \geq \epsilon\right), \tag{5}$$

where ϵ is a random variable with a symmetric and unimodal probability density function, e.g., the standard normal distribution. To interpret, suppose $A_{ik}^j \geq 0$. Then A_{ik}^j quantifies the

[1]For simplicity, we assume that η is identical across the agents.

[2]In Section III, we explain how (4) relates to its original form presented in [7]. For concise presentation, we omit time dependency of the variables in (4).

influence of noise ϵ on inter-agent interactions: the larger A_{ik}^j, the smaller the effect of noise ϵ.[3] Thus, we can interpret (5) as a probabilistic model of agent i's perception of agent k's preference for strategy j over other strategies.

A. Emergence of the Cooperative Equilibrium

In this section, using the prisoner's dilemma as an illustrative example, we provide intuition for how the equilibria of (4) depend on system parameters, and under what parameter regime a cooperative equilibrium emerges. To simplify the presentation, let $d_i = d$, $u_i = u$, $A_{ii}^j = \alpha$, and $A_{ik}^j = \gamma$ if $i \neq k$. Let $\bar{z}^*$ be an equilibrium of (4) that satisfies

$$\bar{z}_{ij}^* = 2u\left(R(\alpha z_{ij}^*) + \sum_{\substack{k \neq i \\ k=1}}^{N_a} R(\gamma z_{kj}^*)\right) + U_{ij}(x^*), \tag{6}$$

where $x_{ij}^* = \sigma_j(z_i^*)$.

Note that by (5), in a dense subset of the tangent space $T\mathbb{X}$ of $\mathbb{X}$, as the influence of the noise in the social interaction becomes arbitrarily small, i.e., α, γ are arbitrarily large, $R(\alpha z_{ij})$ and $R(\gamma z_{kj})$ converge to a binary ($\{0, 1\}$-valued) function. If α, γ are sufficiently large, we can approximate (6) as $\bar{z}_{ij}^* \approx 2u\, n_j^* + U_{ij}(x^*)$, where n_j^* is the number of agents k having a positive relative opinion z_{kj}^* of strategy j at equilibrium. As the attention u increases, each agent tends to favor the most popular strategy, even though selecting other strategies would return higher payoffs. It follows that the social interaction R incentivizes each agent to reciprocate with other agents in the strategy selection, and the level of reciprocation is determined by the attention parameter u and the number n_j^* of agents preferring the same strategy under consideration.

Example: With two reciprocating agents ($N_a = 2$, $\alpha = 0, \gamma > 0$) playing the prisoner's dilemma ($N_s = 2$), the equilibrium $\bar{z}^*$ satisfies $\bar{z}_{i1}^* \approx 2u\, n_1^* + U_{i1}(x^*)$, where $n_1^* \in \{0, 1\}$ indicates whether the opponent cooperates ($n_1^* = 1$) or defects ($n_1^* = 0$). If the attention parameter satisfies $2u > \max(p_{DC} - p_{CC}, p_{DD} - p_{CD})$, then for sufficiently large γ, cooperation becomes an equilibrium of (4). Moreover, given any arbitrarily large u, there is a minimum value of γ below which cooperation will not be an equilibrium.

B. Rationality and Reciprocity in the Model

In this section, we show how the model (4) captures a range of features observed in human decision-making, including (bounded) rationality [17] and reciprocity [1], [3]. We begin by showing that (4) generalizes the exponentially discounted reinforcement learning (EXP-D-RL) model studied in [9] where every agent makes an individually rational decision by selecting payoff-maximizing strategies. To see this, let $A_{ik}^j = 0$ for $i, k \in \{1, \ldots, N_a\}$ and $j \in \{1, \ldots, N_s\}$ for which the social interaction $R(A_{ik}^j z_{kj})$ becomes constant, i.e., $R(A_{ik}^j z_{kj}) = 0.5, \ \forall z_{kj} \in \mathbb{R}$. By translating $\bar{z}_{ij}$ by constant $u_i N_a$ and since the logit choice function is invariant with respect to the translation of $\bar{z}_{ij}$, (4) specializes to

$$\dot{\bar{z}}_{ij} = -d_i\big(\bar{z}_{ij} - U_{ij}(x)\big), \quad x_{ij} = \frac{\exp\big(\eta^{-1}\bar{z}_{ij}\big)}{\sum_{l=1}^{N_s}\exp\big(\eta^{-1}\bar{z}_{il}\big)},$$

which is the EXP-D-RL model presented in [9]. In this sense, our model (4) realizes rationality.

