Strategy Competition Dynamics of Multi-Agent Systems in the Framework of Evolutionary Game Theory

Zhang, Jianlei; Cao, Ming

Published in:
IEEE Transactions on Circuits and Systems. II: Express Briefs

DOI:
10.1109/TCSII.2019.2910893

Strategy Competition Dynamics of Multi-Agent Systems in the Framework of Evolutionary Game Theory

Jianlei Zhang, *Member, IEEE*, and Ming Cao, *Senior Member, IEEE*

***Abstract*—There is the recent boom in investigating the control of evolutionary games in multi-agent systems, where personal interests and collective interests often conflict. Using evolutionary game theory to study the behaviors of multi-agent systems yields an interdisciplinary topic which has received an increasing amount of attention. Findings in real-world multi-agent systems show that individuals have multiple choices, and this diversity shapes the emergence and transmission of strategy, disease, innovation, and opinion in various social populations. In this sense, the simplified theoretical models in previous studies need to be enriched, though the difficulty of theoretical analysis may increase correspondingly. Here, our objective is to theoretically establish a scenario of four strategies, including competition among the cooperatives, defection with probabilistic punishment, speculation insured by some policy, and loner. And the possible results of strategy evolution are analyzed in detail. Depending on the initial condition, the state converges either to a domination of cooperators, or to a rock-scissors-paper type heteroclinic cycle of three strategies.**

***Index Terms*—Game theory, multi-agent system, evolution dynamics.**

I. INTRODUCTION

THERE is burgeoning study in the networked systems and control theory in applications ranging from distributed robotics to epidemic control and decision making of humans [1]–[3]. When the agents have competing objectives, as is often the case, each agent must consider the actions of her competitors; in such cases single-objective optimization methods fail. Especially, situations in which the private interest can be at odds with the public interest constitute an important class of societal problems. Evolutionary game theory is an interdisciplinary mathematical tool which seems to be able to embody several relevant features of the problem and, as such, is used in much cooperation-oriented research. In particular, the oft-cited public goods game [4]–[7] is a paradigm example for investigating the emergence of cooperation in spite of the fact that self-interest seems to dictate defective behavior.

As a cross-cutting topic, many solutions for this multi-agent cooperative dilemma in multi-agent systems have been discussed [8], [9]. The theory of kin selection focuses on cooperation among individuals that are genetically closely related, whereas theories of direct reciprocity focus on the selfish incentives for cooperation in bilateral long-term interactions [10]–[13]. The theories of indirect reciprocity and costly signalling indicate how cooperation in larger groups can emerge when the cooperators can build a reputation [14], [15]. Current research has also highlighted two factors boosting cooperation in public goods interactions, namely, punishment of defectors [16], [17] and the option to abstain from the joint enterprise. Voluntary participation [18] allows individuals to adopt a risk-aversion strategy, termed loner. A loner refuses to participate in unpromising public enterprises and instead relies on a small but fixed payoff.

For the multi-agent systems, the individual heterogeneity and biological or social diversity are also well-known phenomena in nature [19], [20]. It is intriguing to investigate whether and how biodiversity affects the emergence and transmission of strategy, disease, innovation, opinion and so on. The potential difficulties brought by individual heterogeneity in mathematical modeling, raise challenges for existing theoretical models which only consider relatively simple (in strategy types, decision-making modes, etc) agents in games. However, this is an unavoidable direction and many more studies concerning with the individual heterogeneity or diversity, in the framework evolutionary game theory, are expected to appear in the near future. Only in this way could we gain more insight into a series of perplexing puzzles about cooperative phenomena in the multi-agent systems.

In this line of research, based on the punishment in the strategy competition [21], [22], our previous work [23] goes a step further by proposing another behavior type named as speculation. Results indicate scenarios where speculation either leads to the reduction of the basin of attraction of the cooperative equilibrium or even the loss of stability of this equilibrium, if the costs of the insurance are lower than the expected fines faced by a defector.

Further, agents often have multiple choices in decision making due to the individual personality, especially when facing the potential punishment if defecting. For example, resolute defectors will persist in their defection strategy, though taking the risk of being punished with a probability. Speculators incline to buy an insurance policy covering the costs of punishment when caught defecting. While timid loners will

Manuscript received December 15, 2018; revised February 25, 2019 and March 14, 2019; accepted April 4, 2019. This work was supported in part by the National Natural Science Foundation of China under Grant 61603201, Grant 61603199, and Grant 91848203, in part by the Tianjin Natural Science Foundation of China under Grant 18JCYBJC18600, in part by the European Research Council under Grant ERC-CoG-771687, and in part by the Dutch Technology Foundation (STW) under Grant vidi-14134. This brief was recommended by Associate Editor J. Wu. *(Corresponding author: Ming Cao.)*

J. Zhang is with the Department of Automation, College of Artificial Intelligence, Nankai University, Tianjin 300071, China.

