Coevolutionary dynamics of collective cooperation and dilemma strength in a collective-risk game

Shijia Hua, Mingquan Xu, and Linjie Liu*
College of Science, Northwest A & F University, Yangling 712100, China

Xiaojie Chen†
School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 611731, China

(Received 23 April 2024; accepted 4 June 2024; published 24 June 2024)

Human behavioral decision-making influences the gaming environment. In turn, changes in the gaming environmental impact individual strategic choices. However, there is scant exploration into how human behavioral decision-making coevolves with dilemma strength. Here, we propose a coevolutionary game model based on collective-risk social dilemma, where an increase in cooperators within the game group reduces the dilemma strength, and vice versa. Upon examining this coupled system, we find that the system is capable of achieving a relatively optimal state, wherein the population sustains a high level of cooperation and the dilemma strength remains at the lowest level. In addition, we have identified the conditions for the emergence of tristability and bistability in the coupled system and numerically validated our theoretical results. Furthermore, we find that the incorporation of institutional rewards not only promotes the appearance of the system's optimal state, where all individuals choose to cooperate and the dilemma strength is at its lowest level, but it also effectively averts the manifestation of the system's worst state, where all individuals resort to defection and the dilemma strength reaches its highest level. These findings illuminate how cooperation can be sustained when a dynamical coupling exists between individual decision-making and dilemma strength.

DOI: 10.1103/PhysRevResearch.6.023313

I. INTRODUCTION

In the current global landscape, humanity is faced with an array of daunting challenges that necessitate a cooperative approach. These challenges range from the escalating severity of climate change to the uncontrolled spread of infectious diseases, each presenting a complex problem that surpasses the capacity of any single entity to resolve [1–3]. The imperative for collective action is clear. Yet, within the context of natural selection, the dynamics are not always conducive to cooperation. Individuals who opt to cooperate often find themselves disadvantaged relative to those who choose not to cooperate. This paradox, known as the "tragedy of the commons," is a fundamental problem in different disciplines [4–7]. It describes a scenario in which individual actions, while rational from a personal standpoint, collectively lead to a situation that is suboptimal for the group.

Evolutionary game theory provides a theoretical framework for examining the aforementioned issues [8–12]. Over the past few decades, various game models have been proposed to depict real-world scenarios, including the Prisoner's Dilemma [13–15], stag hunt [16,17], and public goods games [18–22]. The collective-risk social dilemma (CRD), also known as the threshold public goods game, has drawn considerable attention due to its implications in areas such as climate negotiations and infectious disease transmission [23–30]. In this type of game, a group of individuals must collectively decide whether to contribute resources towards a common goal. If the total contribution from the group surpasses a predetermined threshold, all individuals reap the benefits. However, if the total contribution falls short of this threshold, all individuals lose all their endowment with a probability r, which is referred to as the risk of collective task failure.

Prior theoretical research has delved into the evolution of cooperation in collective-risk social dilemma games from various perspectives, such as asset heterogeneity [28,31–33], population structure [34], incentives [35–37], and repeated group interactions [38]. However, most of these studies contemplate a static game environment, where an individual's behavioral decisions do not impact the game environment. This assumption is evidently at odds with reality. The coevolutionary game theory proposed in recent years offers a fresh perspective for studying this inconsistency [39–43]. For instance, Weitz *et al.* [39] introduced a feedback-evolving game model, where an individual's behavioral decisions alter the state of the environment, which in turn influences the individual's payoffs. Arefin and Tanimoto [44] constructed feedback-evolving game models that consider replicator dynamics and aspiration dynamics, respectively. Liu *et al.* [45] constructed a feedback-evolving game model based on CRD, capturing the coupled dynamics of cooperation and risk. They

*Contact author: linjieliu1992@nwafu.edu.cn
†Contact author: xiaojiechen@uestc.edu.cn

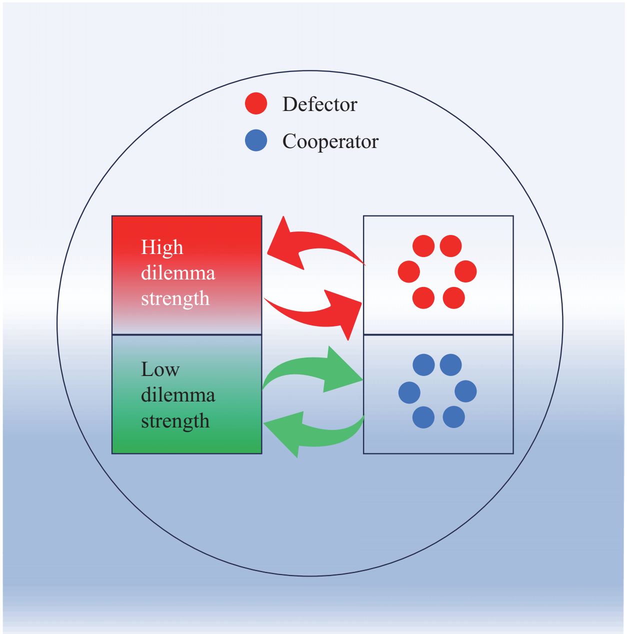

FIG. 1. Illustration of the coevolutionary process of strategies and dilemma strength in the collective-risk social dilemma. The increase of cooperators within a group diminishes the dilemma strength, facilitating the propensity for individual cooperation. Conversely, an increase in the number of defectors incrementally elevates the dilemma strength, thereby inhibiting the inclination towards cooperative behavior among individuals.

revealed that the bidirectional feedback between collective behavior and risk is beneficial for avoiding the tragedy of the commons. Despite previous research that shed light on various aspects of CRD, a crucial reciprocal relationship has remained largely unexplored. This involves the dynamic interplay where collective behaviors shape the dilemma strength, which in turn affects individual behaviors. Indeed, the dilemma strength has recently been explored in various 2×2 games [46–50], but they usually assume that the dilemma strength does not change with the state of the population. Hence, the nature of the coevolutionary dynamics resulting from the unidirectional feedback between collective behavior and dilemma strength remains an enigma. Moreover, previous studies have affirmed an institutional reward as the optimal mechanism to foster cooperation, outperforming punishments or mixed approaches [35,51]. Yet, its role in a game model where there is a feedback relationship between population status and dilemma strength is still undetermined.

In this present study, we thus construct a feedback-evolving game model based on the collective-risk social dilemma game to encapsulate the coupled relationship between strategies and the dilemma strength (see Fig. 1). In a scenario similar to the previously discussed stochastic game [52], we consider that an increase in the proportion of cooperators within the game group will mitigate the dilemma strength, while an increase in defectors will exacerbate it. Through the analysis of the coevolutionary dynamics of this coupled system, we uncover some intriguing dynamical phenomena, such as tristability and bistability. However, the system's optimal state, where all individuals choose to cooperate and the dilemma strength remains at its lowest, is unattainable. Additionally, the worst state of the system, where all individuals choose to defect and the dilemma strength persists at its highest, is invariably stable. By introducing institutional rewards, we observe that not only is the system's optimal state attainable, but the worst state can also be effectively circumvented.