To discuss reciprocity of the opinion dynamics, we consider a two-agent ($N_a = 2$) two-strategy ($N_s = 2$) case. Suppose that $A_{ik}^j = \eta^{-1}$ if $i \neq k$ and $A_{ik}^j = 0$ otherwise, where η is the noise level constant in the logit choice function (1). Then, with $R(\cdot) = (\tanh(\cdot) + 1)/2$, we have $R(A_{ik}^j z_{kj}) = x_{kj}$ if $i \neq k$ and $R(A_{ik}^j z_{kj}) = 0.5$ otherwise.

For small $h > 0$, assuming that U_{ij} is arbitrarily small, we can approximate the opinion dynamics model (4) as

$$\bar{z}_{ij}(t+h) - \bar{z}_{ij}(t) \approx -hd_i\big(\bar{z}_{ij}(t) - 2u_i x_{-ij}(t)\big).$$

For sufficiently large d_i, by evaluating the opinion state at time instant $t+h$ with $h = d_i^{-1}$, we observe that

$$\bar{z}_{ij}(t+h) \approx 2u_i x_{-ij}(t). \tag{7}$$

Recall that x_{-ij} is the j-th entry of the mixed strategy x_{-i} of the opponent of agent i. According to (7), with large u_i, it holds that $x_{ij}(t+h) = 1$ if and only if $x_{-ij}(t) = 1$. In the prisoner's dilemma, under (7), each agent i decides to cooperate (or defect) if its opponent does so at the previous stage. This behavior resembles TFT, a well-known reciprocity-based strategy in discrete-time iterated games [2]. In this sense, our model (4) realizes reciprocity.

C. Further Remarks on the Model

Social interaction encourages reciprocity: When $A_{ik}^j > 0$ for $i \neq k$, the social interaction in (4) encourages reciprocity by incentivizing each agent to select the strategies preferred by other agents. As shown in Section IV, in the prisoner's dilemma and public goods game, such a social interaction mechanism leads to decision-making representative of human behavior; notably, the agents conditionally cooperate. This contrasts with the outcomes of rationality-based models where agents fail to cooperate (or coordinate).

Our model and analysis can be readily extended to a more general case, as in [7], where the social interaction term in (4) is given by $u_i \sum_{k=1}^{N_a}\sum_{l=1}^{N_s} 2R(A_{ik}^{jl} z_{kl})$. In this generalization, agent i's opinion of strategy j may also depend on other agents' opinions of strategies $l \neq j$.

Network structure: The A_{ik}^j in (4) defines a network structure among agents for strategy j. One can specify the presence ($A_{ik}^j > 0$ for reciprocal, $A_{ik}^j < 0$ for antagonistic) or lack ($A_{ik}^j = 0$) of interaction between agents i and k in their selecting strategy j. We prove results on the role of network structure in our model in Section III. See [7], [18] for more on network structure and the nonlinear opinion dynamics.

III. Bistability Analysis of 2-Strategy Games

We present bistability analysis for (4) in two-strategy games with homogeneous parameters.[4] We assume $G = (\mathbb{V}, \mathbb{E})$ and $\hat{G} = (\mathbb{V}, \hat{\mathbb{E}})$, with $\mathbb{V} = \{1, \ldots, N_a\}$, are simple graphs governing the social interaction and game interaction, respectively, and $A = (a_{ik})_{i,k\in\mathbb{V}}$ and $\hat{A} = (\hat{a}_{ik})_{i,k\in\mathbb{V}}$ are the corresponding adjacency matrices. We assume the payoff function has the form:

$$\begin{pmatrix} U_{i1}(x) \\ U_{i2}(x) \end{pmatrix} = \sum_{(i,k)\in\hat{\mathbb{E}}} \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} x_k + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \tag{8}$$

and the parameters of (4) are given by $d_i = d$, $u_i = u$, $A_{ii}^j = \alpha > 0$, and $A_{ik}^j = \gamma a_{ik} \geq 0$ if $i \neq k$.

[3] See Section II-C for more discussions on the parameter A_{ik}^j.

[4] The proofs of all the theorems are provided in the Appendix.

For analysis, we adopt the original form of (4) from [7]:

$$\dot{z}_{ij} = F_{ij}(z) - \frac{1}{N_s}\sum_{l=1}^{N_s} F_{il}(z), \quad \sum_{j=1}^{N_s} z_{ij}(0) = 0,$$

$$F_{ij}(z) = -d\left(z_{ij} - u\left(S(\alpha z_{ij}) + \sum_{(i,k)\in\mathbb{E}} S(\gamma z_{kj})\right) - U_{ij}(x)\right) \quad (9)$$

where $z_{ij}(0) = \bar{z}_{ij}(0) - \frac{1}{N_s}\sum_{l=1}^{N_s}\bar{z}_{il}(0)$ and the saturation function S is given by $S = 2R - 1$. The variable $z = (z_1, \ldots, z_{N_a}) \in T\mathbb{X}$ denotes the relative opinion state. In Theorem 1, we show that models (4) and (9) are related by the projection $z_i = P_0\bar{z}_i$, where $P_0 = I - \frac{1}{N_s}\mathbf{1}\mathbf{1}^T$, and yield the same transient and steady-state mixed-strategy behavior.