M. Cao is with the Research Institute of Engineering and Technology, University of Groningen, 9747AG Groningen, The Netherlands (e-mail: ming.cao@gmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCSII.2019.2910893

conservatively obtain an autarkic income independent of the other players' decision. These mentioned choices can better represent the possible attempts to raise money for public goods in complicated real-life situations. With this formulation, as an extension of our previous work proposing speculation [23], the fourth strategy (i.e., loner, a player can refuse to participate and get some small but fixed income) is also provided for the players. As mentioned, it is based on the assumption that players can voluntarily decide whether to participate in the joint game or not.

So altogether we consider four behavior types, which enrich the model and meanwhile raise the difficulty of theoretical analysis. (a) The cooperators join the group and to contribute their effort. (b) The defectors join, but do not contribute; moreover, defectors are caught with a certain probability and a fine is imposed on them when caught. Here we are less interested in the specific establishment of an effective system of punishment, but rather in the two additional options (speculation and loner) found in several systems. To be more specific, we consider the public goods game with an external punishment system as indicated above. (c) The speculators purchase an insurance policy covering the costs of punishment when caught defecting. It means that by paying a fixed cost for their insurance policy, speculators can defect without paying any fine from punishment. (d) The loners are unwilling to join the game, but prefer to rely on a small but fixed payoff. By means of a theoretical approach, we investigate the joint evolution of multiple strategies and the stability of the evolving system.

II. Problem Formulation

In a typical public goods game (PGG) played in interaction groups of size N, each player receives an endowment c and independently decides how much of it to be contributed to a public goods system. Then the collected sum is multiplied by an amplification factor r ($1 < r < N$) and is redistributed to the group members, irrespective of her strategy. The maximum total benefit will be achieved if all players contribute maximally. In this case each player receives rc, thus the final payoff is $(r-1)c$. Players are faced with the temptation of taking advantage of the public goods without contributing. In other words, any individual investment is a loss for the player because only a portion $r/N < 1$ will be repaid. Consequently, rational players invest nothing-hence a collective dilemma occurs.

This brief is based on the PGG played in interaction groups of size N, consisting of by cooperators, defectors, speculators, and loners. To be precise, each participant (except loners) gains an equal benefit rcx_c ($c > 0$) which is proportional to the fraction of cooperators (x_c, $0 \leq x_c \leq 1$) among the players. Cooperators pay a fixed cost c to the public goods. Defectors contribute nothing, but may be caught and fined by α ($\alpha > 0$). Speculators neither contribute to common goods nor pay a fine when caught, instead they pay an amount λ ($\lambda > 0$) to the insurance policy. Loners obtain a fixed pay-off σ ($0 < \sigma$) from a solitary pursuit without participating and contributing.

Assuming for theoretical analysis, from time to time, sample groups of N such players are chosen randomly from a very large, well-mixed system. Notably, the probability that two players in large populations ever encounter again can be neglected.

Within such a group, if N_c ($0 \leq N_c \leq N$) denotes the number of cooperators and N_l ($0 \leq N_l \leq N$) is the number of loners among the public goods players, the net payoffs of the four strategies are respectively given by

$$\begin{cases} P_c = \frac{rcN_c}{N-N_l} - c \\ P_d = \frac{rcN_c}{N-N_l} - \alpha \\ P_s = \frac{rcN_c}{N-N_l} - \lambda \\ P_l = \sigma. \end{cases} \tag{1}$$

In this game, each unit of investment is multiplied by r ($0 < r < N$) and the product is distributed among all participants (except loners) irrespective of their strategies. The first term in the expression represents the benefit that the agent obtains from the public goods, while the second term denotes cost.

We first derive the probability that n of the N sampled individuals are actually willing to join the public goods game. In the case $n = 1$ (no co-player shows up) we assume that the player has no other option than to play as a loner, and obtains payoff σ. This happens with probability x_l^{N-1}. Here, x_l is the fraction of loners. For a given player (C, D or S) willing to join the public goods game, the probability of finding, among the $N-1$ other players in the sample, $n-1$ co-players joining the group ($n > 1$), is given by

$$\binom{N-1}{n-1}(1-x_l)^{n-1}(x_l)^{N-n}. \tag{2}$$

The probability that m of these players are cooperators is

$$\binom{n-1}{m}\left(\frac{x_c}{x_c+x_d+x_s}\right)^m\left(\frac{x_d+x_s}{x_c+x_d+x_s}\right)^{n-1-m}. \tag{3}$$

where x_c, x_d, x_s respectively denote the fractions of cooperators, defectors and speculators in the population.