II. THEORETICAL MODEL AND METHODS

A. Collective-risk social dilemmas

In a collective-risk social dilemma game, each individual starts with a constant initial endowment b, and it has two strategies available for selection: cooperation (C), which involves contributing c to the public pool, and defection (D), which involves no contribution to the public pool. If the number of individuals opting to contribute falls short of the collective target (n_{pg}), a disaster will occur with a probability r, resulting in all individuals losing their entire endowment. However, if the collective target is met, all individuals are able to retain their initial endowment. Here, n_{pg} represents the minimal number of collective investors to ensure the initial endowment of each individual. Therefore, the payoffs for cooperators and defectors in a collective-risk social dilemma game can be expressed as

$$P_C = -c + b\theta(N_C + 1 - n_{pg}) + (1 - r)b[1 - \theta(N_C + 1 - n_{pg})], \quad (1)$$

$$P_D = b\theta(N_C - n_{pg}) + (1 - r)b[1 - \theta(N_C - n_{pg})], \quad (2)$$

where N_C denotes the number of cooperators in the group, and $\theta(k) = 1$ when $k \geqslant 0$, otherwise it equals 0. Equation (1) represents the payoff for a cooperator, where the first term signifies the cost associated with cooperative behavior, the second term indicates the preservation of the initial endowment if the number of cooperators meets or exceeds the collective target, and the third term reflects that if the number of cooperators does not meet the collective target, the initial endowment will be retained with probability $1 - r$. Similarly, Eq. (2) depicts the payoff for a defector, with the first term denoting the retention of the initial endowment if the number of cooperators meets or exceeds the collective target, and the second term indicating that if the number of cooperators fails to meet the collective target, the initial endowment will be retained with probability $1 - r$.

In an infinite well-mixed population, N individuals are randomly drawn to participate in the collective-risk social dilemma game. A replicator equation can be used to depict the evolution of individual strategy choices over time. Based on previous studies [24,45], the specific replicator equation can be written as

$$\dot{x} = x(1 - x)\left[\binom{N-1}{n_{pg}-1} x^{n_{pg}-1}(1 - x)^{N-n_{pg}} rb - c\right], \quad (3)$$

where $x(t)$ represents the proportion of individuals employing a cooperative strategy at time t, while $\dot{x}$ denotes the time derivative of $x(t)$, capturing the dynamical changes in the proportion of cooperators in the population over time.

Here, we define the dilemma strength in the collective-risk social dilemma game as $a = c/b$, which represents the ratio of the initial budget each cooperator contributes to the initial endowment, also called the "cost-endowment ratio." We know that a higher strength of the dilemma necessitates a larger contribution from each cooperator, given a constant initial endowment. In prior theoretical research, the strength of the dilemma was invariably constant, not subject to change with alterations in individual behavior decision-making within the group [24].

B. Feedback-evolving games

We assume that the strength of the dilemma is influenced by the state of the population. Specifically, an increase in the proportion of cooperators within the population gradually diminishes the dilemma strength, whereas an increase in the proportion of defectors progressively escalates the strength of the dilemma (see Fig. 1). For convenience, we consider a linear feedback form in which the strength of the dilemma decreases linearly with the rise in the proportion of cooperators within the group, and increases linearly with the rise in the proportion of defectors. Correspondingly, the dynamical equation for the dilemma strength can be written as

$$\dot{a} = (a - \alpha)(\beta - a)[(1 - x)u - x], \tag{4}$$

where α and β are the upper and lower limits of the dilemma strength, with $0 < \alpha,\ \beta < 1$, and $u(1 - x) - x$ represents the rate at which the dilemma strength is enhanced by the defection strategy at a rate u, and weakened by cooperation at a relative rate of 1.

Combining Eqs. (3) and (4), we are able to obtain the system equations as

$$\begin{aligned} \varepsilon\dot{x} &= x(1-x)\left[\binom{N-1}{n_{pg}-1}x^{n_{pg}-1}(1-x)^{N-n_{pg}}rb - ba\right],\\ \dot{a} &= (a-\alpha)(\beta-a)[(1-x)u - x], \end{aligned} \tag{5}$$

where $\varepsilon \ll 1$ signifies that the change in the strength of the dilemma is relatively slow compared to the change in individual behavior decisions since the dilemma strength as an environmental index often varies relatively slowly [39,45]. For the sake of convenience, we define Eq. (5) as system I. In Appendix A, we thoroughly analyze the equilibrium points of the system and examine the stability of these equilibrium points.

C. Institutional reward in feedback-evolving games

To better promote the emergence of cooperation, we further introduce institutional reward into the aforementioned feedback-evolving game model. We assume that each individual participating in the game will pay a tax of $b\delta$ before the game begins. The total tax revenue $Nb\delta$ is then evenly distributed to the cooperators at the end of the game, thereby increasing the payoff of each cooperator in the gaming group by $Nb\delta/(N_C + 1)$. In the presence of incentives, the difference between cooperators and defectors participating in the collective-risk social dilemma game can be written as $b\theta(N_C + 1 - n_{pg}) + b[1 - \theta(N_C + 1 - n_{pg})](1 - r) - c + \frac{Nb\delta}{N_C+1} - b\theta(N_C - n_{pg}) + b[1 - \theta(N_C - n_{pg})](1 - r)$. Through calculation, Eq. (3) can be rewritten as

$$\dot{x} = x(1-x)\left[\binom{N-1}{n_{pg}-1}x^{n_{pg}-1}(1-x)^{N-n_{pg}}rb + b\delta\frac{1-(1-x)^N}{x} - c\right].$$

Correspondingly, system (5) can be rewritten as follows:

$$\begin{aligned} \varepsilon\dot{x} &= x(1-x)\left[\binom{N-1}{n_{pg}-1}x^{n_{pg}-1}(1-x)^{N-n_{pg}}rb + b\delta\frac{1-(1-x)^N}{x} - ba\right],\\ \dot{a} &= (a-\alpha)(\beta-a)[(1-x)u - x]. \end{aligned} \tag{6}$$

For the convenience of subsequent analysis, we refer to Eq. (6) as system II. In the following section, we will specifically analyze the evolutionary dynamics generated by the coupled systems I and II. In Appendix B, we provide a detailed theoretical analysis.

To facilitate the comprehension of all parameters delineated in our work, we compile them in Table I.

TABLE I. Symbols and meanings used in this article.

