ISSN 1674-1056
http://cpb.iphy.ac.cn
http://iopscience.iop.org/cpb
中国物理 B
Chinese Physics
中国物理
Chinese Physics B

Stochastic evolutionary public goods game with first and second order costly punishments in finite populations

Ji Quan(全吉), Yu-Qing Chu(储育青), Wei Liu(刘伟), Xian-Jia Wang(王先甲), Xiu-Kang Yang(阳修康)
Citation: Chin. Phys. B . 2018, 27(6): 060203. **doi:** 10.1088/1674-1056/27/6/060203

Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

On fairness, full cooperation, and quantum game with incomplete information

Zhen-Zhou Lei(雷振州), Bo-Yang Liu(刘博阳), Ying Yi(易英), Hong-Yi Dai(戴宏毅), Ming Zhang(张明)
Chin. Phys. B . 2018, 27(3): 030202. **doi:** 10.1088/1674-1056/27/3/030202

Improving the secrecy rate by turning foes to allies: An auction scheme

Ma Ya-Yan, Wang Bao-Yun
Chin. Phys. B . 2015, 24(9): 090209. **doi:** 10.1088/1674-1056/24/9/090209

Spatial snowdrift game in heterogeneous agent systems with co-evolutionary strategies and updating rules

Xia Hai-Jiang, Li Ping-Ping, Ke Jian-Hong, Lin Zhen-Quan
Chin. Phys. B . 2015, 24(4): 040203. **doi:** 10.1088/1674-1056/24/4/040203

Virus spreading in wireless sensor networks with a medium access control mechanism

Wang Ya-Qi, Yang Xiao-Yuan
Chin. Phys. B . 2013, 22(4): 040206. **doi:** 10.1088/1674-1056/22/4/040206

Word diversity can accelerate consensus in naming game

Zhang Xiao-Dan,Wang Zhen,Zheng Fei-Fei,Yang Miao
Chin. Phys. B . 2012, 21(3): 030205. **doi:** 10.1088/1674-1056/21/3/030205

Stochastic evolutionary public goods game with first and second order costly punishments in finite populations*

Ji Quan(全吉)[1,†], Yu-Qing Chu(储育青)[2,‡], Wei Liu(刘伟)[3], Xian-Jia Wang(王先甲)[4], and Xiu-Kang Yang(阳修康)[1]

[1] School of Management, Wuhan University of Technology, Wuhan 430070, China
[2] School of Science, Wuhan University of Technology, Wuhan 430070, China
[3] School of Resources and Environmental Engineering, Wuhan University of Technology, Wuhan 430070, China
[4] School of Economics and Management, Wuhan University, Wuhan 430072, China

(Received 28 November 2017; revised manuscript received 10 February 2018; published online 10 May 2018)

We study the stochastic evolutionary public goods game with punishment in a finite size population. Two kinds of costly punishments are considered, i.e., first-order punishment in which only the defectors are punished, and second-order punishment in which both the defectors and the cooperators who do not punish the defective behaviors are punished. We focus on the stochastic stable equilibrium of the system. In the population, the evolutionary process of strategies is described as a finite state Markov process. The evolutionary equilibrium of the system and its stochastic stability are analyzed by the limit distribution of the Markov process. By numerical experiments, our findings are as follows. (i) The first-order costly punishment can change the evolutionary dynamics and equilibrium of the public goods game, and it can promote cooperation only when both the intensity of punishment and the return on investment parameters are large enough. (ii) Under the first-order punishment, the further imposition of the second-order punishment cannot change the evolutionary dynamics of the system dramatically, but can only change the probability of the system to select the equilibrium points in the "C+P" states, which refer to the co-existence states of cooperation and punishment. The second-order punishment has limited roles in promoting cooperation, except for some critical combinations of parameters. (iii) When the system chooses "C+P" states with probability one, the increase of the punishment probability under second-order punishment will further increase the proportion of the "P" strategy in the "C+P" states.

Keywords: public goods games, stochastic stable equilibrium, punishment, finite population

PACS: 02.50.Le, 02.50.Ga, 05.45.Pq, 05.65.+b **DOI:** 10.1088/1674-1056/27/6/060203

1. Introduction

Social dilemma problems are prevalent in the real world. They essentially depict the conflict between individual rationality and collective rationality.[1] The public goods game as a classic model is used to describe the dilemma between multiplayers' interactions. The "free riding" phenomenon it portrayed has received widespread attention in economics,[2–6] psychology,[7–10] evolutionary biology,[11–13] complexity science,[14–18] and other disciplines of scholars. In the standard public goods game, participants have two choices of cooperation and defection. Cooperators participate in the investment, and defectors do not participate in the investment. The total investment of all cooperators in the public goods game can be amplified, and then distributed equally among all participants. If each individual chooses to cooperate, then everyone can get the maximized benefits, in which situation the system reaches the Pareto optimal state. However, due to the non-exclusive and non-competitive nature of the public goods, defectors can get the benefits of the cooperators' contribution by free riding, resulting in a higher net income than the cooperators. From the perspective of individual rationality, free riding is an optimal strategy for each individual. From the evolutionary point of view, regardless of the cooperator's proportion in the initial population, the cooperation strategy will eventually be replaced by the defection of the free riders. The public goods game cannot form highly efficient social cooperation.

In order to improve the efficiency and promote the evolution of cooperation, scholars have proposed some mechanisms to reduce the free riding behavior, such as the reward and punishment mechanism,[13,19–25] reputation mechanism,[12,26–31] optional participation mechanism,[18,25,32–36] spatial interaction mechanism,[14,37–42] etc.[43–45] We focus on the punishment mechanism in this paper. Punishment, as an efficient way to change human behavior, is regarded as an important mechanism to maintain cooperation in our human society. Through behavioral experiments, experimental economists have found that punishment and reward can significantly enhance the level of cooperation in the public goods game.[46] Different forms of punishment have also attracted the broad interest of schol-

*Project supported by the National Natural Science Foundation of China (Grant Nos. 71501149 and 71231007), the Soft Science Project of Hubei Province, China (Grant No. 2017ADC122), and the Fundamental Research Funds for the Central Universities, China (Grant No. WUT: 2017VI070).