Theorem 1: The following two statements are true.

i) If $\bar{z}(t)$ is a solution of (4), then $z(t)$, satisfying $z_i(t) = P_0\bar{z}_i(t)$, is a solution of (9). Conversely, if $z(t)$ is a solution of (9), then $\bar{z}(t)$ defined as $\bar{z}_{ij}(t) = e^{-dt}\bar{z}_{ij}(0) + d\int_0^t e^{-d(t-\tau)}(2u(R(\alpha z_{ij}(\tau)) + \sum_{(i,k)\in\mathbb{E}} R(\gamma z_{kj}(\tau))) + U_{ij}(x(\tau)))\,\mathrm{d}\tau$ with $x_{ij} = \sigma_j(z_i)$ satisfies $z_i(t) = P_0\bar{z}_i(t)$ and is a solution of (4).

ii) If $\bar{z}^*$ is a stable (unstable) equilibrium of (4), then z^*, satisfying $z_i^* = P_0\bar{z}_i^*$, is a stable (unstable) equilibrium of (9). Conversely, if z^* is a stable (unstable) equilibrium of (9) then $\bar{z}^*$, defined as $\bar{z}_{ij}^* = 2u(R(\alpha z_{ij}^*) + \sum_{(i,k)\in\mathbb{E}} R(\gamma z_{kj}^*)) + U_{ij}(x^*)$ with $x_{ij}^* = \sigma_j(z_i^*)$ satisfies $z_i^* = P_0\bar{z}_i^*$ and is a stable (unstable) equilibrium of (4).

We further assume that S satisfies the following conditions: S is odd sigmoidal and it holds that $S(0) = 0$, $S'(0) > 0$, $\operatorname{sign} S''(a) = -\operatorname{sign}(a)$, $\forall a \in \mathbb{R}$, and $S'''(0) = -2$.[5] Since $z_{i1} = -z_{i2}$, we can simplify the expression (9) as

$$\dot{\mathbf{z}} = -d\Big(\mathbf{z} - u(S(\alpha\mathbf{z}) + AS(\gamma\mathbf{z})) - \frac{1}{4}p\hat{A}\tanh(\eta^{-1}\mathbf{z}) - \frac{1}{4}p^{\perp}\hat{A}\mathbf{1} - (b_1 - b_2)\mathbf{1}\Big) \quad (10)$$

with $\mathbf{z} = (z_{11}, \ldots, z_{N_a 1})$, $p = p_{11} - p_{12} - p_{21} + p_{22}$, $p^{\perp} = p_{11} + p_{12} - p_{21} - p_{22}$, and $S(\gamma\mathbf{z}) = (S(\gamma z_{11}), \ldots, S(\gamma z_{N_a 1}))$.

Theorem 2 (Bistability in games): Consider (10). Let $\zeta_{max}(u, \gamma, p)$ be the largest-real-part eigenvalue of $u\gamma S'(0)A + \frac{1}{4}\eta^{-1}p\hat{A}$ and v_{max} (w_{max}) be its corresponding right (left) eigenvector.

i) Suppose ζ_{max} is real and simple, and $w_{max}^T\gamma S'(0)Av_{max} > 0$ holds. When $p^{\perp} = b_1 = b_2 = 0$, there exists a critical value u^* for which if $u < u^*$, the origin $\mathbf{z} = 0$ is locally exponentially stable, and if $u > u^*$, the origin is unstable and two bistable equilibrium solution branches emerge in a symmetric pitchfork bifurcation along a manifold tangent to the span of v_{max}. When $p^{\perp}$, b_1, and/or b_2 are nonzero, the system is an unfolding of the symmetric pitchfork bifurcation, and the parameter

$$b = \left\langle w_{max}, \frac{1}{4}dp^{\perp}\hat{A}\mathbf{1} + d(b_1 - b_2)\mathbf{1}\right\rangle \quad (11)$$

determines the direction of the unfolding. Furthermore, u^* depends on p according to $\frac{\partial u^*}{\partial p} = -\frac{1}{4\alpha S'(0)}w_{max}^T\hat{A}v_{max}$.

[5] To simplify the notation, without loss of generality, we make the assumption that $S'''(0) = -2$, for instance, by scaling S.

