2024 IEEE 63rd Conference on Decision and Control (CDC)
December 16-19, 2024. MiCo, Milan, Italy

Evolution of Cooperation Among Unequals with Game-Environment Feedback

Xiaotong Yu, Haili Liang, Zhihai Rong, Xiaoqiang Ren, Xiaofan Wang and Ming Cao

***Abstract*—In social activities, the conflict between individual interests and group interests often leads to social dilemmas. The theory of direct reciprocity suggests that repeated interactions can alleviate this dilemma, but it often assumes homogeneity among individuals. However, the widespread heterogeneity in the real world often diminishes cooperation. Fortunately, the dynamics of bidirectional feedback between game mechanisms and environmental conditions can exert a positive influence on the evolutionary progression of cooperative behavior. This paper investigates how heterogeneous individuals affect the evolution of cooperation with environmental feedback. Despite extensive literature showing that endowment inequality among individuals tends to diminish cooperation, our findings suggest that the implementation of appropriate environmental feedback mechanisms can facilitate the development of cooperation. Furthermore, our results demonstrate that appropriate environmental feedback mechanisms can significantly augment the propensity of homogeneous individuals for cooperative behavior. These results provide insights for decision-makers in formulating strategies related to fairness and the sharing of public goods.**

***Index Terms*—Evolutionary games, stochastic games, public goods games, game-environmental feedback, heterogeneity.**

I. INTRODUCTION

In social dilemmas, rational individuals tend to defect due to shortsightedness, yet maximum collective benefits can only be achieved when all individuals choose to cooperate. This conflict between individual rationality and collective welfare is pervasive across various domains [1]–[5]. The theory of direct reciprocity, grounded in repeated interactions, serves as a potent mechanism for mitigating this conflict and fostering the evolution of cooperation [6]–[8]. It typically necessitates equality among individuals, such that those adopting the same strategy receive identical payoffs [9], [10].

Previous models have often assumed uniformity in factors such as individual contributions and the set of optional strategies, with only a limited number of models considering the influence of unequals on individual decision-making. However, heterogeneity is prevalent in the real world, emphasizing the importance of exploring its impact on individual cooperation. So far, the effect of heterogeneity on cooperation remains ambiguous [11]–[14]. Experimental findings in [11] indicate that when income is transparent, participants tend to penalize the wealthy and reward the poor. Conversely, when income is concealed, participants tend to penalize the poor and reward the wealthy. Moreover, disclosing income leads to greater overall contributions to public goods, suggesting that inequality may yield social benefits. Vasconcelos *et al.* [12] examine the impact of wealth inequality in countries on climate policies. Their model suggests that wealth inequality fosters greater global cooperation, with the affluent generally contributing more than the less affluent. In contrast, experiments conducted by [15] suggest that unequal individual endowments diminish cooperation. Similarly, experiments by [16] espouse a similar perspective, albeit designing their experiments to mitigate the influence of the total sum of endowments on the results. Hauser *et al.* [17] develop a model that comprehensively considers the impact of potential multi-source heterogeneity on the evolution of cooperation. They find that extreme inequality impedes cooperation, yet successful cooperation may necessitate a certain degree of unequal endowments if individuals' productivities vary.

The models mentioned above typically assume fixed and unchanging environmental resources. This implies that once players' strategies are determined, they receive the expected returns associated with those strategies. However, in many applications, especially those involving the consumption of common environmental resources, individuals have become aware of the dynamic nature of environmental resources. If environmental resources are limited, the "tragedy of the commons" becomes inevitable. Recent attention has been focused on exploring the dynamics of environmental resources changing over time due to individual behavior [18]–[21]. In these models, individuals' returns are influenced by the external environment, and conversely, individuals' strategies also change the surrounding environment, forming a bidirectional game-environment feedback loop. This line of research originated with [18], whose integrated dynamic model combines replicator dynamics of a two-person, two-strategy game with environmental resource dynamics, demonstrating rich dynamic behaviors. In most cases, the joint state of strategy and environment converges to an equilibrium point on the phase space boundary where the environmental state is zero, reflecting the tragedy of the commons. Furthermore, system dynamics may also converge to limit cycles on the boundary, where the environmental state cycles between low and high values, known as the "oscillating tragedy of the commons". Subsequently, Hilbe *et al.* [19] further introduced environmental feedback into the analysis of finite populations. Their model suggests that when the environmental feedback mechanism is such that cooperation leads to playing more valuable games, while defection leads to playing less

The work was supported in part by the National Natural Science Foundation of China under Grant 62273223, 62373153 and 62336005, and the Project of Science and Technology Commission of Shanghai Municipality under Grant 22JC1401401.

X. Yu, H. Liang, X. Ren and X. Wang are with School of Mechatronic Engineering and Automation, Shanghai University, Shanghai, China xiaotongyu@shu.edu.cn, lianghaili1016@163.com, xqren@shu.edu.cn, xfwang@shu.edu.cn

Z. Rong is with College of Information and Technology, Donghua University, Shanghai, China rongzhh@gmail.com

M. Cao is with ENTEG, Faculty of Science and Engineering, University of Groningen, 9747 AG, Groningen, The Netherlands m.cao@rug.nl

valuable games, it can significantly enhance individuals' tendency to cooperate. Intuitively, some environmental feedback mechanisms favor cooperation, while others may hinder it. The question of which environmental feedback mechanisms are favorable for cooperation is also of particular interest.