†Corresponding author. E-mail: quanji123@163.com
‡Corresponding author. E-mail: chuyuqing@whut.edu.cn

© 2018 Chinese Physical Society and IOP Publishing Ltd

ars. Based on different parties to carry out the punishment, punishment can be divided into three cases: first-party punishment, second-party punishment, and third-party punishment. First-party punishment refers to uncomfortable feelings when an individual violates the group's norms or defection. It is an inner self-conscience punishment, such as guilt, shame or embarrassment.[47] Second-party punishment refers to the punishment imposed by others involved in the game on the free riders.[21,22] Third-party punishment refers to the punishment imposed by spectators on the free riders.[23,48] The spectators here can be a completely independent external group, such as the judicial system, or a potential group that the defectors may meet in the future. The reputation mechanism can be seen as a special case of third-party punishment. In this paper, we study the second-party punishment with cost in the public goods game. It takes punishment as a participant strategy. Participants can choose this strategy to reduce the benefits of some other types of strategies, which will bring additional costs to themselves. This study considers two forms of second-party punishments: punishing only the defection strategy (first-order free ride behavior); punishing not only the defectors, but also the cooperators who do not punish the defective behaviors (second-order free ride behavior). These two forms of punishments are called first-order punishment and second-order punishment respectively in the following.

Evolutionary games which can effectively describe the evolutionary process of strategies have been widely used to study the evolution of cooperation in social dilemmas. In evolutionary games, only bounded rationality is needed for the assumption of participants. The model is based on groups and describes the updating process of individuals' strategies according to their incomes. For the public goods game, in Ref. [12] the replicator dynamics was used to study the influence of reward on the cooperative behavior of the population, and compared with the punishment mechanism. The replicator dynamics is a deterministic evolution model, in which it is assumed that the number of individuals in the population is infinite. Thus, differential equations are used to describe the evolutionary process of different strategies in the system. This model analyzes the evolutionary stable states of the system by analyzing the stability of the differential equations at the equilibrium point. The stable equilibrium point is the evolutionary stable strategy (ESS) of the game. In order to describe a finite size population and the randomness of the system, in Ref. [49] the public goods game was studied based on the Moran process that is widely used in ecology. They investigated the influence of reward on the cooperative behavior of the finite size population. In the Moran-based model the fixation probability was used to analyze the evolutionary dynamics of the system for only two strategies. This analysis is conducted necessarily under the assumption of the weak selection generally.

In order to describe the continuous noise effects caused by mutation in the evolutionary process of the system, Young and Foster[50,51] introduced the concept of stochastic stable equilibrium (SSE) for the first time. This concept can better describe the evolutionary stability of strategy in a stochastic environment. However, since the concept of SSE was proposed, only a few articles have been devoted to the SSE of the games. As far as we know, recently, Quan *et al.*[52,53] studied the SSE of the evolutionary games in a finite size population based on the Markov process. Other models are basically based on the stochastic differential equations for analysis under the assumption of infinite size populations as indicated in the research by Huang *et al.*[54] and Liang *et al.*[55] Unlike the above-mentioned evolution model that aims at the ESS of the game, we study the SSE of the public goods game with punishment in a finite size of population. We adopt the stochastic evolution model proposed in Ref. [56], where the evolutionary process of strategies in the population is described as a finite state multidimensional Markov process. The stochastic stabilities of the system under different punishment parameters are analyzed by the limit distribution of the Markov process. By analyzing the SSE of this stochastic system, the influence of the punishment mechanism on the cooperation behavior of the population is revealed.

The remainder of the paper is organized as follows. In Section 2, we introduce the stochastic evolutionary game of the public goods game with punishment, including the strategy expected payoffs in a finite population, the evolutionary dynamics, and the SSE of the stochastic evolutionary system. In Section 3, the stochastic stability and equilibrium results of the system are discussed under different punishment parameters, and then the influence of these parameters on the cooperative behavior of the population is analyzed. Finally, some conclusions are drawn in Section 4.

2. Model

2.1. Public goods with costly punishment

Two forms of punishing strategies are introduced into the classic public goods game, which can be called the first-order and the second-order punishment respectively. The first-order punishing strategies participate in the investment and only punish the first-order free riders (defection strategies). The second-order punishing strategies participate in the investment, and punish not only the first-order free riders but also the second-order free riders (the sole cooperators who does not punish the defectors). Let α_1 and α_2 $(\alpha_1 > \alpha_2)$ denote the probabilities of the punishing strategies to punish the defection and cooperation strategies respectively. We assume that the processes of punishing the defectors and the sole cooperators are separate. Therefore, we can consider these two random events independently. The punishment is costly. Let γ

denote the unit cost of punishment for the punisher, and β refer to the corresponding unit penalty for the individual who is to be punished ($\gamma < \beta$).

Suppose that the population contains three types of strategies, that is, cooperation, defection, and punishment strategies. Each time N individuals from the population are randomly selected to participate in the public goods game. Let c denote the cost of investment, and r the return on investment ($1 < r < N$) of the public goods. Each individual chooses the corresponding strategy according to its type. In a sample, when the numbers of cooperation, defection, and punishment type individuals are $n_{\rm C}$, $n_{\rm D}$, and $n_{\rm P}$ respectively, then the cooperation-type individual's payoff is $[rc(n_{\rm C}+n_{\rm P})]/N - c - \alpha_2\beta n_{\rm P}$, the defection-type individual's payoff is $[rc(n_{\rm C}+n_{\rm P})]/N - \alpha_1\beta n_{\rm P}$, and the punishment-type individual's payoff is $[rc(n_{\rm C}+n_{\rm P})]/N - c - \alpha_1\gamma n_{\rm D} - \alpha_2\gamma n_{\rm C}$.