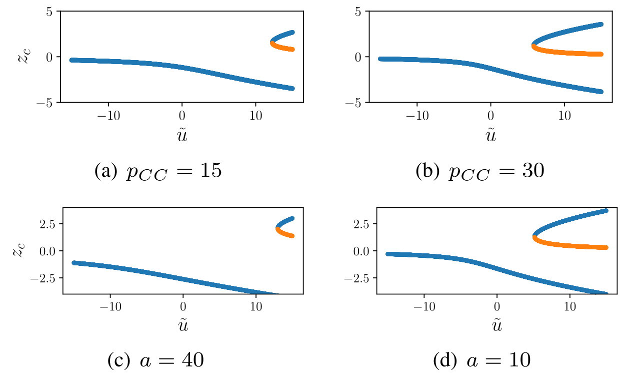


Fig. 1. Bifurcation diagram of (16) in the prisoner's dilemma with (a) $p_{CC} = 15$ and (b) $p_{CC} = 30$, where $\alpha = 0$, $d = \gamma = \eta = 1$, $p_{CD} = 0$, $p_{DC} = 40$, $p_{DD} = 5$, and in the public goods game with (c) $a = 40$ and (d) $a = 10$, where $\alpha = 0$, $d = \gamma = \eta = 1$, $\rho = 2$, and $N_a = 20$. Blue (orange) curves are stable (unstable) equilibrium states. Solutions with $z_c < 0$ ($z_c > 0$) correspond to mutual defection (cooperation). Mutual defection is always stable; for large enough $\tilde{u} = u - u^*$, mutual cooperation is stable.

ii) Suppose $u\gamma S'(0)A + \frac{1}{4}\eta^{-1}p\hat{A}$ is an irreducible nonnegative matrix.[6] Near u^*, for the bistable equilibria, $\operatorname{sign}(z_{i1}) = \operatorname{sign}(z_{k1})$, $\forall i, k$, i.e., all agents favor the same strategy.

iii) Suppose v_{max} (w_{max}) is also a left (right) eigenvector of both A and $\hat{A}$. Denote by λ, $\hat{\lambda}$ the eigenvalues of A, $\hat{A}$, respectively, corresponding to $v_{max}(w_{max})$. Then $u^* = \frac{1 - \frac{1}{4}\eta^{-1}p\hat{\lambda}}{S'(0)(\alpha + \gamma\lambda)}$, and the unfolding parameter (11) simplifies to

$$b = d\left(\frac{1}{4}p^{\perp}\hat{\lambda} + b_1 - b_2\right)\langle w_{max}, \mathbf{1}\rangle. \quad (12)$$

The following theorem shows how the bifurcation depends on degree (number of neighbors) for G, $\hat{G}$ regular graphs.

Theorem 3: Suppose $\gamma > 0$, $p \geq 0$, and G, $\hat{G}$ are undirected, connected, and regular with degrees K, $\hat{K}$, respectively. The bifurcation point u^* and unfolding parameter b satisfy $\operatorname{sign}(\frac{\partial u^*}{\partial K}) = \operatorname{sign}(\frac{1}{4}\eta^{-1}p\hat{K} - 1)$, $\operatorname{sign}(\frac{\partial u^*}{\partial \hat{K}}) = \operatorname{sign}(-p)$, and $\operatorname{sign}(\frac{\partial b}{\partial \hat{K}}) = \operatorname{sign}(p^{\perp})$.

Remark 1: For games with more than 2 strategies and heterogeneous payoff functions, the analysis can be generalized using $U_{ij}(x) = \sum_{(i,k)\in\hat{\mathbb{E}}}\begin{pmatrix} p_{j1}^{ik} & \cdots & p_{jN_s}^{ik}\end{pmatrix}x_k + b_j^i$.

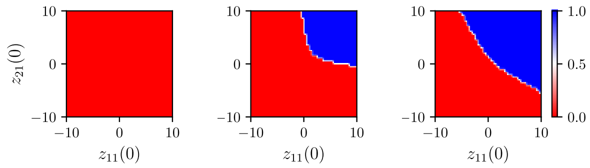
In what follows, we discuss implications of Theorems 2, 3 in social dilemmas using the prisoner's dilemma and public goods game. From now on, we take $S(\cdot) = \tanh(\cdot)$.

Prisoner's dilemma: Let $p_{11} = p_{CC}$, $p_{12} = p_{CD}$, $p_{21} = p_{DC}$, $p_{22} = p_{DD}$ and $b_1 = b_2 = 0$ so (8) specializes to (2).

Corollary 1: For $\gamma > 0$, the following hold: $v_{max} = w_{max} = (1, 1)$, $\lambda = \hat{\lambda} = 1$, and $p^{\perp} < 0$. Hence, we have $u^* = \frac{1 - \frac{1}{4}\eta^{-1}p}{\alpha + \gamma}$ and $b = \frac{1}{2}dp^{\perp} < 0$.