For simplicity and without loss of generality, we set the cost c of cooperation equal to 1. In the above case, the payoff for a defector is $rm/n - \alpha$, while the payoffs for a cooperator and a speculator are respectively specified by $r(m+1)/n - 1$ and $rm/n - \lambda$. Hence, the expected payoff for a defector in such a group is:

$$\begin{aligned}&\left(\frac{rm}{n}-\alpha\right)\sum_{m=0}^{n-1}\binom{n-1}{m}\left(\frac{x_c}{1-x_l}\right)^m\left(1-\frac{x_c}{1-x_l}\right)^{n-m-1}\\ &= \frac{r}{n}\cdot(n-1)\frac{x_c}{1-x_l}-\alpha.\end{aligned}$$

The payoff of a cooperator in a group of n players is:

$$\begin{aligned}&\left[\frac{r(m+1)}{n}-1\right]\sum_{m=0}^{n-1}\binom{n-1}{m}\left(\frac{x_c}{1-x_l}\right)^m\left(1-\frac{x_c}{1-x_l}\right)^{n-m-1}\\ &= \frac{r}{n}\cdot(n-1)\frac{x_c}{1-x_l}+\frac{r}{n}-1.\end{aligned}$$

The payoff of a speculator in a group of n players is:

$$\begin{aligned}&\left(\frac{rm}{s}-\lambda\right)\sum_{m=0}^{N-1}\binom{n-1}{m}\left(\frac{x_c}{1-x_l}\right)^m\left(1-\frac{x_c}{1-x_l}\right)^{n-m-1}\\ &= \frac{r}{n}\cdot(n-1)\frac{x_c}{1-x_l}-\lambda.\end{aligned}$$

The payoff of a loner is the constant value of σ.

Then, the expected payoff for a defector in the population is,

$$\begin{aligned}P_d &= \sigma x_l^{N-1} + \sum_{n=2}^{N}\left[\frac{r}{n}\cdot(n-1)\frac{x_c}{1-x_l}-\alpha\right]\binom{N-1}{n-1}\\ &\quad (1-x_l)^{n-1}(x_l)^{N-n}\\ &= \sigma x_l^{N-1} + \frac{rx_c}{1-x_l}\left[1-\frac{1-x_l^N}{N(1-x_l)}\right]-\alpha(1-x_l^{N-1}).\end{aligned} \tag{4}$$

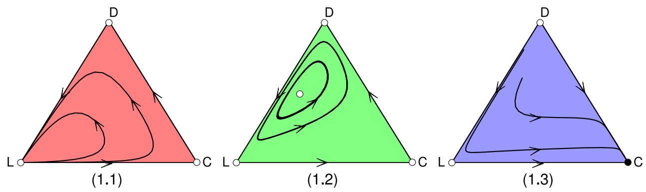

Fig. 1. The evolution dynamics results of $T = (C, D, L)$, where in the absence of speculation. (1.1): $r < 2-2\alpha$. (1.2): $r > 2-2\alpha$; and (1.3): $1-r/N-\alpha < 0$. Parameters: $N = 5, \sigma = 0.3$, and $r = 1.6, \alpha = 0.1$ for (1.1); $r = 3, \alpha = 0.1$ for (1.2); $r = 3, \alpha = 0.5$ for (1.3). Open dots are unstable equilibrium points and closed dots are stable equilibrium points. Three corners represent a rock-scissors-paper type heteroclinic cycle if $1 - r/N - \alpha > 0$ (cases 1.1 and 1.2) while full-C is a global attractor if $1 - r/N - \alpha < 0$ (case 1.3).

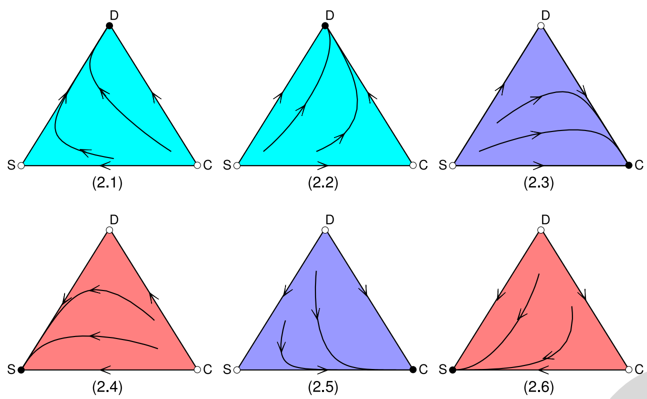

Fig. 2. The evolution dynamics results of $T = (C, D, S)$, where in the absence of defection. We consider six cases, which are discussed in cases 2.1 till 2.3 in the upper panel of Fig. 2. Fig. 2 focuses on the situation $\lambda - \alpha > 0$ implying that the fine for defectors is higher than the costs of cooperation. Lower panels of Fig. 2 considers the opposite case $\lambda - \alpha < 0$, where defection is the dominating strategy. Results show that there is always a global attractor in the system, and the outcome of the game dynamics depends on model parameters. Parameters: $N = 5, r = 3, \sigma = 0.3$, and $\alpha = 0.1, \lambda = 0.2$ for (2.1); $\alpha = 0.1, \lambda = 0.8$ for (2.2); $\alpha = 0.5, \lambda = 0.8$ for (2.3); $\alpha = 0.1, \lambda = 0.2$ for (2.4); $\alpha = 0.8, \lambda = 0.5$ for (2.5); $\alpha = 0.8, \lambda = 0.1$ for (2.6).