2.2. Strategy expected payoff in a finite population

Assuming that a finite population size is M, the three types of strategies for cooperation, defection, and punishment are well mixed. Each time N individuals from the population are randomly selected to participate in the public goods game. Obviously, $M \geq N$. As in the other literature on the finite size population, such as in Refs. [32] and [49], here in our work we do not discuss the trivial case of $M = N$. Let i, j, k ($i + j + k = M$) denote the numbers of the three types of individuals respectively. In the following, we analyze the expected payoff for each type of strategy in this finite size population.

We do not consider the penalty items in our analysis at first. In this case, the cooperation and punishment strategies are the same, including their payoffs. For a defector, when it encounters another $N-1$ individual in the population, the probability with which there are m cooperators or punishers, and the rest are $N-1-m$ defectors, is

$$p(n_{\rm C}+n_{\rm P}=m) = \frac{\binom{i+k}{m}\binom{j-1}{N-1-m}}{\binom{M-1}{N-1}}. \tag{1}$$

The expected payoff for this defection-type individual, which does not consider the penalty items, is

$$\begin{aligned}\sum_{m=0}^{N-1}\frac{rcm}{N}p(n_{\rm C}+n_{\rm P}=m) &= \sum_{m=0}^{N-1}\frac{rc}{N}m\frac{\binom{i+k}{m}\binom{j-1}{N-1-m}}{\binom{M-1}{N-1}}\\ &= \frac{rc}{N}(N-1)\frac{i+k}{M-1}.\end{aligned} \tag{2}$$

The expected payoff for a cooperation- or punishment-type individual, which does not consider the penalty items, is

$$\sum_{m=1}^{N}\left(\frac{rcm}{N}-c\right)p(n_{\rm C}+n_{\rm P}=m-1)$$

$$= \frac{rc}{N}(N-1)\frac{i+k}{M-1}+\frac{rc}{N}-c. \tag{3}$$

Taking into account the corresponding penalty or punishing cost for each type of strategy, the above-mentioned expected payoff can be corrected as follows.

For a defection-type individual, the total penalty brought by the punishment-type strategies is $k/(M-1)(N-1)\alpha_1\beta$, where α_1 is the probability with which a defector is punished by a punishment-type strategy, β is the penalty intensity, and $k/(M-1)(N-1)$ is the expected number of punishment-type individuals in a sample.

For a cooperation-type individual, the total penalty brought by the punishment-type strategies is $k/(M-1)(N-1)\Psi(j)\alpha_2\beta$, where α_2 is the probability with which a cooperator (second-order free rider) is punished by a punishment-type strategy:

$$\Psi(j) = 1 - \frac{\binom{M-j-2}{N-2}}{\binom{M-2}{N-2}},$$

$1-\Psi(j)$ is the probability with which there is no defector but there are cooperators and punishers in a sample. In this case, there is no reason for punishing the cooperators (because there is no defector in the game).

For a punishment-type individual, the total punishing cost is $j/(M-1)(N-1)\alpha_1\gamma + i/(M-1)(N-1)\Psi(j)\alpha_2\gamma$, the former and latter items represent the costs of punishing the defectors and the cooperators respectively.

Thus, when the size of the population is M, and the numbers of cooperators, defectors, and punishers in the group are i, j, k, the expected payoffs of the three types of strategies are respectively

$$\pi_{\rm C}^{(i,j,k)} = B - \varphi c - \frac{k}{M-1}(N-1)\Psi(j)\alpha_2\beta, \quad (i \neq 0), \tag{4}$$

$$\pi_{\rm D}^{(i,j,k)} = B - \frac{k}{M-1}(N-1)\alpha_1\beta, \quad (j \neq 0), \tag{5}$$

$$\begin{aligned}\pi_{\rm P}^{(i,j,k)} &= B - \varphi c - \frac{j}{M-1}(N-1)\alpha_1\gamma\\ &\quad - \frac{i}{M-1}(N-1)\Psi(j)\alpha_2\gamma, \quad (k \neq 0),\end{aligned} \tag{6}$$

where

$$B = \frac{rc}{N}(N-1)\frac{i+k}{M-1}, \quad \varphi = 1 - \frac{r}{N},$$

$$\Psi(j) = 1 - \frac{\binom{M-j-2}{N-2}}{\binom{M-2}{N-2}}.$$

2.3. Evolutionary dynamics in discrete system

The number of each type of strategies will evolve with the quantity of their payoffs. In order to describe the evolutionary process, we introduce a stochastic process $z(t)$.

Let $z(t) = (z_1(t),\ z_2(t),\ M - z_1(t) - z_2(t))$ denote the number of cooperation-, defection- and punishment-type strategies in the population at time t, and define $z(t)$ as the system state. For convenience, we abbreviate it as $(z_1(t), z_2(t))$. The state space of the system is $S = \{(i,j)|0 \le i+j \le M; i,j \in \mathbb{N}\}$, and the number of elements in the state space is $|S| = (M+1)(M+2)/2$. Each time, individuals in the population adjust their strategies according to their expected payoffs. The strategy adjustment leads to the change of the system state. The three assumptions: inertia, myopic, and mutation in Ref. [56] about bounded rationality of individuals in the population are used in our model. Due to inertia, we can assume that it is impossible to have more than two individuals adjust their strategies simultaneously once. Myopic refers to the individual when choosing its strategy: it will only consider the current payoff, regardless of the payoff in the future. Mutation refers to the possibility with which individuals may choose a non-optimal strategy with a small probability because of the complex decision-making environment and the limited nature of individual cognitive capability.