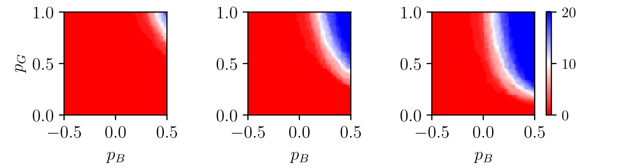
Figs. 1(a), 1(b) show the bifurcation diagram (plot of equilibria as a function of bifurcation parameter $\tilde{u} = u - u^*$) of the Lyapunov-Schmidt reduction (16) of (9), for two values of p_{CC}. p and $p^{\perp}$ have a two-fold effect: *i)* p changes the location of the pitchfork bifurcation point in the $\tilde{u}$-axis. *ii)* Since $p^{\perp} < 0$, the pitchfork bifurcation unfolds favoring the branch of solutions corresponding to mutual defection. For sufficiently large $\tilde{u}$ (equivalently, u), a branch of solutions corresponding

[6] This holds, e.g., when $\gamma > 0$, $p \geq 0$, and at least one of A, $\hat{A}$ corresponds to a connected graph.



(a) $r = 20$ (b) $r = 10$ (c) $r = 0$

Fig. 2. Heatmaps depict the probability of mutual cooperation in the prisoner's dilemma for three different values of r in (13).



(a) $a = 40$ (b) $a = 20$ (c) $a = 10$

Fig. 3. Heatmaps depict the average number of agents cooperating in the public goods game for three different values of a in (14).

to mutual cooperation emerges, and the larger the u the larger its region of attraction. A smaller u is required for larger p_{CC}, since larger p_{CC} decreases incentive to defect.

If $p_{11} > p_{21} = 40$, the game is the Stag Hunt where the strategy to hunt a stag replaces cooperation and the strategy to hunt a hare replaces defection. Coordinated stag hunting and coordinated hare hunting are both Nash equilibria, the former payoff-dominating and the latter risk-dominating. The model predicts the larger the u, the larger the region of attraction to coordinated stag hunting.

Public goods game: Let $N_s = 2$, i.e., each agent decides to cooperate and contributes its entire wealth ($j = 1$), or defect and contributes nothing ($j = 2$). Note that (8) specializes to (3) by selecting $p_{11} = p_{21} = b_1 = a\rho/N_a$, $p_{12} = p_{22} = 0$, and $b_2 = a$ with all-to-all game interaction graph $\hat{G}$.

Corollary 2: With $N_s = 2$, $p = p^{\perp} = 0$. For $\gamma > 0$ and connected social interaction graph G, the following hold: *i)* v_{max}, w_{max} have all nonzero same-sign entries, $u^* = \frac{1}{\alpha+\gamma\lambda}$ and $b = -da(1-\rho/N_a)\langle w_{max}, \mathbf{1}\rangle < 0$. *ii)* When A is regular with degree K, it holds that $\frac{\partial u^*}{\partial K} < 0$, i.e., with larger K (more social interactions), bistability requires less attention u.

Figs. 1(c), 1(d) show the bifurcation diagram for two values of total wealth a. Since $p = 0$, p has no effect. However, $b_1 - b_2 = -a(1 - \rho/N_a) < 0$; hence, for reciprocating agents ($\gamma > 0$), the pitchfork bifurcation unfolds towards the branch of solutions corresponding to no agent contributing to the public pool. Since the strength of the unfolding is proportional to a, emergence of the mutually cooperative solution, when all agents contribute, requires a smaller u for smaller a and for a fixed u its region of attraction grows as a decreases.

IV. Numerical Studies

A. Prisoner's Dilemma

We set $d_i = \eta = 1$, $u_i = 10$, $A^j_{ii} = 0$, $A^j_{ik} = 1$ for $i \neq k$, and $S(\cdot) = \tanh(\cdot)$. Consider the payoff matrix (2) given by

$$\begin{pmatrix} p_{CC} & p_{CD} \\ p_{DC} & p_{DD} \end{pmatrix} = \begin{pmatrix} 35 & 0-r \\ 40+r & 5 \end{pmatrix} \tag{13}$$

with $r > 0$ an extra reward (penalty) an agent receives if it defects (cooperates) while its opponent cooperates (defects).

Using simulations, we illustrate limit points of the opinion state trajectories predicted by the theory. In Fig. 2 each heatmap illustrates the probability of both agents cooperating and the two axes represent the initial opinion states $z_{11}(0), z_{21}(0)$ of the agents associated with the cooperation strategy. Since the two agents are reciprocating, for all cases, we observe that the heatmaps for both agents are identical, and hence we only present that of agent 1.