In the continuous time model, the evolution of the fractions of the four strategies proceeds according to

$$\dot{x}_i = x_i(P_i - \bar{P}), \tag{5}$$

where i can be c, d, s, l, P_i is the payoff of strategy i, and $\bar{P} = x_c P_c + x_d P_d + x_s P_s + x_l \sigma$.

III. Theoretical Analysis

We firstly focus on the replicator dynamics starting from a three-strategy state in the population, then we pay attention to analyzing the output when all the four strategies initially exist in the population. For the replicator dynamics of three-strategy evolution, we comprehensively consider four scenarios depicted in Figs. 1-4 as follows. The advantage of one strategy over another depends on the payoff difference between them, hence

$$\begin{aligned} P_d - P_c &= \sum_{n=2}^{N} [1 - \frac{r}{n} - \alpha] \binom{N-1}{n-1} (1 - x_l)^{n-1} (x_l)^{N-n} \\ &= 1 - \alpha + (r - 1 + \alpha) x_l^{N-1} - \frac{r}{N} \frac{1 - x_l^N}{1 - x_l}, \end{aligned} \tag{6}$$

$$\begin{aligned} P_d - P_s &= \sum_{n=2}^{N} [\lambda - \alpha] \binom{N-1}{n-1} (1 - x_l)^{n-1} (x_l)^{N-n} \\ &= (\lambda - \alpha)(1 - x_l^{N-1}), \end{aligned} \tag{7}$$

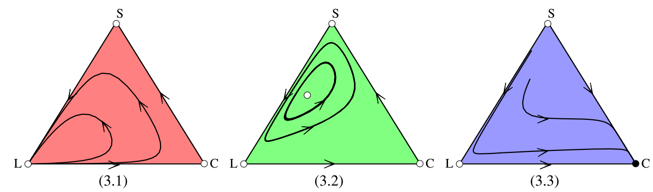

Fig. 3. The evolution dynamics results of $T = (C, S, L)$, where in the absence of speculation. (3.1): $r < 2-2\lambda$. (3.2): $r > 2-2\lambda$; and (3.3): $1-r/N-\lambda < 0$. Parameters: $N = 5, \sigma = 0.3$, and $r = 1.6, \lambda = 0.1$ for (3.1); $r = 3, \lambda = 0.1$ for (3.2); $r = 3, \lambda = 0.5$ for (3.3). Three corners here represent a rock-scissors-paper type heteroclinic cycle if $1 - r/N - \lambda > 0$ (cases 3.1 and 3.2) while pure cooperation is a global attractor if $1 - r/N - \lambda < 0$ (case 3.3).

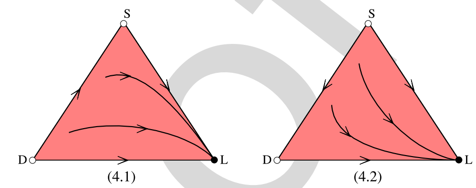

Fig. 4. The evolution dynamics results of $T = (D, L, S)$ where in the absence of cooperation.(4.1) resulting game dynamics in the absence of speculation, where pure loners is the only global attractor in the system. Parameters: $N = 5$, $r = 3$, $\sigma = 0.3$, and $\alpha = 0.4, \lambda = 0.1$ for (3); $\alpha = 0.4, \lambda = 0.1$ for (4.1); $\alpha = 0.1, \lambda = 0.4$ for (4.2).

$$P_s - P_c = 1 - \lambda + (r - 1 + \lambda) x_l^{N-1} - \frac{r}{N} \frac{1 - x_l^N}{1 - x_l}. \tag{8}$$

In the above calculations, $N > 1$, $1 < r < N$ and $\alpha > 0$. The sign of $P_i - P_j$ in fact determines whether it pays to switch from cooperation to defection or vice versa, $P_i - P_j = 0$ being the equilibrium condition, where i, j can be strategy C, D, S, and L.

We now proceed to the study of evolutionary dynamics when $\lambda \neq \alpha$ where four strategies coexist in the population; the point in the phase space corresponding to such a state is, referred to as an interior point. We make the following three assumptions and want to show the results that at least one strategy will become extinct with the evolution of the system initialized from an interior point.