According to the above assumptions, when the system state is $(z_1(t), z_2(t)) = (i,j) \in S$, the transfer rate of the strategy x towards strategy y can be described as

$$p_{x\to y}^{(i,j)} = \varepsilon + \kappa \cdot (\pi_y^{(i,j)} - \pi_x^{(i,j)})^+,\ \ x,y \in \{C,D,P\},\ \ x \ne y, \quad (7)$$

where

$$f^+ = \begin{cases} f, & f > 0, \\ 0, & f \le 0, \end{cases}$$

$\varepsilon > 0$ is a small positive number, and $\kappa > 0$. For example, when $\pi_x^{(i,j)} > \pi_y^{(i,j)}$, the individual in the (i,j) state has more incentive to move from strategy y to strategy x, but because of the mutation, the transfer rate of strategy x to strategy y is $p_{x\to y}^{(i,j)} = \varepsilon$. Thus, the parameter ε can be seen as the noise intensity in the environment, and κ can be understood as the speed at which the individual responds to the environment.

When i takes zero, $\pi_{\rm C}^{(0,j,k)}$ is meaningless; when j takes zero, $\pi_{\rm D}^{(i,0,k)}$ is meaningless; when k takes zero, $\pi_{\rm P}^{(i,j,0)}$ makes no sense. At this point, the payoffs of cooperation-, defection- and punishment-type strategies are defined as the average payoff of the population, i.e.,

$$\pi_{\rm C}^{(0,j,k)} = \frac{M-k}{M}\pi_{\rm D}^{(0,j,k)} + \frac{k}{M}\pi_{\rm P}^{(0,j,k)}, \quad (8)$$

$$\pi_{\rm D}^{(i,0,k)} = \frac{M-k}{M}\pi_{\rm C}^{(i,0,k)} + \frac{k}{M}\pi_{\rm P}^{(i,0,k)}, \quad (9)$$

$$\pi_{\rm P}^{(i,j,0)} = \frac{i}{M}\pi_{\rm C}^{(i,j,0)} + \frac{M-i}{M}\pi_{\rm D}^{(i,j,0)}. \quad (10)$$

Let $I = (i,j)$, then $I' = (i',j')$. Due to homogeneity, $p_{I,I'}(t)$ denotes the probability with which the system transfers from state I to state I' after time t, that is,

$$p_{I,I'}(t) = p\{z(s+t) = (i',j')|z(s) = (i,j)\}, \quad \forall s > 0. \quad (11)$$

According to the system evolutionary rules and the transition rate between different strategies, after a small enough time t, the probabilities with which the system transfers from state $I = (i,j)$ to states $(i-1,j)$ and $(i-1,j+1)$ are $p_{\rm C\to P}^{(i,j)}t + o(t)$ and $p_{\rm C\to D}^{(i,j)}t + o(t)$ respectively; the probabilities with which the system transfers from state $I = (i,j)$ to states $(i,j-1)$ and $(i+1,j-1)$ are $p_{\rm D\to P}^{(i,j)}t + o(t)$ and $p_{\rm D\to C}^{(i,j)}t + o(t)$ respectively; the probabilities with which the system transfers from state $I = (i,j)$ to states $(i+1,j)$ and $(i,j+1)$ are $p_{\rm P\to C}^{(i,j)}t + o(t)$ and $p_{\rm P\to D}^{(i,j)}t + o(t)$ respectively; the probability of keeping the same state is $1 - \sum_{\substack{x,y\in\{C,D,P\}\\ x\ne y}} p_{x\to y}^{(i,j)}t - o(t)$; where $o(t)$ is a high-order infinitesimal of t. Figure 1 shows all possible state transfer processes of the system.

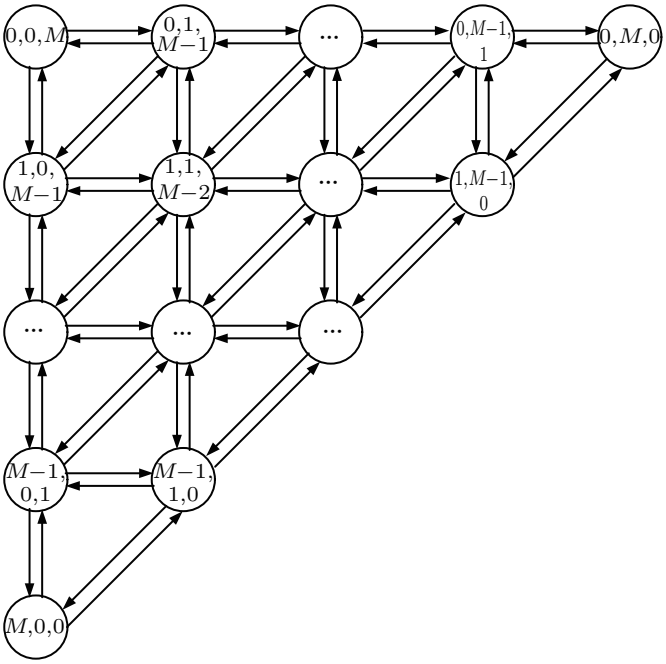


Fig. 1. Possible state transfer processes of the system.