Research on game-environment feedback has been extensive, but there has been relatively little study on the effects of game-environment feedback on cooperation in the presence of heterogeneity among individuals. This paper contributes to the field in several ways. First, we propose a general model framework to study the effects of environmental feedback on the cooperation behavior of heterogeneous individuals. Next, we conduct simulation analyses to examine the outcomes of homogeneous individuals under various environmental feedback conditions and derived the condition under which the WSLS strategy is a subgame perfect Nash equilibrium.

The remaining sections of this paper are organized as follows. Section II introduces the stochastic game model, the calculation of payoffs for reactive strategies, and concludes with the strategy evolution process. Section III presents the main conclusions: firstly, the analysis of the impact of introducing appropriate environmental feedback on cooperation; secondly, the comparison of all possible environmental feedback modes; and finally, the analysis of the impact of introducing appropriate environmental feedback on individual cooperation behavior in the presence of heterogeneity among individuals. The last section summarizes our contributions and discusses future research directions.

II. MODEL

In this section, we begin by outlining a model for the repeated public goods game that incorporates time-varying environmental factors within a stochastic game framework. Subsequently, we introduce a memory-one strategy and detail the method for calculating the payoff function. Lastly, we employ "introspection" dynamics to elucidate the evolutionary process of strategies.

A. Stochastic Games

In this study, we consider a repeated public goods game with n players. At the beginning of each game round, player i receive a fixed endowments $e_i \geq 0$. To simplify the analysis, we assume $\sum_{i=1}^{n} e_i = 1$. Each player independently determines the proportion x_i of their received endowments e_i to contribute to the public goods, based on their previous contributions. Additionally, we assume that all players have identical productivity c_i, resulting in each player's actual contribution to the public pool being $c_i x_i e_i$. The total contribution in the public pool is then multiplied by time-varying environmental factors r_k and evenly distributed among all participating players. Player i's payoff u_i is contingent upon both the distribution of endowments $e_1, \ldots, e_n$ and the contributions of all players $x_1, \ldots, x_n$. The specifics are detailed as follows:

1) The set of players. The set $\mathcal{N} = \{1, 2, \ldots, n\}$ denotes the collection of players engaged in a public goods game.
2) The set of environment states. Let $\mathcal{S} = \{s_1, s_2, \ldots, s_m\}$ denote the set, where the environmental state represents the aggregate of all external and endogenous factors. The environmental factor r_k indicates the multiplication factor when the environmental state is s_k, with changes in the environmental state depending on the players' historical contributions. This paper assumes, without loss of generality, the presence of only two environmental states, s_1 and s_2, with the game initiating in state s_1.
3) The set of optional contribution for players. The set $\mathcal{X} = \{0, 1\}$ represents the optional contribution for players, where $x_i = 1$ indicates full contribution of the endowments (cooperation), and $x_i = 0$ indicates no contribution (defection). In each round of the game, players independently determine their contributions based on their historical actions. The contribution of all players at time t are represented by the vector $x(t) = (x_1(t), \ldots, x_n(t))$, where $x_i(t) \in \mathcal{X}$ denotes the contribution of player i at time t.
4) State transition function. We assume that the environmental state in which players are situated changes over time, and these changes depend on players' historical contributions. This is represented by the following function:
$$Q(x) = (z_1, \ldots, z_m),$$
where z_k represents the probability of the environmental state transitioning to s_k given the contributions x, and it satisfies $\sum_{k=1}^{m} z_k = 1$. Here, $x(t)$ is abbreviated as x, and the same notation applies throughout. we consider only the history of the previous round. When $m = 2$, the state transition function can be expressed as $q(x) = (q_n, q_{n-1}, \ldots, q_1, q_0)$, where q_n denotes the probability of transitioning to state s_1 in the current round when n players cooperated in the previous round, and $1 - q_n$ represents the probability of transitioning to state s_2 accordingly.
5) Payoff function. The payoff for player i is represented by the following function:
$$u_i^k(e, x) = \frac{r_k}{n} \sum_{j=1}^{n} c_j x_j e_j + (1 - x_i) e_i, \qquad (1)$$
where c_j represents player j's productivity, r_k denotes the multiplication factor when a player is in environmental state s_k. x_i and e_i represent player i's contribution proportion and the initial endowment received per round, respectively. To ensure that this game conforms to the standards of a social dilemma, the following three conditions need to be satisfied:
   a) Positive externalities (PE): Each player prefers that other players contributing to the game contribute more. For two contribution vectors x and x', if $x_i = x'_i$ and $x_j > x'_j$ for $i \neq j$, then $u_i(e, x) > u_i(e, x')$.
   b) Incentive to free-ride (IF): Each player strives to minimize their own contribution while benefiting

from others' contributions. For two contribution vectors x and x', if $x_i < x'_i$ and $x_j = x'_j$ for $i \neq j$, then $u_i(e, x) > u_i(e, x')$.

c) Optimality of cooperation (OC): Cooperation results in greater collective benefits. For two contribution vectors x and x', if $x < x'$, then $U(e, x) < U(e, x')$, where $U(e, x) = \sum_{i=1}^{n} u_i(e, x)$.