In Figs. 2(b) and 2(c), we can observe that when both agents are *nice*, i.e., the agents' initial opinion states $z_{11}(0) > 0$, $z_{21}(0) > 0$ for the cooperation are large enough, they can maintain mutual cooperation. Also, a sufficiently nice agent ($z_{i1}(0) > 0$) *forgives* the exploiting behavior (defection) of its opponent that initially is not nice ($z_{-i1}(0) < 0$). However, when its opponent has a strong intention to defect ($z_{-i1}(0) < 0$ substantially large in magnitude), the agent also defects to avoid being exploited and is thus *provocable*.

An increase in r motivates the agents to defect (Fig. 2). When $r = 20$, since $p_{DC} - p_{CC} = p_{DD} - p_{CD} = 25 > 2u_i = 20$, the cooperation strategy is dominated by the defection strategy, and both agents eventually defect (Fig. 2(a)). Thus, as predicted by the theory and illustrated in Figs. 1(a), 1(b), when there is a strong enough incentive to defect, the level of attention u to opinion exchanges, which translates into the level of reciprocity, may be insufficient to prevent the agents from pursuing individually rational decision-making.

B. Public Goods Game

For the 2-strategy public goods game, we adopt the same parameters of (9) as in Section IV-A except that $u_i = 5$, the inter-agent interactions are governed by the Erdös-Rényi graph with parameter p_G (for $i \neq k$, $A^j_{ik} = 1$ with probability p_G and $A^j_{ik} = 0$ with probability $1 - p_G$), and the initial opinion state of each agent is uniformly randomly selected as $z_{i1}(0) \sim \text{Uniform}(-0.5 + p_B, 0.5 + p_B)$, where p_B is a bias in favor of cooperation. Let $\rho = 2$, $N_a = 20$ so (3) is

$$U_{ij}(x) = \begin{cases} \frac{a}{10} + \frac{a}{10}\sum_{k\neq i, k=1}^{20} x_{k1} & \text{if } j = 1 \\ a + \frac{a}{10}\sum_{k\neq i, k=1}^{20} x_{k1} & \text{if } j = 2. \end{cases} \tag{14}$$

We evaluate opinion state trajectories over a range of values of p_G, p_B, and a to explore how the network structure of the social interaction, initial opinion states, and total wealth tune the emergence of cooperation as predicted by the theory.

Each heatmap in Fig. 3 depicts, for a given a, the average number of agents that cooperate at steady-state for a range of p_G, p_B. Both network structure, determined by p_G, and agents' initial preference to contribute to the public pool, determined by p_B, play important roles: The cooperation among the 20 agents is more likely to be sustained if each agent has a greater chance to interact with others (p_G large) and favors cooperation at the beginning of the game (p_B large). Interestingly, even if they prefer to cooperate at the beginning (p_B large), when the agents are interacting less and cannot perceive the opinion state of others (p_G small), they decide to defect over time. The advantage of large p_G is as for large K for regular graphs, as predicted by Corollary 2.

The payoff difference $U_{i2}(x) - U_{i1}(x) = 0.9a$ between the two strategies depends on the total wealth a and quantifies the incentive for the agents to defect. Consequently, the more wealth agents have, the higher incentive they receive to not contribute. This is illustrated in Fig. 3, where mutual defection (cooperation) is more (less) likely as a increases.

V. Final Remarks

We have shown that the nonlinear opinion dynamics model of [7], [8], [18] provides an analytically tractable framework for studying cooperative behavior in repeated multi-agent games, where agents rely on rationality *and* reciprocity, both of which are central to human decision-making. The opinion update depends on a saturated function of inter-agent opinion exchanges, which allows mutual cooperation (or coordination) to emerge as one of two bistable equilibria in two-strategy games. For the prisoner's dilemma and multi-agent public goods game, mutual cooperation emerges when the attention u to social interaction, and thus reciprocity, is sufficiently strong. The bistability provides a possible mathematical account for how reciprocity enables stable cooperative behavior, as observed in experimental studies, and a principled approach for tuning cooperative behavior.

Building on coupled opinion-attention dynamic analysis of [7], [18], we will design feedback dynamics for u to reflect, for instance, agents' growing appreciation of social interactions. This will allow opportunities to influence behavior, e.g., to elicit cooperation or coordination among agents. We will also leverage the versatility of the model to investigate games with more than two strategies and heterogeneity.