2.4. Stochastic stable equilibrium

As $\varepsilon > 0$, this process is ergodic. According to the properties of the stochastic process, when $t \to +\infty$, the limit of $p_{I,I'}(t)$ exists and it is independent of the initial state I. Let

$$\lim_{t\to+\infty} p_{I,I'}(t) = v_{I'}^{\varepsilon}, \quad (12)$$

and $v_{I'}^{\varepsilon}$ will be the limit distribution of the two-dimensional Markov process reaching arbitrary state $I'(I' \in S)$ when the system noise is ε. According to the limit distribution, it is possible to determine the evolutionary stable state of the system under arbitrary noise intensity. Further, when the noise parameter gradually reduces to zero, let

$$\lim_{\varepsilon\to 0^+} v_{I'}^{\varepsilon} = v_{I'}. \quad (13)$$

According to $v_{I'}$, we can determine the limit states of the system and their corresponding limit probabilities. According to

Young's description in Ref. [51], state $I' \in S$ is stochastically stable if and only if $v_{I'} > 0$.

3. Results and discussion

The Gauss–Seidel iterative algorithm introduced by Stewart in his monograph[57] can be used to calculate the limit distribution of the above Markov process. We show the effects of game parameters on the stochastic stability equilibrium of the system in the following. By vast numerical calculations, we find that when ε gradually decreased to zero, the system has limit distributions of more than zero only in $(0,M,0)$, $(i,0,M-i)$ $(i=0,1,2,\ldots,M)$ states. Thus, only the states of $(0,M,0)$, $(i,0,M-i)$ $(i=0,1,2,\ldots,M)$ may be the stochastic stable states of the evolutionary system. Among them, state $(0,M,0)$ indicates that all the individuals choose the defection strategy, denoting it as the "All D" state. States $(i,0,M-i)$ $(i=0,1,2,\ldots,M)$ indicate the co-existence of the cooperation and punishment strategies, or all individuals choosing the punishment strategy, or all individuals choosing the cooperation strategy, we denote them as the "C+P" states. In the following, we fix parameters $M=20$, $N=5$, $c=1$, $\kappa=1$, $\alpha_1=1$ $\gamma=0.2$, and study the influences of β, r, α_2 on the probabilities of the system to choose different stable equilibria.

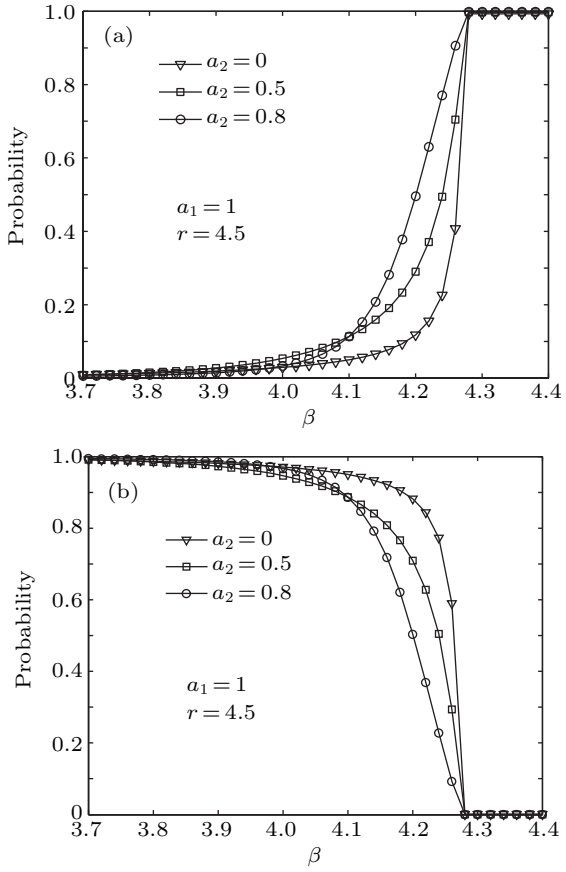


Fig. 2. : Variations of limit distribution with β of two states for $\alpha_2=0$, 0.5, 0.8 with fixing $\alpha_1=1$ and $r=4.5$. (a) The limit distribution of "C+P" states, (b) the limit distribution of "All D" state.

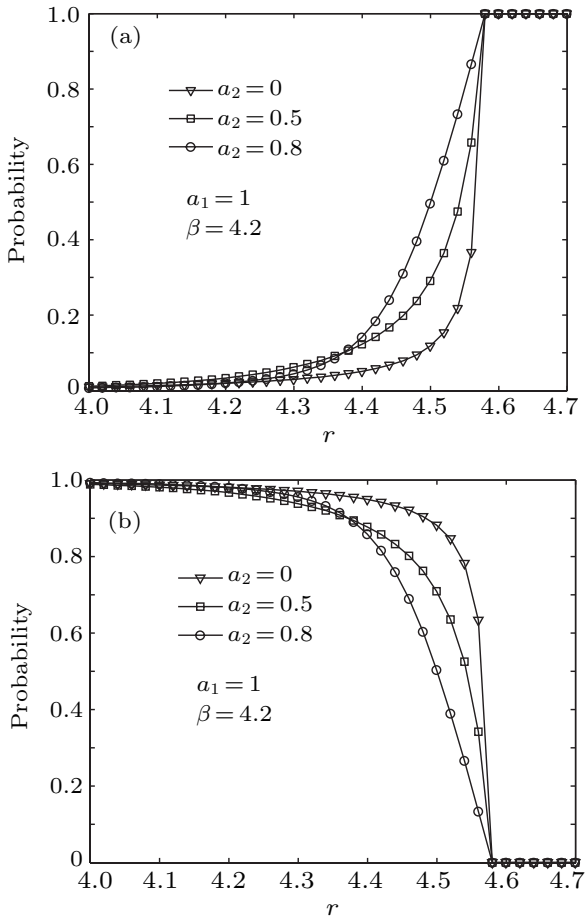


Fig. 3. Variations of probability with r of the two states for $\alpha_2=0$, 0.5, 0.8 with fixing $\alpha_1=1$ and $\beta=4.2$. (a) The limit distribution of "C+P" states, (b) the limit distribution of "All D" state.