Symbol	Meaning
N	Group size
n_{pg}	Collective target
r	Collective risk
b	Initial endowment
c	Cost of cooperation
a	Dilemma strength
α	Lower limit of dilemma strength
β	Upper limit of dilemma strength
u	Growth rate of dilemma strength with fraction of defectors
δ	Per capita incentive
ε	Feedback speed

III. RESULTS

A. Coevolutionary dynamics without incentives

We begin by investigating the coevolutionary dynamics of the coupled system I without considering any incentives. Through analysis, we find that system I can have up to nine equilibrium points. These consist of (i) four corner equilibrium points $(x, a) = (0, \alpha), (0, \beta), (1, \alpha), (1, \beta)$; (ii) four boundary equilibrium points $(x_1, \alpha), (x_2, \alpha), (x_3, \beta), (x_4, \beta)$, where x_1 and x_2 (with $x_2 > x_1$) are the two roots of the equation $\binom{N-1}{n_{pg}-1}x^{n_{pg}-1}(1-x)^{N-n_{pg}}r = \alpha$ [similarly, x_3 and x_4 (with $x_4 > x_3$) are the two roots of the equation $\binom{N-1}{n_{pg}-1}x^{n_{pg}-1}(1-x)^{N-n_{pg}}r = \beta$]; and (iii) one interior equilibrium point $(\frac{u}{1+u}, a^*)$, where $a^* = \binom{N-1}{n_{pg}-1}(\frac{u}{1+u})^{n_{pg}-1}(\frac{1}{1+u})^{N-n_{pg}}r$. For convenience, we set $F(x) = \binom{N-1}{n_{pg}-1}x^{n_{pg}-1}(1-x)^{N-n_{pg}}$. Based on the stability of the equilibrium points, we identify three representative types of evolutionary outcomes, namely tristable, bistable, and monostable states.

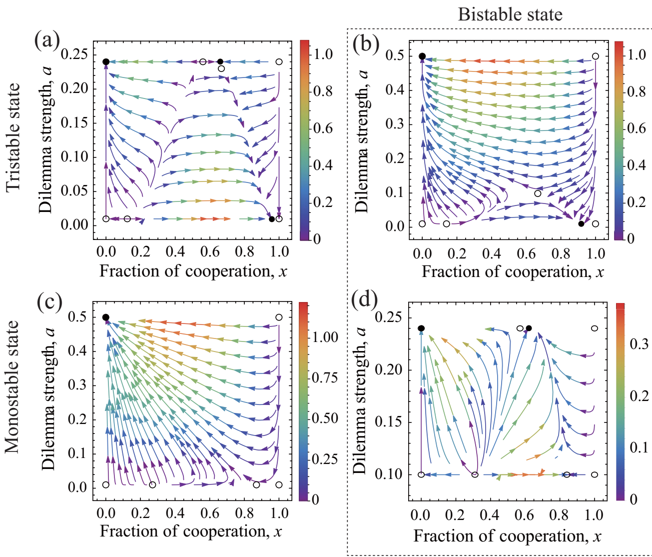

FIG. 2. Coevolutionary dynamics of the coupling system I. Panel (a) presents results of a tristable state, panels (b) and (d) exhibit bistable state outcomes, and panel (c) shows results of a monostable stable state. Solid dots represent stable equilibrium points, while empty circles denote unstable equilibrium points. Colored lines illustrate the evolution trajectories of the coupled system, with arrows indicating the direction of evolution. The color signifies the gradient value, where red represents the maximum gradient value and blue represents the minimum gradient value. Parameters are $N=6$, $n_{pg}=4$, $r=0.7$, $\varepsilon=0.1$, $b=1$, $\alpha=0.01$, $\beta=0.24$, and $u=2$ in panel (a); $N=6$, $n_{pg}=4$, $r=0.3$, $\varepsilon=0.1$, $b=1$, $\alpha=0.01$, $\beta=0.5$, and $u=2$ in panel (b); $N=6$, $n_{pg}=4$, $r=0.1$, $\varepsilon=0.1$, $b=1$, $\alpha=0.01$, $\beta=0.5$, and $u=8$ in panel (c); $N=6$, $n_{pg}=4$, $r=0.7$, $\varepsilon=0.1$, $b=1$, $\alpha=0.1$, $\beta=0.24$, and $u=25$ in panel (d).

1. Tristability

When $\max\{\frac{\alpha}{F(\frac{u}{1+u})}, \frac{\beta}{F(\frac{n_{pg}-1}{N-1})}\} < r < \frac{\beta}{F(\frac{u}{1+u})}$, we know that all nine equilibrium points exist in the system. In this case, if $\frac{x_4}{1-x_4} < u < \frac{x_2}{1-x_2}$, the coupled system I will generate a tristable state (more detailed theoretical analysis can be found in Appendix A). We further provide numerical examples to validate our theoretical results. As shown in Fig. 2(a), there are nine equilibrium points in the phase plane, among which $(0, \beta)$, (x_2, α), and (x_4, β) are stable. In the phase plane, the vast majority of trajectories converge to (x_2, α), indicating that cooperators can be maintained in the population, accompanied by the lowest dilemma strength. A portion of the trajectories converge to $(0, \beta)$, implying the disappearance of cooperators and the dilemma strength reaching its maximum value. The remaining few trajectories converge to (x_4, β), depicting that cooperation can still be sustained under the highest dilemma strength.

2. Bistability

When all nine equilibrium points exist and if $u < \min\{\frac{x_2}{1-x_2}, \frac{x_4}{1-x_4}\}$, the coupled system I will exhibit a bistable outcome, characterized by the presence of two stable states, namely (x_2, α) and $(0, \beta)$. Furthermore, when the risk $r < \frac{\beta}{F(\frac{n_{pg}-1}{N-1})}$, the boundary equilibrium points (x_3, β) and (x_4, β) do not exist (a detailed theoretical analysis is provided in Appendix A). To verify the theoretical analysis, we provide a numerical example. As shown in Fig. 2(b), the phase plane exhibits seven equilibrium points, including four corner equilibrium points, one interior equilibrium point, and two boundary equilibrium points (x_1, α) and (x_2, α). The corner equilibrium point $(0, \beta)$ and boundary equilibrium point (x_2, α) are stable. Depending on the initial conditions, some trajectories converge to the boundary equilibrium point, while the remaining ones converge to the corner equilibrium point.

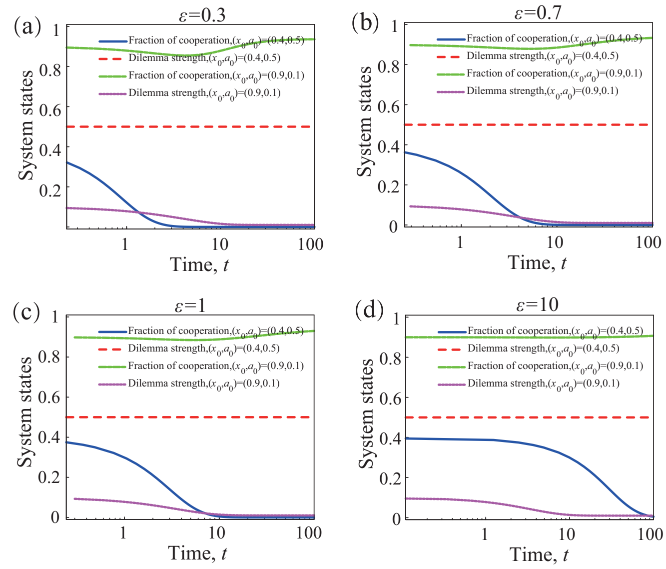

FIG. 3. The evolutionary outcomes of coupled systems I for different feedback speed ε. Parameters are $N=6$, $n_{pg}=4$, $r=0.3$, $b=1$, $\alpha=0.1$, $u=0.08$, and $\beta=0.5$.