Appendix

Proof of Theorem 1:

i) The first statement is verified by comparing (4) and (9). For the second statement, by the definition of S and $z_{ij}(0) = \bar{z}_{ij}(0) - \frac{1}{N_s}\sum_{l=1}^{N_s}\bar{z}_{il}(0)$ we get $\frac{\mathrm{d}}{\mathrm{d}t}(\bar{z}_{ij}(t) - \frac{1}{N_s}\sum_{l=1}^{N_s}\bar{z}_{il}(t)) = F_{ij}(z) - \frac{1}{N_s}\sum_{l=1}^{N_s}F_{il}(z)$. Therefore, $\bar{z}_{ij}(t) - \frac{1}{N_s}\sum_{l=1}^{N_s}\bar{z}_{il}(t)$ is a solution of (9) and hence $z_{ij}(t) = \bar{z}_{ij}(t) - \frac{1}{N_s}\sum_{l=1}^{N_s}\bar{z}_{il}(t)$. Thus, $\sigma_j(z_i(t)) = \sigma_j(\bar{z}_i(t))$ for all $t \geq 0$ and $\bar{z}(t)$ is a solution to (4).

ii) If $\bar{z}^*$ is an equilibrium of (4) then $z_i^* = P_0\bar{z}_i^*$ satisfies $P_0F_i(z^*) = 0$ and hence is an equilibrium of (9). To prove the second statement, suppose z^* is an equilibrium of (9). As in the proof for **i)**, we can establish that $z_{ij}^* = \bar{z}_{ij}^* - \frac{1}{N_s}\sum_{l=1}^{N_s}\bar{z}_{il}^*$ for $\bar{z}_{ij}^*$ defined as in the statement. Thus, $\sigma_j(z_i^*) = \sigma_j(\bar{z}_i^*)$ and $\bar{z}^*$ is an equilibrium of (4). The stability of the equilibria follows from **i)**. ∎

Proof of Theorem 2:

i) When $p^\perp = b_1 = b_2 = 0$, the neutral state $\mathbf{z} = 0$ is always an equilibrium of (10). The Jacobian of the linearization of (10) at $\mathbf{z} = 0$ is

$$J(0) = -d\left((1 - uS'(0)\alpha)I - u\gamma S'(0)A - \frac{1}{4}\eta^{-1}p\hat{A}\right) \quad (15)$$

and its eigenvalues take the form $\mu_i = -d(1 - uS'(0)\alpha - \zeta_i(u, \gamma, p))$ where ζ_i is an eigenvalue of the matrix $u\gamma S'(0)A + \frac{1}{4}\eta^{-1}p\hat{A}$. By [19], we can derive that $\frac{\partial\zeta_{max}}{\partial u} = w_{max}^T\gamma S'(0)Av_{max}$, and $\frac{\partial\mu_{max}}{\partial u} = dS'(0)\alpha + w_{max}^T\gamma S'(0)Av_{max} > 0$ for any u, p, γ. Hence, there exists a critical value u^* for which if $u < u^*$, all eigenvalues of (15) have negative real part, and if $u > u^*$, μ_{max} is positive, real, and simple. By Lyapunov-Schmidt reduction [20], the one-dimensional dynamics projected onto span of v_{max} are

$$\begin{aligned}\dot{z}_c = &-2d\langle w_{max}, \tilde{v}\rangle z_c^3 + dS'(0)\langle w_{max}, (\alpha I + \gamma A)v_{max}\rangle \\ &\times \tilde{u}z_c + \langle w_{max}, \frac{1}{4}dp^\perp\hat{A}\mathbf{1} + d(b_1 - b_2)\mathbf{1}\rangle + h.o.t. \quad (16)\end{aligned}$$

where $\tilde{v} = v_{max} \odot (\alpha I + \gamma A + \frac{1}{4}p\hat{A})v_{max} \odot (\alpha I + \gamma A + \frac{1}{4}p\hat{A})v_{max}$ and $\tilde{u} = u - u^*(\alpha, \gamma, p, \eta)$. By the recognition problem [20, Ch. 2, Proposition 9.2], (16) describes an unfolding of the pitchfork bifurcation. The last statement follows by implicit differentiation of $-1 + \alpha S'(0)u^* + \zeta_{max} = 0$.

ii) By the Perron-Frobenius theorem, v_{max} and w_{max} have all same-sign entries. The rest follows from part **i)** and the center manifold theorem.

iii) By the assumptions on v_{max} (w_{max}), $\zeta_{max} = u\gamma S'(0)\lambda + \frac{1}{4}\eta^{-1}p\hat{\lambda}$, $\mu_{max} = -d(1 - uS'(0)\alpha - u\gamma S'(0)\lambda - \frac{1}{4}\eta^{-1}p\hat{\lambda})$. Thus, $u^* = \frac{1 - \frac{1}{4}\eta^{-1}p\hat{\lambda}}{S'(0)(\alpha + \gamma\lambda)}$. The rest follows from (11) since w_{max} is the left eigenvector of $\hat{A}$. ∎