Theorem 1: If $\lambda \neq \alpha$, at least one strategy will become extinct with the evolution of the system initialized from an interior point. Here, an interior point means that the fraction of every strategy is larger than zero.

Proof: We now analyze the system in different situations.

(1) When $\lambda \neq \alpha$, supposing $\lambda > \alpha$ (i.e., $P_d > P_s$), when $x_l \neq 0$. We suppose that there is a closed set, meaning that the subsequent evolving state of each initial state in this set also belongs to this set. So $x_c > 0$, $x_d > 0$, $x_s > 0$ and $x_l > 0$ in this closed set.

(1.1) We first take one point $(x_c^*, x_d^*, x_s^*, x_l^*)$ in this closed set such that $x_c^* > 0, x_d^* > 0, x_s^* > 0, x_c^* > 0$, and $\dot{x}_c^* = \dot{x}_d^* = \dot{x}_s^* = \dot{x}_l^* = 0$, thus

$$\begin{cases} \dot{x}_d^* = x_d^*(p_d^* - \bar{p}^*) \\ \dot{x}_s^* = x_s^*(p_s^* - \bar{p}^*). \end{cases} \tag{9}$$

Herein, the result $\dot{x}_d^* = \dot{x}_s^* = 0$ needs $\dot{p}_d^* = \bar{p}^* = \dot{p}_s^*$, which contradicts with $\dot{p}_d^* - \dot{p}_s^* > 0$. Therefore we can safely get the conclusion that there is no interior stable point.

(1.2) We next assume that the interior domain is a limit cycle. In this case, the four strategy players will gain the same average payoffs driven by the replicator equation, where $\bar{p}_c = \bar{p}_d = \bar{p}_s = \bar{p}_l$. However, $\bar{p}_d = \bar{p}_s$ contradicts with $p_d > p_s$, indicating that the closed set is not a limit cycle.

(1.3) We then verify whether the interior domain contains chaotic solutions, where also $x_c > 0, x_d > 0, x_s > 0, x_l > 0$. By introducing the fraction of defections in a population consisting of defectors and speculators, $f = \frac{x_d}{x_d+x_s}$, thus

$$\dot{f} = \left(\frac{x_d}{x_d + x_s}\right)' = \frac{\dot{x}_d x_s - x_d \dot{x}_s}{(x_d + x_s)^2} = \frac{x_d x_s (p_d - p_s)}{(x_d + x_s)^2} > 0. \quad (10)$$

Then, $\lim_{t\to\infty}(\frac{x_d}{x_d+x_s}) = 1$ and $x_s \to 0$.

The above mentioned results suggest that, when $\lambda > \alpha$ there is no such a closed set, in which the evolving state of each initial state which consist of these four strategies in this set also belongs to this set.

(2) When $\lambda < \alpha$ and according to the results in (1), there is no internal domain.

(3) When $\lambda = \alpha$ and thus $p_d = p_s$, the four-strategy system was reduced to the simplex $T = (C, D, L)$ or $T = (C, S, L)$. We will discuss this situation in the following.

Summing up the above dynamics, we can safely get the following conclusions: $\lambda = \alpha$ reduce the system to a three-strategy game, and $\lambda \neq \alpha$ will lead to the distinction of at least one strategy. ∎

A. Scenario 1: The Corners of the Simplex $T = (C, D, L)$

Theorem 2: If $r > 2 - 2\alpha$ holds, there exists a threshold value of x_l in the interval $(0, 1)$, above which $P_d - P_c < 0$.

Proof: Here, we employ the function $G(x_l) = (1 - x_l)(P_d - P_c)$ which has the same roots as $P_d - P_c$. For $x_l \in (0, 1)$,

$$\begin{aligned} G(x_l) &= (1 - x_l)(P_d - P_c) \\ &= (1 - \frac{r}{N} - \alpha) - (1 - \alpha)x_l + (r - 1 + \alpha)x_l^{N-1} \\ &\quad + (\frac{r}{N} + 1 - \alpha - r)x_l^N, \end{aligned} \quad (11)$$

$$\begin{aligned} G'(x_l) &= (\alpha - 1) + (N - 1)(r - 1 + \alpha)x_l^{N-2} \\ &\quad + N(\frac{r}{N} + 1 - \alpha - r)x_l^{N-1}. \end{aligned} \quad (12)$$

Note that $G(1) = G'(1) = 0$,

$$\begin{aligned} G''(1) &= (N - 1)(N - 2)(r - 1 + \alpha)x_l^{N-3} \\ &\quad + N(N - 1)(\frac{r}{N} + 1 - \alpha - r)x_l^{N-2}, \end{aligned} \quad (13)$$

$$G''(1) = (N - 1)(2 - 2\alpha - r). \quad (14)$$

We have

$$\begin{aligned} G(x_l) &\simeq G(1) + G'(1)(z - 1) + \frac{1}{2}G''(1)(z - 1)^2 \\ &= \frac{1}{2}(N - 1)(2 - 2\alpha - r)(1 - x_l)^2. \end{aligned} \quad (15)$$