Figure 2(b) illustrates a bistable outcome. The equilibrium point positioned in the upper left corner delineates a scenario characterized by a high strength of the dilemma, wherein the choices of all individuals converge towards the defection strategy, representing the most unfavorable state. In contrast, the boundary equilibrium point located in the lower right corner characterizes a scenario marked by a low strength of the dilemma, where the majority of individuals converge towards the cooperate strategy, exemplifying a relatively advantageous state. Nevertheless, it is important to note that the basin of attraction of the boundary equilibrium point in the lower right corner is significantly smaller than that of the upper left corner equilibrium point. This implies that achieving a favorable state has specific requirements regarding the initial proportion of cooperators and the strength of the dilemma. Furthermore, our results are not affected by the feedback speed (see Fig. 3).

Based on theoretical analysis, we know that the corner equilibrium point $(0, \beta)$ is always stable. When the coupled system exhibits multiple stable states, it is crucial to maximize the basin of attraction of the more favorable state. In light of this, we investigate the variation of the basin of attraction of the equilibrium point (x_2, α) as the model parameters change. We observe that an increase in the risk of collective task

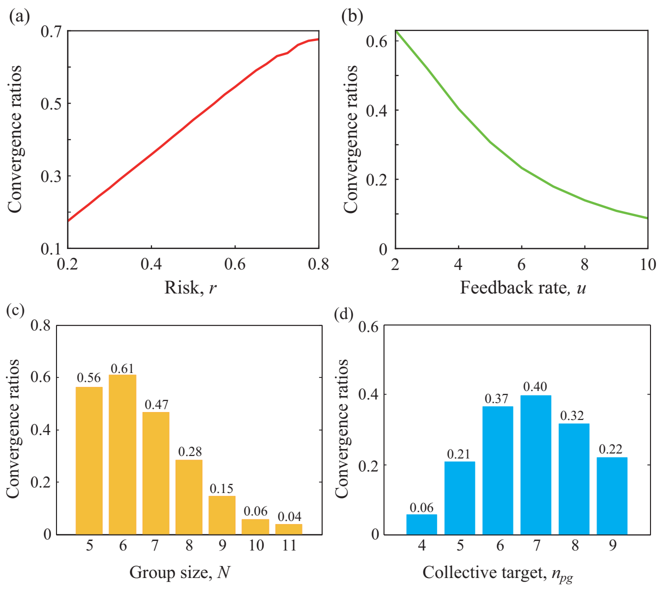

FIG. 4. The basin of attraction of the equilibrium point (x_2, α) in dependence of the risk value r (a), the feedback rate u (b), the group size N (c), and the collective target n_{pg} (d). The Monte Carlo method is used to calculate the convergence rate. Specifically, in the phase plane, we randomly and uniformly select a large number of initial points. After the system evolves to a steady state, we calculate the proportion of initial points that converge to the point (x_2, α). This proportion represents the basin of attraction for (x_2, α). Parameters are $N = 6,\ n_{pg} = 4,\ \varepsilon = 0.1,\ b = 1,\ \alpha = 0.01,\ \beta = 0.24$, and $u = 2$ in panel (a); $N = 6,\ n_{pg} = 4,\ r = 0.7,\ \varepsilon = 0.1,\ b = 1,\ \alpha = 0.01$, and $\beta = 0.24$ in panel (b); $n_{pg} = 4,\ r = 0.7,\ \varepsilon = 0.1,\ b = 1,\ \alpha = 0.01,\ \beta = 0.24$, and $u = 2$ in panel (c); $N = 10,\ r = 0.7,\ \varepsilon = 0.1,\ b = 1,\ \alpha = 0.01,\ \beta = 0.24$, and $u = 2$ in panel (d).

failure expands the basin of attraction of (x_2, α) [see Fig. 4(a)]. This implies that an elevation in risk better stimulates individuals to make cooperative decisions, thereby reducing the strength of the dilemma. In addition, the basin of attraction of this stable state decreases with an increase in the values of u (rate of escalation of the dilemma strength inducing by defection) [see Fig. 4(b)]. This implies that when the proportion of defectors has a greater impact on the dilemma strength, the willingness of individuals to cooperate is significantly reduced. The impact of the group size on the basin of attraction for the equilibrium point (x_2, α) is presented in Fig. 4(c). We find that for different group size values, the basin of attraction for this stable equilibrium point first increases and then decreases. This nonlinear phenomenon implies that an intermediate group size can best promote the emergence of cooperation and thereby maintain a lower strength of the dilemma. Too high or too low a group size cannot achieve this effect. We further present the impact of different collective goal values on the basin of attraction for the stable equilibrium point (x_2, α), as depicted in Fig. 4(d). We observe a nonlinear effect of the collective goal on the basin of attraction for this equilibrium point. That is, with the increase in collective goal values, the basin of attraction initially enlarges, reaches a maximum, and then gradually decreases. This phenomenon subtly suggests that an intermediate collective goal can better facilitate the system's stability in this more favorable state.

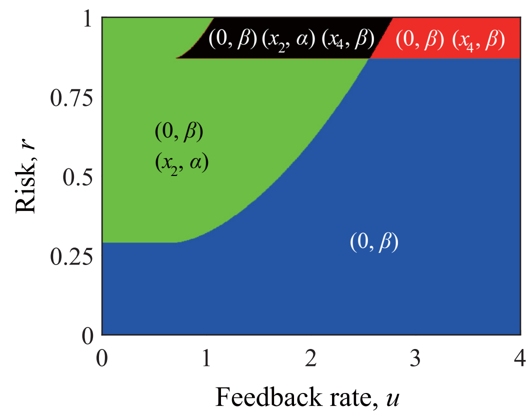

FIG. 5. The evolutionary outcomes of coupled system I, differentiated by various colors, corresponding to different feedback rate u and risk r. Parameters are $N = 6,\ n_{pg} = 3,\ \varepsilon = 0.1,\ b = 1,\ \alpha = 0.1$, and $\beta = 0.3$.

Moreover, we present another bistable outcome of the coupled system I in Fig. 2(d). When $r > \frac{\beta}{F(\frac{n_{pg}-1}{N-1})}$, we know that all four boundary equilibrium points exist. Then if $u > \frac{x_2}{1-x_2}$, the boundary fixed point (x_2, α) is unstable, while (x_4, β) is stable. Therefore, the coupled system exhibits two stable equilibrium points, namely $(0,\ \beta)$ and (x_4, β). The former remains the worst outcome for the coupled system: cooperation cannot occur, and the strength of the dilemma reaches its maximum. The latter avoids the disappearance of cooperation, but the strength of the dilemma still remains at the highest level, indicating that the cooperators need to bear a high cost of cooperation.