Proof of Theorem 3: By the assumptions on p, G, and $\hat{G}$, we can verify that v_{max}, w_{max}, λ, and $\hat{\lambda}$, given in Theorem 2 **iii)**, satisfy $v_{max} = w_{max} = \mathbf{1}$, $\lambda = K$, and $\hat{\lambda} = \hat{K}$; then $\frac{\partial u^*}{\partial K} = \frac{-\gamma(1 - \frac{1}{4}\eta^{-1}p\hat{K})}{S'(0)(\alpha + \gamma K)^2}$ and $\frac{\partial u^*}{\partial \hat{K}} = \frac{-\frac{1}{4}\eta^{-1}p}{S'(0)(\alpha + \gamma K)}$. From (12), $\frac{\partial b}{\partial \hat{K}} = \frac{1}{4}N_a dp^\perp$, and the theorem follows. ∎

References

[1] A. W. Gouldner, "The norm of reciprocity: A preliminary statement," *Amer. Sociol. Rev.*, vol. 25, no. 2, pp. 161–178, 1960.

[2] R. Axelrod, *The Evolution of Cooperation*. New York, NY, USA: Basic Books, 1984.

[3] E. Fehr and S. Gächter, "Reciprocity and economics: The economic implications of homo reciprocans," *Eur. Econ. Rev.*, vol. 42, nos. 3–5, pp. 845–859, 1998.

[4] P. Kollock, "Social dilemmas: The anatomy of cooperation," *Annu. Rev. Sociol.*, vol. 24, no. 1, pp. 183–214, 1998.

[5] L. Heuer and A. Orland, "Cooperation in the prisoner's dilemma: An experimental comparison between pure and mixed strategies," *Royal Soc. Open Sci.*, vol. 6, no. 7, 2019, Art. no. 182142.

[6] A. Mao, L. Dworkin, S. Suri, and D. J. Watts, "Resilient cooperators stabilize long-run cooperation in the finitely repeated Prisoner's Dilemma," *Nat. Commun.*, vol. 8, no. 1, 2017, Art. no. 13800.

[7] A. Bizyaeva, A. Franci, and N. Leonard, "Nonlinear opinion dynamics with tunable sensitivity," 2020, *arXiv:2009.04332*.

[8] R. Gray, A. Franci, V. Srivastava, and N. E. Leonard, "Multiagent decision-making dynamics inspired by honeybees," *IEEE Trans. Control Netw. Syst.*, vol. 5, no. 2, pp. 793–806, Jun. 2018.

[9] B. Gao and L. Pavel, "On passivity, reinforcement learning, and higher order learning in multiagent finite games," *IEEE Trans. Autom. Control*, vol. 66, no. 1, pp. 121–136, Jan. 2021.

[10] P. Mertikopoulos and W. H. Sandholm, "Learning in games via reinforcement and regularization," *Math. Oper. Res.*, vol. 41, no. 4, pp. 1297–1324, 2016.

[11] M. H. DeGroot, "Reaching a consensus," *J. Amer. Stat. Assoc.*, vol. 69, no. 345, pp. 118–121, 1974.

[12] S. R. Etesami and T. Başar, "Game-theoretic analysis of the Hegselmann-Krause model for opinion dynamics in finite dimensions," *IEEE Trans Autom. Control*, vol. 60, no. 7, pp. 1886–1897, Jul. 2015.

[13] D. Bauso and M. Cannon, "Consensus in opinion dynamics as a repeated game," *Automatica*, vol. 90, pp. 204–211, Apr. 2018.

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

[15] J. Hofbauer and W. H. Sandholm, "Evolution in games with randomly disturbed payoffs," *J. Econ. Theory*, vol. 132, no. 1, pp. 47–69, 2007.

[16] H.-C. Chen, J. W. Friedman, and J.-F. Thisse, "Boundedly rational Nash equilibrium: A probabilistic choice approach," *Games Econ. Behav.*, vol. 18, no. 1, pp. 32–54, 1997.

[17] J. Conlisk, "Why bounded rationality?" *J. Econ. Literature*, vol. 34, no. 2, pp. 669–700, 1996.

[18] A. Franci, A. Bizyaeva, S. Park, and N. E. Leonard, "Analysis and control of agreement and disagreement opinion cascades," *Swarm Intell.*, vol. 15, no. 1, pp. 47–82, 2021.

[19] R. Plaut and K. Huseyin, "Derivatives of eigenvalues and eigenvectors in non-self-adjoint systems," *AIAA J.*, vol. 11, no. 2, pp. 250–251, 1973.

[20] M. Golubitsky and D. Schaeffer, *Singularities and Groups in Bifurcation Theory*, vol. 1. New York, NY, USA: Springer, 1985.