According to the given payoff function (1), if the game meets the requirements of a social dilemma, PE requires $c_i > 0$, IF requires $c_i r_k < n$, and OC requires $c_i r_k > 1$. Therefore, to satisfy the conditions of a social dilemma, the following equations should hold:

$$\begin{cases} c_i > 0, & i \in \mathcal{N} \\ 1 < c_i r_k < n, & i \in \mathcal{N}, k \in \{1, \ldots, m\}. \end{cases} \tag{2}$$

Assuming individuals interact for an infinite number of rounds, with t denoting the round of the game, the average payoff per round of the game for individual i is

$$\pi_i = \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} u_i^{k(t)}(e, x(t)). \tag{3}$$

B. Memory-One Strategy

In this paper, the memory-one strategy is represented by a vector $p = (p_0, p_x)$, where p_0 denotes the probability of a player choosing cooperation in the first round of the game, and p_x represents the probability of a player choosing cooperation in the current round given that the contribution distribution of players in the previous round was $x \in \mathcal{X}^n$. The strategy p encodes the conditional probabilities of choosing cooperation in the current round for all possible contribution distributions in the previous rounds. When there are n players, the number of dimensions in the memory-one strategy is $2^n + 1$. We assume that strategies have a execution error ϵ, a player with strategy p effectively implements the strategy $(1 - \epsilon)p + \epsilon(1 - p)$. Games with errors exhibit a favorable property in long-term dynamics: they can disregard the influence of players' initial actions [22]. Therefore, the construction of the strategy can disregard p_0, and the number of dimensions in the strategy becomes 2^n.

When all n players employ memory-one strategies $p^1, p^2, \ldots, p^n$, the dynamics of the stochastic game can be modeled as a Markov chain with state (s, x). Here, (s, x) represents the combination of the environmental state s and the possible contribution distribution x. In the case of m environmental states, this Markov chain encompasses $m2^n$ possible states. The elements of the corresponding state transition probability matrix consist of two parts: the transition probabilities Q between environmental states in consecutive game rounds, and the probability p^n that a player chooses cooperation in the current round, given the contribution distribution x from the previous round. The specific calculation method is outlined as follows:

$$M_{(s_k, x) \to (s_{k'}, x')} = Q(s_{k'}|x) \prod_{i \in \mathcal{N}} y_i,$$

where

$$y_i = \begin{cases} p_x^i & \text{if } x'_i = 1 \\ 1 - p_x^i & \text{if } x'_i = 0, \end{cases}$$

$Q(s_{k'}|x)$ denotes the probability of transitioning to state $s_{k'}$ in the current round given the contribution proportion x of players in the previous round, and p_x^i represents the probability that player i chooses cooperation in the current round given the contribution distribution x from the previous round. According to our setup, the steady distribution is independent of the initial distribution. The steady distribution is the left eigenvector of the state transition probability matrix M with an eigenvalue of 1. The elements of vector v, denoted as $v_{(s,x)}$, represent the expected frequency of each Markov state. According to equation (3), the expected payoff per round is

$$\pi_i = \sum_{s,x} v_{(s,x)} u_i^s(e, x).$$

C. Evolutionary Process

In this subsection, we primarily explore the evolution of players' strategies over time. In scenarios where payoffs are symmetric, players employing identical strategies receive equivalent payoffs. Evolutionary advancements can occur through the imitation of more successful strategies or by successful strategies leading to increased offspring. However, in games characterized by unequal payoffs, the intuitive impracticality of imitating other players' strategies arises because a strategy yielding high payoffs for one player may result in low payoffs for others. While numerous studies have examined evolutionary games with unequal payoffs, their frameworks often categorize populations, restricting interactions to occur solely among individuals within the same category. This approach lacks realism when considering populations where every individual differs in some aspect. Hauser *et al.* [17] introduced a novel evolutionary process in which individuals do not imitate others' strategies but instead update their own. An individual randomly generates a new strategy, compares the payoffs of the new and original strategies, and adopts the new strategy with a certain probability if it yields a higher payoff.

In this paper, at each evolutionary time step t, one individual i is randomly selected from the population of n individuals to update their strategy. Individual i randomly selects a new strategy from the set of available strategies (while the strategies of other players remain unchanged), then compares the payoffs of the new and old strategies. The selected player switches to the new strategy with the following probability [23]:

$$\rho = \frac{1}{1 + exp\{-\sigma(\tilde{\pi}_i - \pi_i)\}}, \tag{4}$$

where π_i and $\tilde{\pi}_i$ represent the payoffs of individual i for the old and new strategies, respectively, and the parameter $\sigma \geq 0$ denotes the selection intensity. As $\sigma \to 0$, the switching probability $\rho = 1/2$, indicating a random strategy switch independent of the payoff. When $\sigma \to \infty$, an individual adopts the new strategy only if its payoff exceeds that of the old strategy.

Assuming individuals employ pure one-step memory strategies, the strategy space is finite. It is possible to calculate the expected trajectory of introspective dynamics. With a total of l strategies, introspective dynamics can be modeled as a Markov process with l^n states, denoted as $P = (p^1, p^2, \ldots, p^n)$. The state transition probabilities can be represented as:

$$W_{P \to \tilde{P}} = \begin{cases} \frac{1}{n(l-1)} \rho(\pi_i, \tilde{\pi}_i) & \text{if } p^i \neq \tilde{p}^i, \text{and } p^j = \tilde{p}^j \text{ for all } j \neq i, \\ 1 - \sum_{\tilde{p}^i \neq p'} \frac{1}{n(l-1)} \rho(\pi_i, \tilde{\pi}_i) & \text{if } P = \tilde{P}, \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

Due to the random selection of only one player i from the n players to update their strategy at each evolutionary time step, only one player's strategy changes at each time step, with player i being selected with a probability of $1/n$. Once selected, player i can randomly choose a new strategy from the strategy set (if different from the old strategy) with a probability of $\frac{1}{l-1}$. Subsequently, the player switches to the new strategy with a probability of $\rho(\pi_i, \tilde{\pi}_i)$. The matrix W in (5) describes the state transition probabilities of the Markov process representing introspective dynamics. Assuming the initial distribution of given strategies is $v(0)$, the distribution of strategies after t time steps is $v(t) = v(0)W^t$. Since the selection intensity is finite, thus, no strategy can guarantee non-replacement. When t is sufficiently large, the distribution of strategies converges to a unique steady distribution v. This steady distribution, representing the expected frequency of strategies in introspective dynamics, is obtained by solving for the left eigenvector of W corresponding to the eigenvalue 1.