By numerical calculation, we find that for fixed $r=4.5$, and for different values of $\alpha_2=0,0.5,0.8$, there are common critical values of $\underline{\beta}=3.78$ and $\bar{\beta}=4.28$. When $\beta\le\underline{\beta}$, the system selects the "All D" state with a probability close to one. When $\beta\ge\bar{\beta}$, the system selects "C+P" states with a probability close to one. When $\underline{\beta}<\beta<\bar{\beta}$, the system chooses the "All D" state and "C+P" states with different probabilities; as β increases, the probability of choosing the "All D" state gradually decreases, and the probability of choosing "C+P" states gradually increases; when β increases to a critical value $\tilde{\beta}$, the probability of the system choosing one of the two states rapidly decreases from nearly one to zero (or increases from nearly zero to one). The smaller the value of α_2 and the greater the critical value of $\tilde{\beta}$, the steeper the curve of the probability reduction (or increase) is. Figure 2 shows the relationship between the limit distributions of the system choosing the two stable states and β under different parameters of α_2. For fixed $\beta=4.2$, and for different values of $\alpha_2=0,0.5,0.8$, there are common values of $\underline{r}=4.06$ and $\bar{r}=4.58$. When $r\le\underline{r}$, the system selects the "All D" state with a probability close to one. When $r\ge\bar{r}$, the system selects "C+P" states with a probabil-

ity close to one. When $\underline{r} < r < \bar{r}$, the system chooses the "All D" state and "C+P" states with different probabilities; as r increases, the probability of selecting the "All D" state gradually decreases, and the probability of selecting "C+P" states gradually increases, when r increases to a certain critical value $\tilde{r}$, the probability of the system choosing one of the two states rapidly decreases from nearly one to zero (or from nearly zero to one). The smaller the value of α_2 and the greater the critical value of $\tilde{r}$, the steeper the curve of the probability reduction (or increase) is. Figure 3 shows the relationship between the limit distributions of the system choosing the two stable states and r under different values of α_2 .

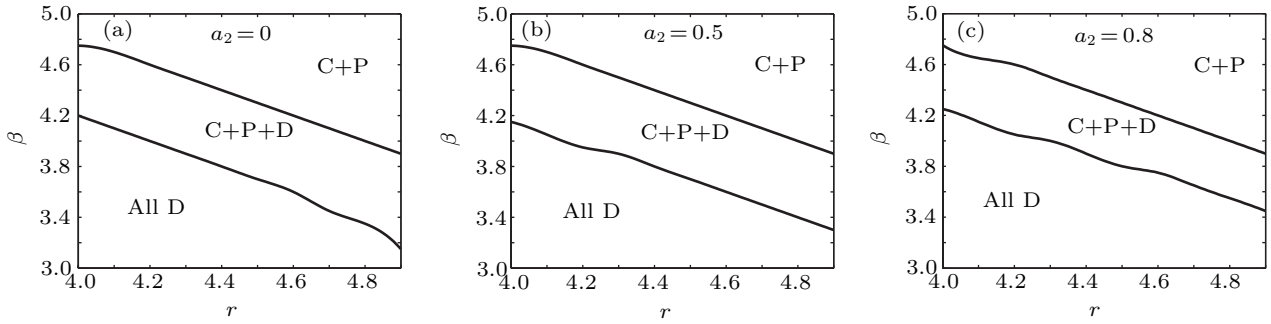


Fig. 4. Ranges of (β, r) for the system reaching the three different kinds of stable states for fixed $\alpha_1 = 1$ when $\alpha_2 = 0, 0.5, 0.8$ respectively.