For $r > 2 - 2\alpha$, $\lim_{x_l\to 1^-} G(x_l) < 0$,

$$\begin{aligned} G''(x_l) &= x_l^{N-3}(N - 1)[(N - 2)(r - 1 - \alpha) \\ &\quad + x_l(r + N - N\alpha - Nr). \end{aligned} \quad (16)$$

Since $G''(x_l)$ changes sign at most once in the interval $(0, 1)$, we claim that there exists a threshold value of x_l in the interval $(0, 1)$, above which $P_d - P_c < 0$.

From the above analysis, we get

$$\begin{cases} G(x_l) = (1 - x_l)(P_d - P_c) \\ G(0) = 1 - \frac{r}{N} - \alpha \\ G(1) = 0. \end{cases} \quad (17)$$

As illustrated in Fig. 1, the game dynamics takes on three qualitatively different cases, which will be discussed as follows.

Case 1.1 ($1 - r/N - \alpha > 0$, i.e., $G(0) > 0$):

$$\lim_{x_l\to 1^-} G(x_l) = \frac{1}{2}(N - 1)(2 - 2\alpha - r)(1 - x_l)^2. \quad (18)$$

When $r < 2 - 2\alpha$, $G(x_l) > 0$, $x_l \in (0, 1)$, the three corners represent a rock-scissors-paper type heteroclinic cycle, and there is no stable equilibrium of the game dynamics in this case.

Case 1.2 ($1 - r/N - \alpha > 0$, $r > 2 - 2\alpha$, $G(1^-) > 0$): the three corners represent a heteroclinic cycle. It is a center surrounded by closed orbits. Being similar to case 1.1, there is no stable equilibrium of the game dynamics in this case.

Case 1.3 ($1 - r/N - \alpha < 0$, i.e., $r > 2 - 2\alpha$): In this case, for all x_s, pure speculation (S) and pure defection (D) are both unstable equilibria of the game dynamics. The cooperation equilibrium (C) is stable and in fact a global attractor.

Summarizing the three cases in this scenario corresponding to the simplex $T = (C, D, L)$, we can conclude that the three corners represent a rock-scissors-paper type heteroclinic cycle if $1 - r/N - \alpha > 0$ (cases 1.1 and 1.2) while pure cooperation is a global attractor if $1 - r/N - \alpha < 0$ (case 1.3). ∎

Proposition 1: When $T = (C, D, L)$, under the replicator dynamics of (6.5), it holds that

if $1 - r/N - \alpha > 0$ and $r < 2 - 2\alpha$, there is no inner fixed point in T;

if $1 - r/N - \alpha > 0$ and $r > 2 - 2\alpha$, there is one inner fixed point in T;

if $1 - r/N - \alpha < 0$, full-C is only stable fixed point in T.

Proof: When $r > 2 - 2\alpha$, there exists a fixed point $x_l \in (0, 1)$ that $P_d = P_c$. Since we can get the only x_c and $x_d = 1 - x_l - x_c$, hence there is one inner fixed point in T. If $1 - r/N - \alpha > 0$ and $r < 2 - 2\alpha$, $P_d > P_c$ for all $x_l \in (0, 1)$, so there is no fixed point in T. If $1 - r/N - \alpha < 0$, we have $r > 2 - 2\alpha$, $(N > 2)$. Then it must be true that $P_c > P_d$, so full-C is only stable fixed point in T. ∎

B. Scenario 2: The Corners of the Simplex $T = (C, D, S)$

$$\begin{cases} P_d - P_c = 1 - \alpha - \frac{r}{N} \\ P_d - P_s = \lambda - \alpha \\ P_c - P_s = \lambda + \frac{r}{N} - 1. \end{cases} \quad (19)$$

Case 2.1 ($\lambda - \alpha > 0$, $1 - \alpha - r/N > 0$ and $1 - \lambda - r/N > 0$): Here, pure cooperation and pure speculation are both unstable equilibria of the game dynamics. Full defection equilibrium (D) is stable and in fact a global attractor.

Case 2.2 ($\lambda - \alpha > 0$, $1 - \alpha - r/N > 0$ and $1 - \lambda - r/N < 0$): In this case, pure cooperation and pure speculation are both unstable equilibria of the game dynamics. Pure defection equilibrium (D) is stable and a global attractor. The difference between case 2.1 and case 2.2 is that when there are only cooperators and speculators in the population, pure cooperation is the attractor in case 2.2 while pure speculation is the attractor in case 2.1.