3. Monostability

In the coupled system I, the corner equilibrium point $(0,\ \beta)$ is always stable. Ultimately, we present a monostable outcome [see Fig. 2(c)]. When $\frac{\alpha}{F(\frac{n_{pg}-1}{N-1})} < r < \frac{\beta}{F(\frac{n_{pg}-1}{N-1})}$ and $u > \frac{x_2}{1-x_2}$, the system has six equilibrium points: four corner equilibrium points and two lower boundary equilibrium points (x_1, α) and (x_2, α), among which only $(0,\ \beta)$ is stable. This implies that the strength of the dilemma reaches its peak, with no individual willing to cooperate.

Now we present the evolutionary dynamics of system I under various parameter conditions. To provide readers with a clearer understanding of the system's evolutionary stable states across different parameter regions, we depict the evolutionary outcomes of the coupled system I in Fig. 5 where diverse colors are employed to illustrate the distinct evolutionary results.

Despite the potential for a variety of dynamic outcomes, two unavoidable conundrums persistently emerge. The first is that the optimal state of the coupled system can never stabilize, that is, all individuals choose to cooperate, reducing the strength of the dilemma to its lowest point. The second is that the worst state of the coupled system is always stable, that is, all individuals choose to defect, pushing the strength of the dilemma to its maximum value.

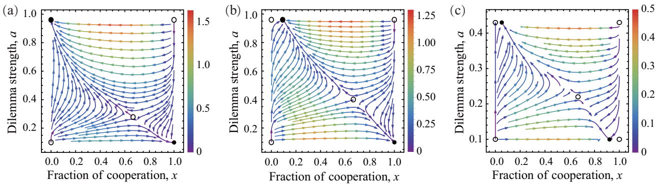

FIG. 6. Coupled systems with institutional rewards exhibit bistability. Parameters are $N = 6$, $n_{pg} = 4$, $r = 0.3$, $\varepsilon = 0.1$, $b = 1$, $\alpha = 0.1$, $\beta = 0.96$, $\delta = 0.12$, and $u = 2$ in panel (a); $N = 6$, $n_{pg} = 4$, $r = 0.3$, $\varepsilon = 0.1$, $b = 1$, $\alpha = 0.1$, $\delta = 0.2$, $u = 2$, and $\beta = 0.96$ in panel (b); $N = 6$, $n_{pg} = 4$, $r = 0.3$, $\varepsilon = 0.1$, $b = 1$, $\alpha = 0.1$, $u = 2$, $\delta = 0.08$, and $\beta = 0.43$ in panel (c).

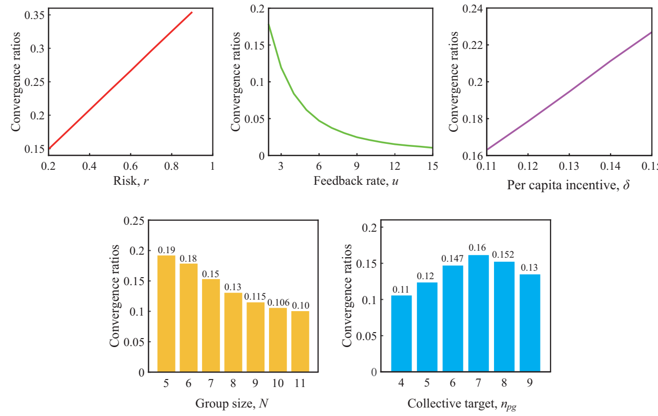

FIG. 7. The basin of attraction of the equilibrium point $(1, \alpha)$ varies with the changes in risk (r), feedback rate (u), per capita incentive (δ), group size (N), and collective target (n_{pg}). Parameters are $N = 6$, $n_{pg} = 4$, $\varepsilon = 0.1$, $b = 1$, $\alpha = 0.1$, $\beta = 0.96$, $\delta = 0.12$, and $u = 2$ in the first panel of the top row; $N = 6$, $n_{pg} = 4$, $r = 0.3$, $\varepsilon = 0.1$, $b = 1$, $\alpha = 0.1$, $\delta = 0.12$, and $\beta = 0.96$ in the second panel of the top row; $N = 6$, $n_{pg} = 4$, $r = 0.3$, $\varepsilon = 0.1$, $b = 1$, $\alpha = 0.1$, $u = 2$, and $\beta = 0.96$ in the third panel of the top row; $n_{pg} = 4$, $r = 0.3$, $\varepsilon = 0.1$, $b = 1$, $\alpha = 0.1$, $\beta = 0.96$, and $u = 2$ in the first panel of the bottom row; $N = 10$, $r = 0.3$, $\varepsilon = 0.1$, $b = 1$, $\alpha = 0.1$, $\beta = 0.96$, and $u = 2$ in the second panel of the bottom row.

B. Coevolutionary dynamics with institutional reward

To address the aforementioned two issues inherent in a coupled system, we introduce institutional reward. Through the analysis of the coupled system II, we find that the four corner equilibrium points $(0, \alpha)$, $(0, \beta)$, $(1, \alpha)$, $(1, \beta)$ and the interior equilibrium point $(\frac{u}{1+u}, \bar{a}^*)$ still exist, where $\bar{a}^* = \binom{N-1}{n_{pg}-1}(\frac{u}{1+u})^{n_{pg}-1}(\frac{1}{1+u})^{N-n_{pg}} r + \delta\sum_{i=0}^{N-1}(\frac{1}{u+1})^i$. Additionally, the boundary equilibrium points (x^*, α) and $(\bar{x}^*, \beta)$ are determined by the equations $\binom{N-1}{n_{pg}-1}(x^*)^{n_{pg}-1}(1-x^*)^{N-n_{pg}} r + \delta\sum_{i=0}^{N-1}(x^*)^i = \alpha$ and $\binom{N-1}{n_{pg}-1}(\bar{x}^*)^{n_{pg}-1}(1-\bar{x}^*)^{N-n_{pg}} r + \delta\sum_{i=0}^{N-1}(\bar{x}^*)^i = \beta$, respectively. Due to the complexity of these two higher-order equations, it is challenging to theoretically analyze the distribution and quantity of the boundary equilibrium points. Below, we numerically study the evolutionary dynamics of the system based on the scenarios where the system can reach its optimal state or escape from its worst state.

When $\delta > \alpha$, the optimal state $(1, \alpha)$ of the coupled system II is stable, implying that all individuals choose cooperative behavior, accompanied by the lowest dilemma strength. At this time, all individuals only need to pay the lowest cooperation cost of $b\alpha$. Furthermore, when $\delta < \frac{\beta}{N}$, the worst state $(0, \beta)$ of the coupled system can also achieve stability, implying that all individuals choose to defect, accompanied by the highest dilemma strength. In Fig. 6(a), we provide a numerical example that satisfies the conditions mentioned above. We find that the phase plane exhibits bistability, that is, depending on the initial conditions, part of the system trajectories converge to the optimal state, while the vast majority of the remaining trajectories converge to the worst state.