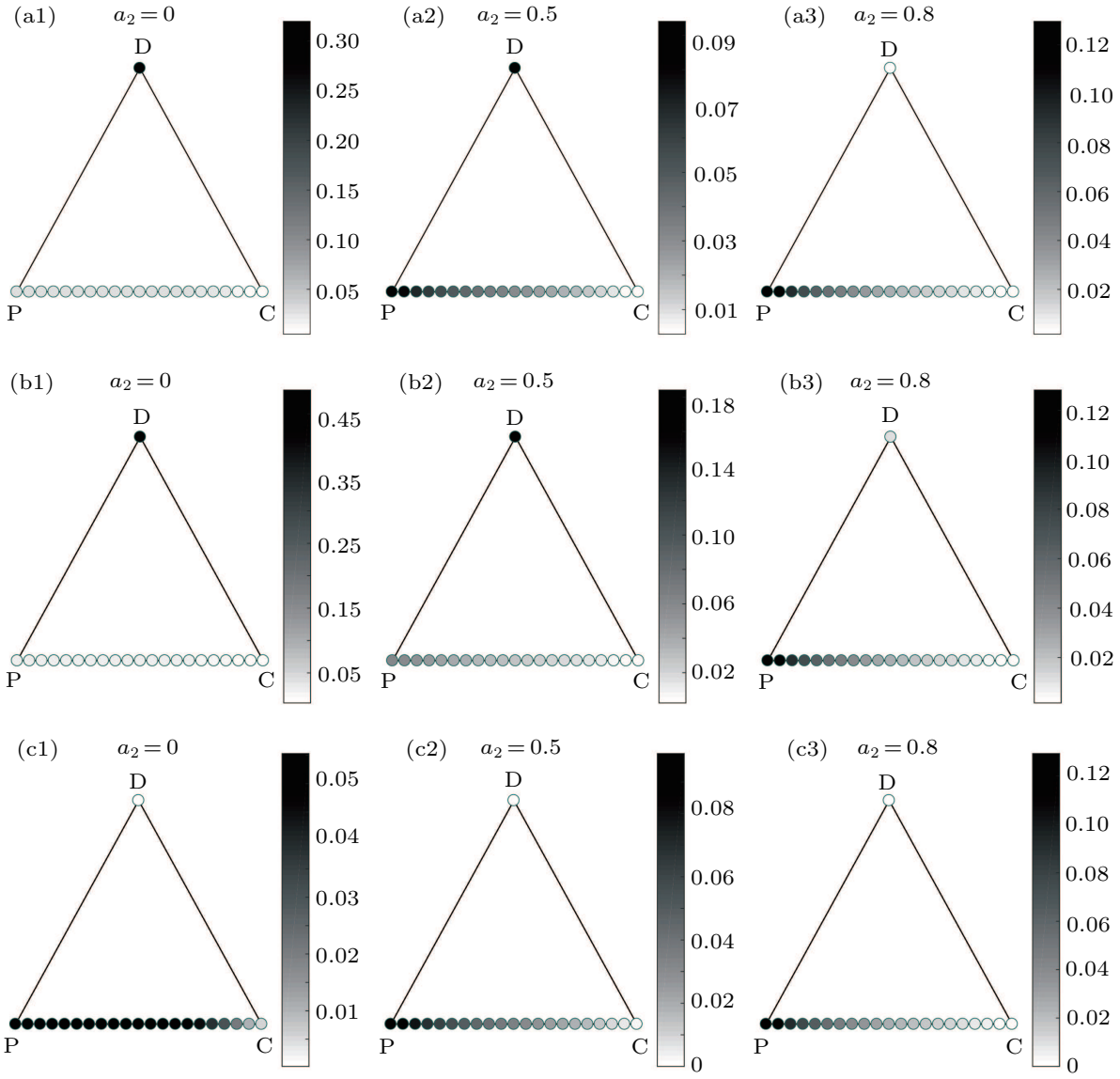


Fig. 5. Limit distributions of the stable states with $\alpha_2 = 0, 0.5, 0.8$ for fixed $\alpha_1 = 1$ and three different combinations of (β, r): (a) $\beta = 4.65$, $r = 4.1$; (b) $\beta = 4.55$, $r = 4.2$; (c) $\beta = 4.5$, $r = 4.5$.

Figure 4 shows the corresponding ranges of parameters (β, r) for the system reaching the three different kinds of stable states of "C+P", "All D", "C+P+D" for fixed $\alpha_1 = 1$ when $\alpha_2 = 0$, 0.5, 0.8 respectively. Comparing the three sub-figures in Fig. 4, we can see that under the first-order punishment, the further imposition of the second-order punishment cannot change the evolutionary dynamics of the system dramatically. In most cases, α_2 has a limited effect on the evolution of cooperation in the population. We choose two special combinations of (β, r), in which case the change of α_2 can have a great influence on the system to choose different stable equilibria. Figures 5(a) and 5(b) show the results of $r = 4.1$, $\beta = 4.65$, and $r = 4.2$, $\beta = 4.55$ for fixed $\alpha_1 = 1$ in the simplex S3 respectively. From the figures, we can see that there are critical combinations of (β, r). For these combinations, the value of α_2 can change the stochastic stable equilibrium of the system.

Figure 5(c) shows the effects of different values of α_2 on the limit probabilities with which the system chooses different states $(i, 0, 20-i)$ $(i = 0, 1, 2, \ldots, 20)$, when the system reaches the "C+P" states with probability one. In this figure, the vertices C, D, P of the simplex correspond to the three extreme states of $(20, 0, 0)$, $(0, 20, 0)$, and $(0, 0, 20)$ respectively. It can be seen from the figure that when α_2 increases, the limit probabilities of the system selecting $(i, 0, 20-i)$ states increase for small i, and decrease for large i. The figure shows the skewing of the states from vertex C to P on the PC edge of the simplex (the coexistence of the P and C strategies). Thus, the increase of α_2 will increase the proportion of the punishment strategy when the system chooses the "C+P" states with probability one.

According to Figs. 2–5, the conclusions can be drawn as follows. (i) The first-order costly punishment can change the evolutionary dynamics and equilibria of the public goods game, and it can promote cooperation only if both the punishment intensity and the return on investment are large enough. (ii) On the basis of the first-order punishment, the further imposition of the second-order punishment cannot change the evolutionary dynamics of the system dramatically, but only change the probability with which the system selects the exact equilibrium points in the "C+P" states. The second-order punishment has a limited role in promoting the cooperation in most cases, except for some critical combinations of (β, r). (iii) When the system chooses "C+P" states with probability one, the increase of the punishment probability under the second-order punishment will further increase the proportion of the punishment type strategies.

4. Conclusions

In this paper, we introduce two types of costly punishment strategies in the traditional public goods game. Considering continuous noise in the strategy evolution process, a stochastic dynamic model in a finite size population is established. The evolution of the system is described as a continuous and finite state multidimensional Markov process. We analyze the stochastic stable state of the evolutionary system by the limit distribution of the Markov process. The influences of parameters such as the return on investment coefficient, punishment intensity, and punishment probability on the cooperative behavior of the population are studied under the first and second order punishment respectively. Unlike the most commonly used replicator dynamic models in infinite populations, the equilibrium state based on the Markov process is stochastically stable, and it does not depend on the initial state of the system. Unlike the Moran process based model in finite populations, the model proposed in this paper is applicable to all possible states, while the Moran process using the fixation probability can only analyze the probabilities of extreme states that all individuals choose the same strategy. Therefore, the model in this paper has strong adaptability. The analysis framework can also be applied to many other cases.