D. Problem of Interests

In this study, we investigate the influence of dynamic environments relative to fixed environments on individual cooperation and strategy selection throughout the evolutionary process. Furthermore, we analyze interactions among individuals with varying endowments and examine how the introduction of environmental feedback affects the decisions of individuals with unequal endowments.

III. MAIN RESULTS

"The tragedy of the commons" represents a fundamental problem. To alleviate this issue, we propose a general framework featuring environmental feedback to avoid social dilemmas. Initially, we illustrate the positive impact of suitable environmental feedback on individual cooperative behavior through three examples. These examples are all executed within the framework of section II, with differences solely in the state transition function $q(x)$.

Consider the simplest scenario: three individuals participate in a public goods game, interacting for an infinite number of rounds. In each round, individuals are in one of two environmental states, s_1 or s_2, determined by the state transition function $q(x)$. Each individual can choose to contribute fully (cooperate) or contribute nothing (defect), with their payoff determined by (1). In the first example, individuals only play in the environmental state with a lower payoff, regardless of their contributions in the previous round. The state transition function is $q_2(x) = (0, 0, 0, 0)$ (see Fig. 1a), where all elements are 0, indicating that the probability of transitioning to the environmental state with a higher payoff is 0. This means that the environmental state remains fixed at s_2, regardless of how many individuals chose to cooperate in the previous round. In the second example, individuals only play in the environmental state with a higher payoff, independent of their contributions in the previous round. The state transition function is given by $q_1(x) = (1, 1, 1, 1)$ (see Fig. 1a). Both of the above examples are considered instances of fixed environment scenarios, both of which are within the context of a social dilemma. Additionally, in the third example, we consider that the environmental state changes over time, depending on the contributions in the previous round. The state transition function is $q_s(x) = (1, 0, 0, 0)$ (see Fig. 1a), where the transition probability to the higher-payoff environmental state is 1 only if all individuals contribute fully in the previous round; otherwise, it is 0. This implies that individuals aiming to play in the higher-payoff environment require full cooperation in the previous round; otherwise, the environmental state deteriorates.

When simulating the evolutionary dynamics of the three scenarios, we observed a continuous learning process towards cooperation among individuals, both in situations where they exclusively played in a high-yield environment and in scenarios involving a changing environment (see Fig. 1b). However, the latter exhibited a higher cooperation rate. Although it might appear that maintaining a fixed environment state at s_1 would yield higher returns, the introduction of an additional, relatively poorer environment state s_2 as part of the environmental feedback actually favored the evolution of cooperation.

Furthermore, we noted that compared to a fixed environment state, introducing environmental feedback resulted in a quicker attainment of a cooperative equilibrium. To elucidate this outcome, we numerically computed the prevalence of individual strategies in the evolutionary process of the three scenarios (see Fig. 1c). In the first scenario, no strategy emerged as dominant, thereby impeding the sustainability of cooperation among individuals. In contrast, in the second and third scenarios, individuals predominantly adopted the well-known WSLS strategy $p = (1, 0, 0, 0, 0, 0, 0, 1)$ [10]. The distinction lies in the fact that in the third scenario, the prevalence of the WSLS strategy was higher than that in the second scenario. As a population comprised entirely