When $\delta > \max\{\alpha, \frac{\beta}{N}\}$, the best equilibrium state of the system is stable while the worst state is unstable. As shown in Fig. 6(b), the corner equilibrium point $(1, \alpha)$ in the phase plane is stable, whereas the corner equilibrium point $(0, \beta)$ is unstable. A stable equilibrium point exists on the upper boundary of the phase plane, which signifies the presence of a minority of individuals choosing cooperative behavior, accompanied by the highest dilemma strength. It should be noted that these few cooperators need to pay higher cooperation cost $b\beta$.

When $\frac{\beta}{N} < \delta < \alpha$, both the best state $(1, \alpha)$ and the worst state $(0, \beta)$ of the coupled system II are unstable. In Fig. 6(c), we provide a specific numerical example in which there exist seven equilibrium points in the phase plane, including four vertex equilibrium points, one interior equilibrium point, and two boundary equilibrium points. The two boundary equilibrium points are stable, where the upper boundary equilibrium point suggests that only a minority of individuals opt for cooperative behavior, reaching the highest strength of the dilemma, and the lower boundary equilibrium point implies that the vast majority of individuals choose to cooperate, accompanied by the lowest strength of the dilemma, indicating that cooperators only need to pay a lower cost of cooperation ($b\alpha$).

When the coupled system II is capable of exhibiting multistable states, the optimal state $(1, \alpha)$ can reach stability. We further provide results of how the basin of attraction for this optimal state changes with model parameters. As shown in Fig. 7, an increase in risk values or per capita incentive values can expand the basin of attraction for the optimal state, while the growth of feedback rates and group size can shrink it. Additionally, we observe that the influence of collective goals on the optimal state's basin of attraction is nonlinear, i.e., with the increase of the goal value, the basin of attraction first increases and then decreases, a phenomenon consistent with results under scenarios without incentives.

Based on the stability of corner equilibrium points, we further provide two scenarios of monostability. When $\delta < \min\{\frac{\beta}{N}, \alpha\}$, the corner equilibrium point $(0, \beta)$ is stable, while the other three corner equilibrium points are unstable. Moreover, when the parameters do not satisfy $\alpha < \bar{a}^* < \beta$, the interior equilibrium point does not exist. In Fig. 8(a), we provide a specific numerical example where we find only a single stable point in the phase plane, located in the upper left corner. This implies the occurrence of the worst-case scenario in the system, where no individuals are willing to bear the cost of cooperation, and the strength of the dilemma reaches

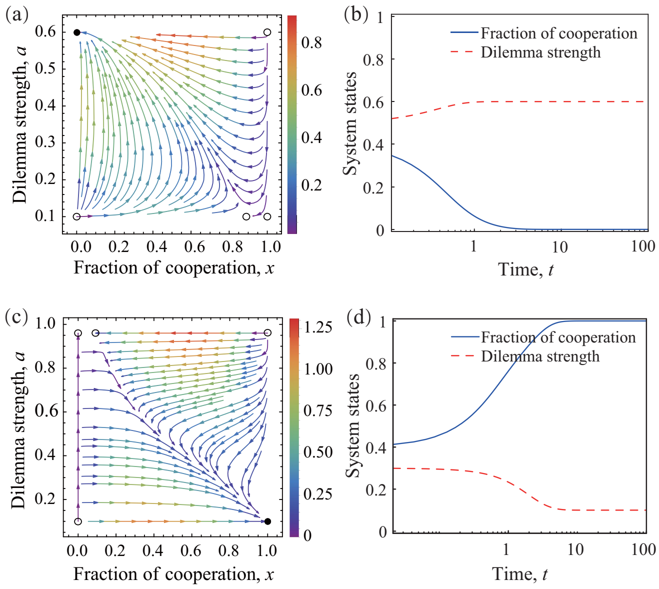

FIG. 8. Two representative numerical examples when the coupled system exhibits monostability. Parameters are $N = 6,\ n_{pg} = 4,\ r = 0.3,\ \varepsilon = 0.1,\ b = 1,\ \alpha = 0.1,\ \beta = 0.6,\ \delta = 0.08$, and $u = 10$ in panels (a) and (b); $N = 6,\ n_{pg} = 4,\ r = 0.3,\ \varepsilon = 0.1,\ b = 1,\ \alpha = 0.1,\ \beta = 0.96,\ \delta = 0.2$, and $u = 0.08$ in panels (c) and (d).

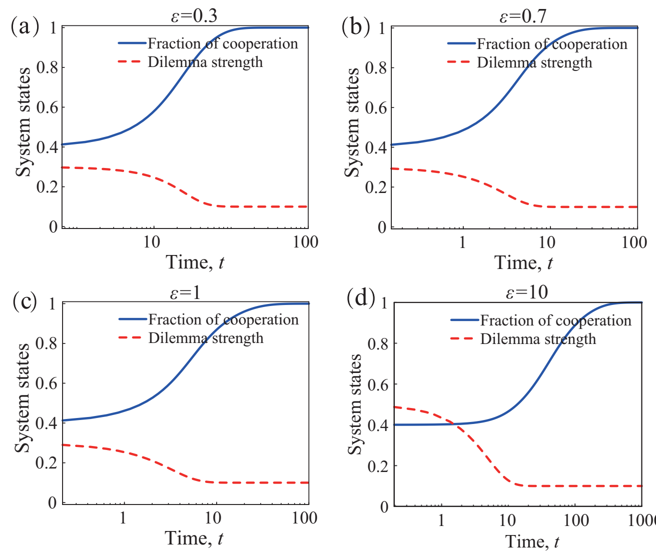

FIG. 9. The evolutionary outcomes of coupled system II for different feedback speed ε. Parameters are $N = 6,\ n_{pg} = 4,\ r = 0.3,\ b = 1,\ \alpha = 0.1,\ \delta = 0.2,\ u = 0.08$, and $\beta = 0.96$.

its maximum. In Fig. 8(c), we present another result of a single stable state. We find that when the model parameters satisfy $\delta > \max\{\alpha, \frac{\beta}{N}\}$, the corner equilibrium point $(1, \alpha)$ is stable, while the other three corner equilibrium points are unstable. Moreover, our numerical results reveal the absence of an interior equilibrium point, with an unstable boundary equilibrium point existing on the upper border. In this scenario, the best state of the system is the sole stable state, implying that all individuals are willing to bear the minimum cost of cooperation to achieve the collective goal. This result is also unaffected by the feedback speed (see Fig. 9).