References

[1] Johnson D D P, Stopka P and Knights S 2003 *Nature* **421** 911
[2] Cox C A 2015 *Econ. Lett.* **126** 63
[3] Croson R, Fatas E and Neugebauer T 2005 *Econ. Lett.* **87** 95
[4] Heap S P H, Ramalingam A and Stoddard B V 2016 *Econ. Lett.* **146** 4
[5] Wagener A 2016 *Econ. Lett.* **138** 34
[6] Kosfeld M, Okada A and Riedl A 2009 *Am. Econ. Rev.* **99** 1335
[7] Skatova A and Ferguson E 2011 *Pers. Individ. Dif.* **51** 237
[8] Volk S, Thoeni C and Ruigrok W 2011 *Pers. Individ. Dif.* **50** 810
[9] Pfattheicher S, Keller J and Knezevic G 2017 *Personality and Social Psychology Bulletin* **43** 337
[10] Abele S, Stasser G and Chartier C 2010 *Personality and Social Psychology Review* **14** 385
[11] Archetti M and Scheuring I 2012 *J. Theor. Biol.* **299** 9
[12] Hauert C 2010 *J. Theor. Biol.* **267** 22
[13] Sasaki T and Unemi T 2011 *J. Theor. Biol.* **287** 109
[14] Santos F C, Santos M D and Pacheco J M 2008 *Nature* **454** 213
[15] Bazzan A L C and Argenta V F 2012 *Adv. Complex Syst.* **15** 1250027
[16] The Anh H, Pereira L M and Lenaerts T 2017 *Auton. Agent Multi Agent Syst.* **31** 561
[17] Hauert C, Haiden N and Sigmund K 2004 *Discrete Cont. Dyn.-B* **4** 575
[18] Hauert C, De Monte S, Hofbauer J and Sigmund K 2002 *Science* **296** 1129
[19] Nikiforakis N 2010 *Games Econ. Behav.* **68** 689
[20] Brandt H, Hauert C and Sigmund K 2006 *Proc. Natl. Acad. Sci. USA* **103** 495
[21] Fowler J H 2005 *Proc. Natl. Acad. Sci. USA* **102** 7047
[22] Hauert C, Traulsen A, Brandt H, Nowak M A and Sigmund K 2007 *Science* **316** 1905
[23] Zhou Y, Jiao P and Zhang Q 2017 *Appl. Econ. Lett.* **24** 54
[24] Wang Z, Xu Z J, Huang J H and Zhang L Z 2010 *Chin. Phys. B* **19** 010204
[25] Wang Z, Xu Z J and Zhang L Z 2010 *Chin. Phys. B* **19** 010201
[26] Wang C, Wang L, Wang J, Sun S and Xia C 2017 *Appl. Math. Comput.* **293** 18
[27] Wang X, Chen X, Gao J and Wang L 2013 *Chaos, Solitons and Fractals* **56** 181
[28] Li A, Wu T, Cong R and Wang L 2013 *Europhys.* **103** 30007
[29] Chen M, Wang L, Sun S, Wang J and Xia C 2016 *Phys. Lett. A* **380** 40
[30] Sigmund K, Hauert C and Nowak M A 2001 *Proc. Natl. Acad. Sci. USA* **98** 10757
[31] Brandt H, Hauert C and Sigmund K 2003 *Proc. R. Soc. Lond. Ser. B* **270** 1099

[32] Hauert C, Traulsen A, Brandt H, Nowak M A and Sigmund K 2008 *Biol. Theory* **3** 114
[33] Page T, Putterman L and Unel B 2005 *Econ. J.* **115** 1032
[34] Hong F and Lim W 2016 *J. Econ. Behav. Organ.* **126** 102
[35] Shinohara R 2009 *Soc. Choice Welfare* **32** 367
[36] Szabo G and Hauert C 2002 *Phys. Rev. Lett.* **89** 118101
[37] Cao X B, Du W B and Rong Z H 2010 *Physica A* **389** 1273
[38] Fan R, Zhang Y, Luo M and Zhang H 2017 *Physica A* **465** 454
[39] Liu R R, Jia C X and Wang B H 2010 *Physica A* **389** 5719
[40] Xia H J, Li P P, Ke J H and Lin Z Q 2015 *Chin. Phys. B* **24** 040203
[41] Yönaç M, Huang Z G, Wang S J, Xu X J and Wang Y H 2008 *Europhys.* **81** 28001
[42] Yönaç M, Menon R and Korolev K S 2015 *Phys. Rev. Lett.* **114** 168102
[43] Wang X P, Jiang L L and Wang B H 2012 *Chin. Phys. B* **21** 070210
[44] Chan S, Reid M D and Ficek Z 2012 *Chin. Phys. B* **21** 018701
[45] Quan J and Wang X J 2011 *Chin. Phys. B* **20** 030203
[46] Chaudhuri A 2011 *Exper. Econ.* **14** 47
[47] Interis M G and Haab T C 2014 *J. Environ. Psychol.* **38** 271
[48] Fehr E and Fischbacher U 2004 *Evolution and Human Behavior* **25** 63
[49] Forsyth P A I and Hauert C 2011 *J. Math. Biol.* **63** 109
[50] Foster D and Young P 1990 *Theor. Popul. Biol.* **38** 219
[51] Young P 1993 *Econometrica* **61** 57
[52] Quan J and Wang X J 2013 *Commun. Theor. Phys.* **60** 37
[53] Quan J, Liu W, Chu Y and Wang X 2017 *Sci. Rep.* **7** 16110
[54] Huang W, Hauert C and Traulsen A 2015 *Proc. Natl. Acad. Sci. USA* **112** 9064
[55] Liang H, Cao M and Wang X 2015 *Syst. Control. Lett.* **85** 16
[56] Amir M and Berninghaus S K 1996 *Games Econ. Behav.* **14** 19
[57] Stewart W J 1994 *Introduction to the Numerical Solution of Markov Chains* (Princeton: Princeton University Press)