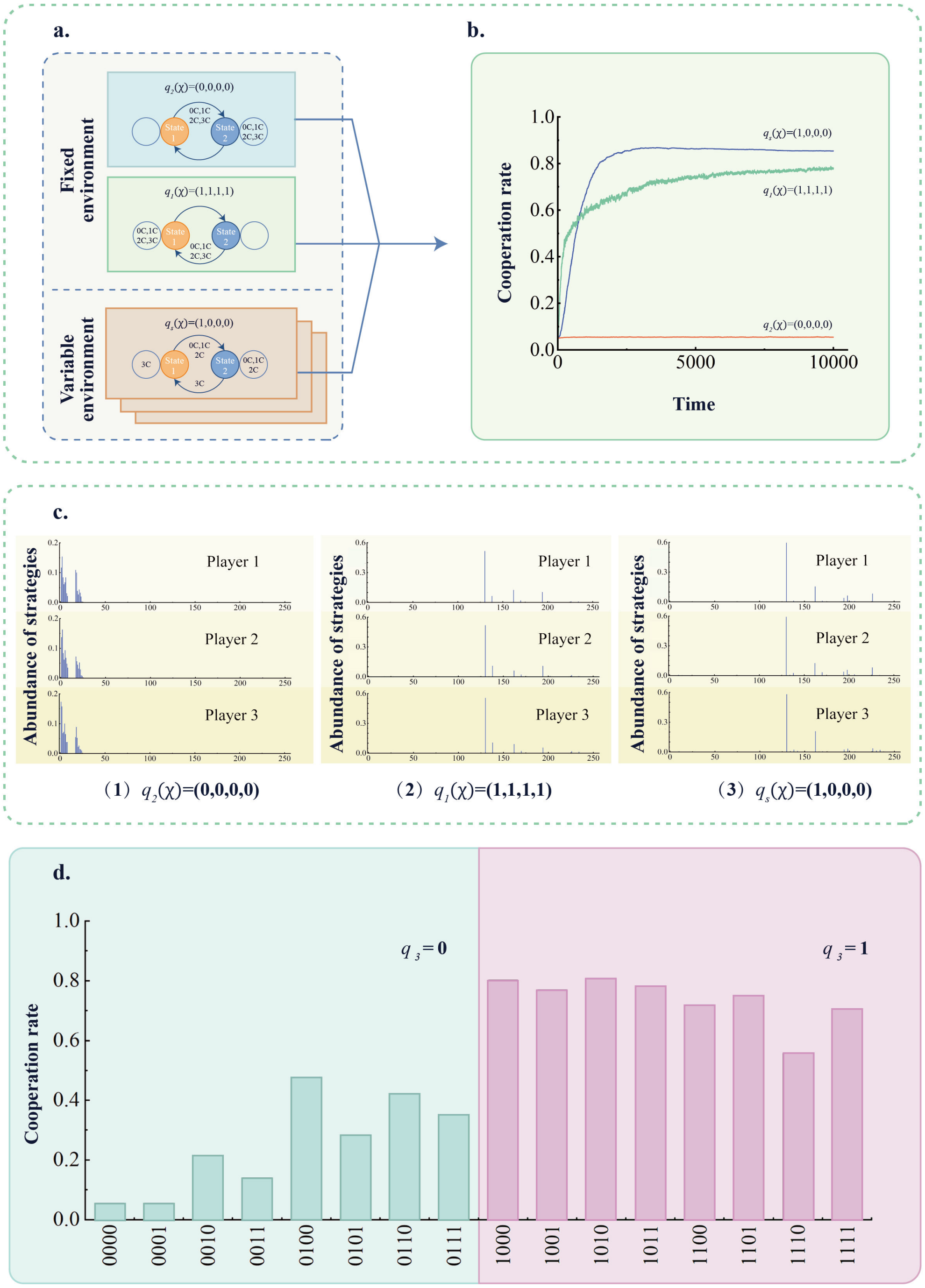
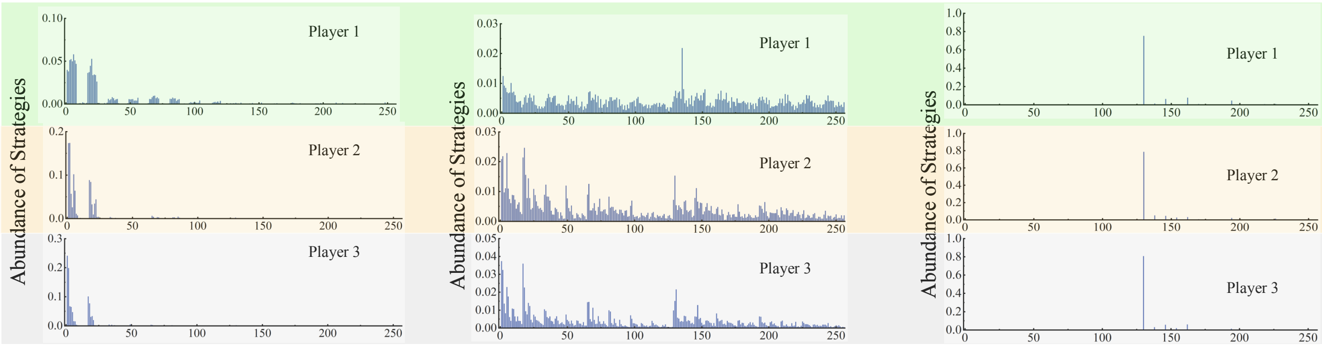


Fig. 1: Introducing appropriate environmental feedback can enhance population learning for cooperation. **a**, We examine the impact of three types of environmental state transition functions on individual cooperative behavior: exclusive play in a low-yield environment, exclusive play in a high-yield environment, and alternating between high-yield and low-yield environments. **b**, The different evolutionary trajectories correspond to state transition functions $q_2(x) = (0,0,0,0)$ (red line), $q_1(x) = (1,1,1,1)$ (green line), and $q_s(x) = (1,0,0,0)$ (blue line). **c**, The prevalence of strategies among individuals across 10^5 time steps under three types of state transition functions. The horizontal axis represents the decimal number corresponding to strategies converted from an 8-bit binary representation. For instance, the WSLS strategy $p = (1,0,0,0,0,0,0,1)$ is converted to the decimal number 130. **d**, The mean cooperation rate across all potential state transition functions during the evolutionary process. Parameters: $e_1 = e_2 = e_3 = 1/3$, $c_1 = c_2 = c_3 = 2$, $r_1 = 1$, $r_2 = 0.5$, $\epsilon = 0.05$, $nGen = 10^4$, $s = 1000$.



(1) $q_2(x) = (0,0,0,0)$ (2) $q_1(x) = (1,1,1,1)$ (3) $q_s(x) = (1,0,0,0)$

Fig. 2: The distribution of strategies among individuals over 10^5 time steps is analyzed under the three state transition functions. The horizontal axis corresponds to the decimal number derived from an 8-bit binary representation of strategies. For instance, the WSLS strategy $p = (1,0,0,0,0,0,0,1)$ is represented by the decimal number 130. The parameters are the same as in Fig. 3.

of WSLS strategies can achieve full cooperation, a higher prevalence of the WSLS strategy during the evolutionary process corresponds to a higher average cooperation rate among individuals.