IV. DISCUSSION

In complex social systems, individual behaviors are often influenced by the surrounding environment, such as the abundance of resources and the risk level of the game environment. Previous theoretical research has revealed that individuals are more inclined to choose defection strategy in resource-rich environments [39] and cooperation in high-risk environments [24]. At the same time, collective behavior can also change the surrounding environment. For example, excessive defectors can lead to resource depletion and increase the risk of disaster. In recent years, coevolutionary game theory has provided a powerful theoretical tool for studying such strategy-environment coupled systems [39]. Along these lines, in this work we have constructed a feedback-evolving game model to capture the coupling relationship between strategy and dilemma strength. We have considered a form of feedback where strategies have linear effects on the dilemma strength, that is, an increase in the proportion of defectors in the population linearly increases the dilemma strength, while an increase in the proportion of cooperators linearly decreases the dilemma strength. We have identified intriguing coevolutionary dynamics, including tristable and bistable states, for which we have provided the analytical conditions and performed numerical validations. Our results illustrate that cooperation can persist, accompanied by either the highest or the lowest dilemma strength. However, under the highest dilemma strength, the state of full defection is always stable. This indicates that the optimal state of this coupled system is unattainable, whereas the least desirable state remains stable.

Although the system's optimal state is unattainable, a suboptimal state can be achieved. In this suboptimal state, cooperation can be maintained at a certain level, while the strength of the dilemma can be kept at a minimum. This implies that the cost incurred by the cooperators is minimized ($c = b\alpha$). Since the worst state is always stable, it is worth exploring how to expand the basin of attraction of the suboptimal state when it reaches stability. Our results reveal that increasing the risk r or reducing the feedback rate u can expand the basin of attraction of this suboptimal state. Furthermore, an intriguing finding is that the influence of group size and collective target on the basin of attraction of the suboptimal state is nonlinear. Specifically, an intermediate group size or an intermediate collective target value can more effectively facilitate the evolution of cooperation. However, previous theoretical research [24] has demonstrated that decision-making within small groups can significantly enhance the likelihood of coordinating actions and avoiding the tragedy of the commons. Furthermore, it is also unveiled that a larger collective target is conducive to the emergence of high-level cooperation. The introduction of dynamic feedback between strategy and dilemma strength primarily contributes to this inconsistency phenomenon.

Building on the understanding that the system's worst state remains persistently stable, a surprising revelation from our research is that the introduction of feedback mechanisms does not categorically facilitate the resolution of cooperation dilemmas. Despite the high risks of collective failure, our model uncovers a stubbornly stable worst-case scenario $(0, \beta)$ (see Fig. 5). This observation is especially intriguing in light of the substantial body of literature that suggests elevated risk levels can effectively enhance cooperation. Consequently, our research highlights the need for a more nuanced comprehension of the dynamics involved. It proposes that a comprehensive approach should be adopted, one that extends beyond focusing solely on the inherent risks of the game, but also takes into account the complex interplay between group decision-making processes and the gaming environment.

The implementation of incentives is a crucial measure in regulating collective actions. To modulate the aforementioned coupled system, we have introduced institutional reward with taxation. We have demonstrated that the system can reach its optimal state, where all individuals choose to cooperate, accompanied by the lowest dilemma strength. Simultaneously, the worst-case scenario, where all individuals choose to defect, accompanied by the highest dilemma strength, can be avoided. We theoretically demonstrate that the occurrence of this state hinges on strong incentives ($\delta > \max\{\alpha, \frac{\beta}{N}\}$). Additionally, we have identified other bistable dynamics where, when the system's optimal state is stable, we have examined the variation of its basin of attraction with model parameters. Our results indicate that an increase in risk r and incentive value δ both individually enhance the basin of attraction of the optimal state, while an increase in group size N and feedback rate u both exert a suppressive effect. The influence of the collective target n_{pg} on this optimal state's basin of attraction manifests as a nonlinear phenomenon, consistent with conclusions drawn in previous scenarios without incentive.

In this study, we have presented a rudimentary feedback-evolving game model to illustrate the interconnectedness between individual behavioral decisions and the state of the gaming environment. Specifically, the model suggests that an upsurge in the ratio of cooperators within the group linearly mitigates the dilemma's strength, whereas an increase in defectors proportionally exacerbates it. A logical progression of this work would be to explore the impact of nonlinear feedback mechanisms. This is primarily due to the fact that in real-world scenarios, individuals' responses to environmental shifts are not always incremental; they can often be sudden and dramatic. For instance, the expression $-x^p + u(1-x)^p$, where p is greater than 1, could be considered to model such nonlinear dynamics.

While the system's optimal state can be stabilized and the worst-case scenario circumvented, achieving this outcome will necessitate substantial per capita incentives. This raises an intriguing question: can we devise alternative incentive schemes that yield the same evolutionary trajectory, but demand considerably fewer per capita incentives? Moreover, alterations in the population's state have a dual effect: they not only modulate the severity of the dilemma, but they also influence the risk. Consequently, understanding how cooperative behavior manifests in a milieu with multiple concurrent feedback mechanisms is a promising avenue for future research. Integrating measures of social capital (e.g., trust, volunteering), inequality (e.g., Gini coefficient), and risk aversion into this analysis will further elucidate the dynamics of individual versus collective benefits, thereby enriching our comprehension of investment dilemmas within group settings. Furthermore, the impact of population size, such as more realistic finite populations [53,54], and different network structures, such as regular graphs and small-world networks, represent a direction worthy of in-depth exploration in future research.

ACKNOWLEDGMENTS

This research was supported by the National Natural Science Foundation of China (Grants No. 62306243, No. 62036002, and No. 61976048), the Natural Science Foundation of Shaanxi (Grants No. 2024JC-YBQN-0055, No. 2023-JC-QN-0791, 23JK0693, and No. 23JK0692), and the Sichuan Science and Technology Program (Grant No. 2024NSFSC0436).

APPENDIX A: COEVOLUTIONARY DYNAMICS WITHOUT INCENTIVES

We first investigate the evolutionary dynamics of the coupled system when no additional incentives are considered. Correspondingly, the coupled system incorporating strategy-dilemma strength feedback can be depicted by the following equations:

$$\varepsilon\dot{x} = x(1-x)\left[\binom{N-1}{n_{pg}-1}x^{n_{pg}-1}(1-x)^{N-n_{pg}}rb - ba\right],$$
$$\dot{a} = (a-\alpha)(\beta-a)[(1-x)u - x].$$

By solving the aforementioned coupled system, we can obtain all possible equilibrium points of the system, which are $(x, a) = (0, \alpha)$, $(0, \beta)$, $(1, \alpha)$, $(1, \beta)$, (x_1, α), (x_2, α), (x_3, β), (x_4, β), and $(\frac{u}{1+u}, a^*)$, where x_1 and x_2 are the two roots of the equation $\binom{N-1}{n_{pg}-1}x^{n_{pg}-1}(1-x)^{N-n_{pg}}r = \alpha$, x_3 and x_4 are the two roots of the equation $\binom{N-1}{n_{pg}-1}x^{n_{pg}-1}(1-x)^{N-n_{pg}}r = \beta$, and $a^* = \binom{N-1}{n_{pg}-1}(\frac{u}{1+u})^{n_{pg}-1}(\frac{1}{1+u})^{N-n_{pg}}r$. Then we determine the stability by analyzing the signs of the eigenvalues of the Jacobian matrix at each equilibrium point. Accordingly, the detailed analysis of all possible evolutionary outcomes is presented below.