Case 2.3 ($\lambda - \alpha > 0$, $1 - \alpha - r/N < 0$, and $1 - \lambda - r/N < 0$): Herein, pure defection and pure speculation are both unstable equilibria of the game dynamics. Pure cooperation is a stable and global attractor.

Case 2.4 ($\lambda - \alpha < 0$, $1 - \alpha - r/N > 0$, and $1 - \lambda - r/N > 0$): In this case, pure speculation is the only stable and global attractor.

Case 2.5 ($\lambda-\alpha < 0$, $1-\alpha-r/N < 0$, and $1-\lambda-r/N < 0$): Pure cooperation is thus the only stable and global attractor.

Case 2.6 ($\lambda-\alpha < 0$, $1-\alpha-r/N < 0$, and $1-\lambda-r/N > 0$): Pure speculation is the only stable and global attractor. The difference between case 2.6 and 2.4 is that when the population consists of only cooperators and defectors, pure cooperation is the attractor in case 2.6 while pure defection is the attractor in case 2.4.

Proposition 2: When $T = (C, D, S)$, under the adopted replicator dynamics, it holds that

if $\lambda - \alpha > 0$ and $1 - \alpha - r/N > 0$: full-D is only stable fixed point in T;

if $1 - \alpha - r/N < 0$ and $1 - \lambda - r/N < 0$: full-C is only stable fixed point in T;

if $\lambda - \alpha < 0$ and $1 - \lambda - r/N >$: full-S is only stable fixed point in T;

Proof: When $x_l = 0$, if $1 - \alpha - r/N > 0$, $P_d > P_c$; if $\lambda - \alpha > 0$, $P_d > P_s$, therefore if $x_d > 0$, $P_d > \bar{P}$. That means full-D ($x_d = 1$) is only stable fixed point in T. When $x_l = 0$, if $1 - \alpha - r/N <$, $P_c > P_d$; if $1 - \lambda - r/N < 0$, $P_c > P_s$, therefore if $x_c > 0$, $P_c > \bar{P}$. That means full-C ($x_c = 1$) is only stable fixed point in T. When $x_l = 0$, if $\lambda - \alpha < 0$, $P_s > P_d$; if $1 - \lambda - r/N > 0$, $P_s > P_c$, therefore if $x_s > 0$, $P_s > \bar{P}$. That means full-S ($x_s = 1$) is only stable fixed point in T. ■

C. Scenario 3: The Corners of the Simplex $T = (C, L, S)$

It is easily observed that $x_l = 0$ leads to $P_c - P_s = \lambda - 1 < 0$. Thus, the three corners represent a rock-scissors-paper type heteroclinic cycle. There is no stable equilibrium in this case.

Proposition 3: When $T = (C, S, L)$, under the adopted replicator dynamics, it holds that if $1 - r/N - \lambda > 0$ and $r < 2 - 2\lambda$, there is no inner fixed point in T; if $1 - r/N - \lambda > 0$ and $r > 2 - 2\lambda$, there is one inner fixed point in T; if $1 - r/N - \lambda < 0$, full C is only stable fixed point in T.

Proof: By using λ takes the place of α, we can get the similar results with proposition 1.3. ■

D. Scenario 4: The Corners of the Simplex $T = (D, L, S)$

Case 4.1 ($\lambda - \alpha < 0$): In this case, pure loners is the only stable and in fact the only global attractor.

Case 4.2 ($\lambda - \alpha > 0$): Still, pure loners remains the only stable and in fact the only global attractor. The difference between case 4.1 and 4.2 is that when there are only speculators and defectors in the population, pure speculation is the attractor in case 4.1 while pure defection is the attractor in case 4.2.

Summarizing the two cases in scenario 4 corresponding to the simplex $T = (C, D, S)$, we can conclude that pure-L is the only global attractor in the system.

Proposition 4: When $T = (S, D, L)$, under the replicator dynamics of (6.5), it holds that full-L is only stable fixed point in T.

Proof: When $x_c = 0$, $P_l - P_d = (\alpha + \sigma)(1 - N_l^{N-1}) > 0$ and $P_l - P_s = (\lambda + \sigma)(1 - N_l^{N-1}) > 0$, therefore full-$L$ ($x_l = 1$) is only stable fixed point in T. ■

IV. Conclusion

How to effectively coordinate the cooperation between agents with conflicts of interest is a hot topic, and its solutions can be applied to a wide range of applications. For such a biology-inspired topic, only when individual heterogeneity and diversity are taken into account in theoretical modeling can the core of the problem be better addressed. In the face of possible punishment and loss of benefits, the individual's strategy choices show diversity. Here, we extend the theoretical analysis to a model in which four strategies coexist, and they are respectively derived from actual behaviors in real world. A theoretical explanation about the evolutionary fate of the system is provided. An interesting future direction would be to address whether the presence of more strategy options altogether affect the dynamics of behaviors in multi-agent systems.