Theorem 1. *Consider a public goods game with n players and a decreasing multiplication factor $r_1 > r_2 > \cdots > r_m$. The strategy WSLS is subgame perfect Nash equilibrium if and only if the following condition are met in all environment s_o*

$$(2q_n - q_0)(r_1 - r_2) \geq \frac{(2n - r_o c_i)e_i}{\sum_{j=1}^{n} c_j e_j} - r_2. \tag{6}$$

Proof. We use a one-shot deviation principle to prove that the WSLS strategy is a subgame perfect Nash equilibrium: in a population entirely adopting the WSLS strategy, if one player deviates from WSLS in a single round due to error but returns to WSLS in subsequent rounds, they cannot achieve a higher payoff. We assume that future payoffs will be discounted by a factor of δ. We consider two cases based on the nature of the WSLS strategy: 1) in the initial or pre-deviation round, all players contribute equally; 2) in the pre-deviation round, players' contributions are unequal.

First, we consider the first case. Assuming that in the initial or previous round of the game, all players contributed equally and the current environmental state is s_o, the payoff of the WSLS strategy in the current and subsequent rounds of the game can be expressed as follows::

$$\pi_R = (1-\delta)u_i^o(e,1) + \delta \sum_{k=1}^{m} Q(s_k|n) u_i^k(e,1).$$

If a mutation occurs at this time, causing a deviation from the WSLS strategy in the current round of the game, and then returning to the WSLS strategy in the next round, the payoff of the mutation is

$$\begin{aligned}\pi_M = &(1-\delta)u_i^o(e,1_{-i}) \\ &+ \delta \sum_{k=1}^{m} Q(s_k|n-1)\Big\{(1-\delta)u_i^k(e,0) \\ &+ \delta \sum_{w=1}^{m} Q(s_w|0)\Big\{(1-\delta)u_i^w(e,1) \\ &+ \delta \sum_{p=1}^{m} Q(s_p|n)u_i^p(e,1)\Big\}\Big\}.\end{aligned}$$

For the second case, assuming that in the initial or previous round of the game, players' contributions were different and the current environmental state is s_v, the payoff of the WSLS strategy in the current round and subsequent rounds is

$$\begin{aligned}\hat{\pi}_R = &(1-\delta)u_i^o(e,0) \\ &+ \delta \sum_{k=1}^{m} Q(s_k|0)\Big\{(1-\delta)u_i^k(e,1) \\ &+ \delta \sum_{w=1}^{m} Q(s_w|n)u_i^w(e,1)\Big\}.\end{aligned}$$

If a mutation occurs at this point, causing a deviation in the individual's decision: instead of choosing betrayal, the individual chooses cooperation, and then returns to the WSLS strategy in the next round. The player's payoff becomes

$$\begin{aligned}\hat{\pi}_M = &(1-\delta)u_i^o(e,0_i) \\ &+ \delta \sum_{k=1}^{m} Q(s_k|1)\Big\{(1-\delta)u_i^k(e,0) \\ &+ \delta \sum_{w=1}^{m} Q(s_w|0)u_i^w(e,1)\Big\}.\end{aligned}$$

According to the definition of subgame perfect Nash equilibrium, the WSLS strategy is a subgame perfect Nash equilibrium if and only if $\pi_R \geq \pi_M$ and $\hat{\pi}_R \geq \hat{\pi}_M$. That is,

$$\begin{aligned} & u_i^o(e,1) - u_i^o(e,1_{-i}) \\ & + \delta \sum_{k=1}^{m} \Big\{ Q(s_k|n) u_i^k(e,1) - Q(s_k|n-1) u_i^k(e,0) \Big\} \\ & + \delta^2 \sum_{k=1}^{m} \Big\{ Q(s_k|n) u_i^k(e,1) - Q(s_k|0) u_i^k(e,1) \Big\} \\ & \geq 0, \\ & u_i^o(e,0) - u_i^o(e,0_i) \\ & + \delta \sum_{k=1}^{m} \Big\{ Q(s_k|0) u_i^k(e,1) - Q(s_k|1) u_i^k(e,0) \Big\} \\ & + \delta^2 \sum_{k=1}^{m} \Big\{ Q(s_k|n) u_i^k(e,1) - Q(s_k|0) u_i^k(e,1) \Big\} \\ & \geq 0. \end{aligned} \tag{7}$$

Combining equation (1) with $m = 2$ and $\delta = 1$, condition(7) can be simplified to

$$(2q_n - q_0)(r_1 - r_2) \geq \frac{(2n - r_o c_i)e_i}{\sum_{j=1}^{n} c_j e_j} - r_2. \tag{8}$$

□

The preceding three scenarios demonstrate the significant impact of environmental state-switching mechanisms, which modify players' strategy selections and thereby their inclination to cooperate. In essence, introducing appropriate environmental feedback may be more favorable for cooperation compared to a fixed environmental. In our model, with a game involving three players, there are a total of $2^{3+1} = 16$ state transition functions. To investigate the effects of environmental feedback mechanisms on cooperation, we enumerated all possible environmental feedback mechanisms and computed their average cooperation rates over 10^4 evolutionary steps (see Fig. 1d). Surprisingly, apart from the state transition function $q_s(x)$ employed in the third scenario, there exist 7 other state transition functions that can drive the population toward higher cooperation rates. Their common characteristic is that when all individuals opt to cooperate in the preceding round, the environmental state improves. This indicates that introducing specific positive incentives when all individuals cooperate can facilitate population-wide learning to cooperate.

Heterogeneity is prevalent in social systems, and numerous experiments have demonstrated that unequal endowments diminish cooperation. However, introducing environmental feedback under conditions of unequal endowments yields unexpected outcomes. It is noteworthy that, due to endowment heterogeneity, individuals' payoffs become asymmetric. Cooperation rates and the prevalence of strategies among individuals may vary throughout the evolutionary process. In addition to varying endowments, we conducted numerical simulations using the three aforementioned scenarios. We observed that, compared to the scenario depicted by the green line in Fig. 1b, which represents equal endowments,

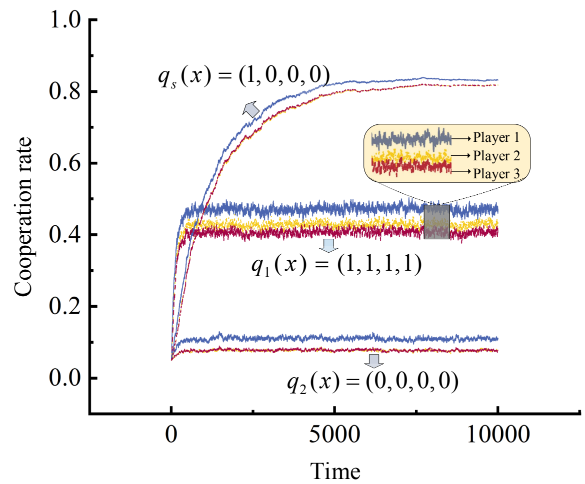


Fig. 3: Introducing appropriate environmental feedback facilitates population learning to cooperate, even in the presence of heterogeneity among individuals. We examine the influence of the three environmental state transition functions depicted in Fig. 1a on unequal individual cooperative behavior. The parameters are the same as in Fig. 1 except for $e_1 = 0.1$,$e_2 = 0.4$, $e_3 = 0.5$.

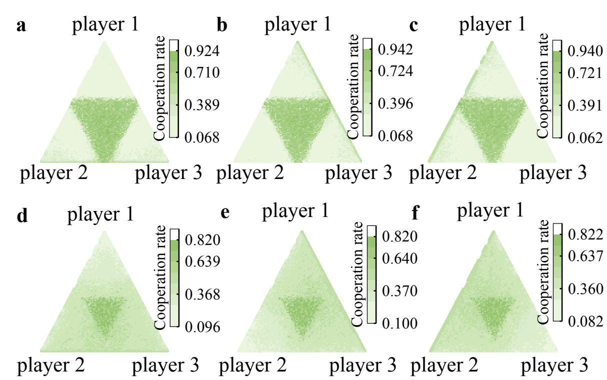


Fig. 4: Introducing appropriate environmental feedback can enhance the tolerance of cooperative individuals to unequal endowments. We consider a public goods game involving three players and compare the effects of heterogeneous endowments on Player 1 (a, d), Player 2 (b, e), and Player 3 (c, f) under the conditions of environmental feedback introduction $(a - c)$ and fixed environment $(d - e)$. The triangle represents the possible initial endowment distributions for players, where each vertex corresponds to one player receiving all the initial endowments, and the edges represent players not receiving any initial endowments relative to the corresponding vertices. The center of the triangle represents all three players receiving the same initial endowments. The cooperation rate is the average cooperation rate of players over 10^4 time steps in the evolutionary process. Parameters: $c_1 = c_2 = c_3 = 2$, $r_1 = 1$, $r_2 = 0.5$, $\epsilon = 0.05$, $nGen = 10^4$, $s = 1000$, $q_1 = (1,1,1,1)$, $q_s = (1,0,0,0)$.

the scenario illustrated by the $q_1(x)$ in Fig. 3, reflecting unequal endowments, indeed diminishes cooperation, thereby validating certain prior experimental findings. Moreover, we discovered that the introduction of environmental feedback significantly enhances individuals' propensity to cooperate, even in the presence of heterogeneity (see Fig. 3). To elucidate this outcome, we computed the prevalence of strategies for each individual over 10^5 evolutionary steps (see Fig. 2). Our findings indicate that in the presence of heterogeneity, no strategy emerges as dominant in the evolutionary process under a fixed environment. Conversely, the introduction of suitable environmental feedback leads individuals to predominantly adopt the WSLS strategy, facilitating cooperative behavior.

To further investigate the influence of environmental feedback and heterogeneous endowments on cooperation, we utilized the state transition functions $q_1(x)$ and $q_s(x)$ to respectively represent scenarios with fixed environmental states and introduced environmental feedback. We then systematically varied the initial endowment distribution in increments of 0.01 and calculated the average cooperation rate for each distribution during the evolutionary process (see Fig. 4). Our analysis revealed that, compared to a fixed environmental state, the introduction of suitable environmental feedback enables a broader range of initial endowment distributions conducive to the evolution of cooperation. Notably, even when a player's endowment is 0, it does not influence their decision to cooperate. In essence, the introduction of appropriate environmental feedback enhances the capacity of cooperative individuals to withstand greater levels of inequality.