References

[1] P. Ramazi, J. Riehl, and M. Cao, "Networks of conforming or nonconforming individuals tend to reach satisfactory decisions," *Proc. Nat. Acad. Sci. USA*, vol. 113, no. 46, pp. 12985–12990, 2016.

[2] M. Long, H. Su, and B. Liu, "Second-order controllability of two-time-scale multi-agent systems," *Appl. Math. Comput.*, vol. 343, pp. 299–313, Feb. 2019.

[3] H. Su, H. Wu, X. Chen, and M. Z. Chen, "Positive edge consensus of complex networks," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 48, no. 12, pp. 2242–2250, Dec. 2018.

[4] L. Böttcher, J. Nagler, and H. J. Herrmann, "Critical behaviors in contagion dynamics," *Phys. Rev. Lett.*, vol. 118, no. 8, 2017, Art. no. 088301.

[5] P. Ramazi and M. Cao, "Asynchronous decision-making dynamics under best-response update rule in finite heterogeneous populations," *IEEE Trans. Autom. Control*, vol. 63, no. 3, pp. 742–751, Mar. 2018.

[6] E. Fehr and S. Gächter, "Altruistic punishment in humans," *Nature*, vol. 415, pp. 137–140, Jan. 2002.

[7] H. Brandt, C. Hauert, and K. Sigmund, "Punishing and abstaining for public goods," *Proc. Nat. Acad. Sci. USA*, vol. 103, no. 2, pp. 495–497, 2006.

[8] J. Zhang, Y. Zhu, and Z. Chen, "Evolutionary game dynamics of multiagent systems on multiple community networks," *IEEE Trans. Syst., Man, Cybern., Syst.*, to be published.

[9] J. Riehl, P. Ramazi, and M. Cao, "A survey on the analysis and control of evolutionary matrix games," *Annu. Rev. Control*, vol. 45, pp. 87–106, 2018. AQ2

[10] L. A. Imhof and M. A. Nowak, "Stochastic evolutionary dynamics of direct reciprocity," *Proc. R. Soc. London B*, vol. 277, no. 1680, pp. 463–468, 2010.

[11] M. A. Nowak, "Five rules for the evolution of cooperation," *Science*, vol. 314, no. 5805, pp. 1560–1563, 2006.

[12] H. Ohtsuki and M. A. Nowak, "Direct reciprocity on graphs," *J. Theor. Biol.*, vol. 247, no. 3, pp. 462–470, 2007.

[13] J. M. Pacheco, A. Traulsen, H. Ohtsuki, and M. A. Nowak, "Repeated games and direct reciprocity under active linking," *J. Theor. Biol.*, vol. 250, no. 4, pp. 723–731, 2008.

[14] M. A. Nowak and K. Sigmund, "Evolution of indirect reciprocity," *Nature*, vol. 437, no. 7063, pp. 1291–1298, 2005.

[15] U. Berger, "Learning to cooperate via indirect reciprocity," *Games Econ. Behav.*, vol. 72, no. 1, pp. 30–37, 2011.

[16] M. Wubs, R. Bshary, and L. Lehmann, "Coevolution between positive reciprocity, punishment, and partner switching in repeated interactions," *Proc. Roy. Soc. London B*, vol. 283, no. 1832, 2016, Art. no. 20160488.

[17] J. Henrich *et al.*, "Costly punishment across human societies," *Science*, vol. 312, no. 5781, pp. 1767–1770, 2006.

[18] C. Hauert and O. Stenull, "Simple adaptive strategy wins the prisoner's dilemma," *J. Theor. Biol.*, vol. 218, no. 3, pp. 261–272, 2002.

[19] W.-B. Du, W. Ying, G. Yan, Y.-B. Zhu, and X.-B. Cao, "Heterogeneous strategy particle swarm optimization," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 64, no. 4, pp. 467–471, Apr. 2017.

[20] J. Zhan and X. Li, "Cluster consensus in networks of agents with weighted cooperative—Competitive interactions," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 65, no. 2, pp. 241–245, Feb. 2018.

[21] L. Balafoutas, N. Nikiforakis, and B. Rockenbach, "Altruistic punishment does not increase with the severity of norm violations in the field," *Nat. Commun.*, vol. 7, Nov. 2016, Art. no. 13327.

[22] K. Panchanathan and R. Boyd, "Indirect reciprocity can stabilize cooperation without the second-order free rider problem," *Nature*, vol. 432, no. 7016, pp. 499–502, 2004.

[23] J. Zhang, T. Chu, and F. J. Weissing, "Does insurance against punishment undermine cooperation in the evolution of public goods games?" *J. Theor. Biol.*, vol. 321, pp. 78–82, Mar. 2013.