IV. CONCLUSIONS

In summary, we present a novel framework incorporating environmental feedback, applicable across a broad spectrum of human decision-making scenarios in social dilemmas characterized by individual heterogeneity. Initially, we examine the influence of environmental feedback on the decision-making behavior of homogeneous individuals and the prevalence of strategies during the evolutionary process. Our model demonstrates that appropriate environmental feedback fosters cooperation. Furthermore, we derived the conditions under which the strategy that promotes cooperation is subgame perfect Nash equilibrium. To further elucidate the impacts of various environmental feedback mechanisms, we analyze the cooperation rates of individuals across all potential scenarios. Additionally, we investigate the effects of environmental feedback on the decision-making behavior of heterogeneous individuals. Our model corroborates the experimental finding that heterogeneous endowments diminish cooperation. Importantly, even in the presence of heterogeneity, the introduction of suitable environmental feedback can enhance individual cooperation and expand the spectrum of initial endowment distributions conducive to the evolution of cooperation.

REFERENCES

[1] M. Chica, J. M. Hernández, and J. Bulchand-Gidumal, "A collective risk dilemma for tourism restrictions under the covid-19 context," *Scientific Reports*, vol. 11, no. 1, p. 5043, 2021.

[2] S. Vesely, C. A. Klöckner, and C. Brick, "Pro-environmental behavior as a signal of cooperativeness: Evidence from a social dilemma experiment," *Journal of Environmental Psychology*, vol. 67, p. 101362, 2020.

[3] J. Riehl, P. Ramazi, and M. Cao, "A survey on the analysis and control of evolutionary matrix games," *Annual Reviews in Control*, vol. 45, pp. 87–106, 2018.

[4] L. Gong, J. Gao, and M. Cao, "Evolutionary game dynamics for two interacting populations in a co-evolving environment," in *2018 IEEE Conference on Decision and Control (CDC)*. IEEE, 2018, pp. 3535–3540.

[5] H. Liang, M. Cao, and X. Wang, "Analysis and shifting of stochastically stable equilibria for evolutionary snowdrift games," *Systems & Control Letters*, vol. 85, pp. 16–22, 2015.

[6] E. Ostrom, "Tragedy of the commons," *The New Palgrave Dictionary of Economics*, vol. 2, pp. 1–4, 2008.

[7] R. L. Trivers, "The evolution of reciprocal altruism," *The Quarterly Review of Biology*, vol. 46, no. 1, pp. 35–57, 1971.

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

[9] M. Doebeli and C. Hauert, "Models of cooperation based on the prisoner's dilemma and the snowdrift game," *Ecology Letters*, vol. 8, no. 7, pp. 748–766, 2005.

[10] M. Nowak and K. Sigmund, "A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game," *Nature*, vol. 364, no. 6432, pp. 56–58, 1993.

[11] O. P. Hauser, G. T. Kraft-Todd, D. G. Rand, M. A. Nowak, and M. I. Norton, "Invisible inequality leads to punishing the poor and rewarding the rich," *Behavioural Public Policy*, vol. 5, no. 3, pp. 333–353, 2021.

[12] V. V. Vasconcelos, F. C. Santos, J. M. Pacheco, and S. A. Levin, "Climate policies under wealth inequality," *Proceedings of the National Academy of Sciences*, vol. 111, no. 6, pp. 2212–2216, 2014.

[13] C. Deng, L. Wang, Z. Rong, and X. Wang, "Cooperation emergence in group population with unequal competitions," *Europhysics Letters*, vol. 131, no. 2, p. 28001, 2020.

[14] K. Safarzynska and P. E. Smaldino, "Reducing global inequality increases local cooperation: a simple model of group selection with a global externality," *Philosophical Transactions of the Royal Society B*, vol. 379, no. 1893, p. 20220267, 2024.

[15] T. L. Cherry, S. Kroll, and J. F. Shogren, "The impact of endowment heterogeneity and origin on public good contributions: evidence from the lab," *Journal of Economic Behavior & Organization*, vol. 57, no. 3, pp. 357–365, 2005.

[16] S. P. H. Heap, A. Ramalingam, and B. V. Stoddard, "Endowment inequality in public goods games: A re-examination," *Economics Letters*, vol. 146, pp. 4–7, 2016.

[17] O. P. Hauser, C. Hilbe, K. Chatterjee, and M. A. Nowak, "Social dilemmas among unequals," *Nature*, vol. 572, no. 7770, pp. 524–527, 2019.

[18] J. S. Weitz, C. Eksin, K. Paarporn, S. P. Brown, and W. C. Ratcliff, "An oscillating tragedy of the commons in replicator dynamics with game-environment feedback," *Proceedings of the National Academy of Sciences*, vol. 113, no. 47, pp. E7518–E7525, 2016.

[19] C. Hilbe, Š. Šimsa, K. Chatterjee, and M. A. Nowak, "Evolution of cooperation in stochastic games," *Nature*, vol. 559, no. 7713, pp. 246–249, 2018.

[20] L. Gong, W. Yao, J. Gao, and M. Cao, "Limit cycles analysis and control of evolutionary game dynamics with environmental feedback," *Automatica*, vol. 145, p. 110536, 2022.

[21] A. R. Tilman, J. B. Plotkin, and E. Akçay, "Evolutionary games with environmental feedbacks," *Nature Communications*, vol. 11, no. 1, p. 915, 2020.

[22] K. Sigmund, *The calculus of selfishness*. Princeton University Press, 2010.

[23] A. Traulsen, M. A. Nowak, and J. M. Pacheco, "Stochastic dynamics of invasion and fixation," *Physical Review E*, vol. 74, no. 1, p. 011909, 2006.