Case 1. When $\max\{\frac{\alpha}{F(\frac{u}{1+u})}, \frac{\beta}{F(\frac{n_{pg}-1}{N-1})}\} < r < \frac{\beta}{F(\frac{u}{1+u})}$, all nine equilibrium points exist in the system.

For $(x, a) = (0, \alpha)$, the Jacobian matrix is

$$J = \begin{bmatrix} \frac{-b\alpha}{\varepsilon} & 0 \\ 0 & u(\beta-\alpha) \end{bmatrix},$$

thus the fixed equilibrium is unstable.

For $(x, a) = (0, \beta)$, the Jacobian matrix is

$$J = \begin{bmatrix} \frac{-b\beta}{\varepsilon} & 0 \\ 0 & u(\alpha-\beta) \end{bmatrix},$$

thus the fixed equilibrium is stable.

For $(x, a) = (1, \alpha)$, the Jacobian matrix is

$$J = \begin{bmatrix} \frac{b\alpha}{\varepsilon} & 0 \\ 0 & \alpha-\beta \end{bmatrix},$$

thus the fixed equilibrium is unstable.

For $(x, a) = (1, \beta)$, the Jacobian matrix is

$$J = \begin{bmatrix} \frac{b\beta}{\varepsilon} & 0 \\ 0 & \beta - \alpha \end{bmatrix},$$

thus the fixed equilibrium is unstable.

For $(x, a) = (x_1, \alpha)$, the Jacobian matrix is

$$J = \begin{bmatrix} \frac{b\alpha}{\varepsilon}[n_{pg} - 1 - (N-1)x_1] & -\frac{x_1(1-x_1)b}{\varepsilon} \\ 0 & (\beta - \alpha)(u - ux_1 - x_1) \end{bmatrix},$$

thus the fixed equilibrium is unstable since $x_1 < \frac{n_{pg}-1}{N-1}$.

For $(x, a) = (x_2, \alpha)$, the Jacobian matrix is

$$J = \begin{bmatrix} \frac{b\alpha}{\varepsilon}[n_{pg} - 1 - (N-1)x_2] & -\frac{x_2(1-x_2)b}{\varepsilon} \\ 0 & (\beta - \alpha)(u - ux_2 - x_2) \end{bmatrix},$$

thus the fixed equilibrium is stable when $u < \frac{x_2}{1-x_2}$.

For $(x, a) = (x_3, \beta)$, the Jacobian matrix is

$$J = \begin{bmatrix} \frac{b\beta}{\varepsilon}[n_{pg} - 1 - (N-1)x_3] & -\frac{x_3(1-x_3)b}{\varepsilon} \\ 0 & (\alpha - \beta)(u - ux_3 - x_3) \end{bmatrix},$$

thus the fixed equilibrium is unstable since $x_3 < \frac{n_{pg}-1}{N-1}$.

For $(x, a) = (x_4, \beta)$, the Jacobian matrix is

$$J = \begin{bmatrix} \frac{b\beta}{\varepsilon}[n_{pg} - 1 - (N-1)x_4] & -\frac{x_4(1-x_4)b}{\varepsilon} \\ 0 & (\alpha - \beta)(u - ux_4 - x_4) \end{bmatrix},$$

thus the fixed equilibrium is stable when $u > \frac{x_4}{1-x_4}$.

For $(x, a) = (\frac{u}{1+u}, a^*)$, the Jacobian matrix is

$$J = \begin{bmatrix} \frac{ba^*}{\varepsilon}\left[n_{pg} - 1 - \frac{(N-1)u}{1+u}\right] & -\frac{ub}{(1+u)^2\varepsilon} \\ (a^* - \alpha)(\beta - a^*)(-u-1) & 0 \end{bmatrix},$$

thus the fixed equilibrium is unstable.

In summary, we can derive the following two conclusions:

(i) If $\frac{x_4}{1-x_4} < u < \frac{x_2}{1-x_2}$, the coupled system I has three stable points: $(0, \beta)$, (x_2, α), and (x_4, β).

(ii) If $\frac{x_1}{1-x_1} < u < \frac{x_3}{1-x_3}$, the coupled system I has two stable points: $(0, \beta)$ and (x_2, α).

Case 2. When $\frac{\beta}{F(\frac{n_{pg}-1}{N-1})} < r < \frac{\alpha}{F(\frac{u}{1+u})}$ or $r > \frac{\beta}{F(\frac{u}{1+u})}$, apart from the nonexistence of an interior equilibrium point, the other eight equilibrium points indeed exist.

Based on the stability of equilibrium points, we summarize the following three conclusions:

(i) If $r < \frac{\alpha}{F(\frac{u}{1+u})}$ and $u > \frac{x_2}{1-x_2}$, the coupled system I has two stable points: $(0, \beta)$ and (x_4, β).

(ii) If $r < \frac{\alpha}{F(\frac{u}{1+u})}$ and $u < \frac{x_1}{1-x_1}$, the coupled system I has two stable points: $(0, \beta)$ and (x_2, α).

(iii) If $r > \frac{\beta}{F(\frac{u}{1+u})}$ and $\frac{x_3}{1-x_3} < u < \frac{x_4}{1-x_4}$, the coupled system I has two stable points: $(0, \beta)$ and (x_2, α).

Case 3. When $\max\{\frac{\alpha}{F(\frac{n_{pg}-1}{N-1})}, \frac{\alpha}{F(\frac{u}{1+u})}\} < r < \frac{\beta}{F(\frac{n_{pg}-1}{N-1})}$, except for (x_3, β) and (x_4, β), the other seven equilibrium points all exist. According to the stability analysis, the coupled system I has two stable points: $(0, \beta)$ and (x_2, α).

Case 4. When $\frac{\alpha}{F(\frac{n_{pg}-1}{N-1})} < r < \min\{\frac{\beta}{F(\frac{n_{pg}-1}{N-1})}, \frac{\alpha}{F(\frac{u}{1+u})}\}$, except for (x_3, β), (x_4, β), and $(\frac{u}{1+u}, a^*)$, the other six equilibrium points exist.

Based on the stability of equilibrium points, we can identify the following two conclusions:

(i) If $u < \frac{x_1}{1-x_1}$, the coupled system I has two stable points: $(0, \beta)$ and (x_2, α).

(ii) If $u > \frac{x_2}{1-x_2}$, the coupled system I has one stable point: $(0, \beta)$.

Case 5. When $r < \frac{\alpha}{F(\frac{n_{pg}-1}{N-1})}$, only four corner equilibrium points exist. According to the stability analysis, the coupled system has one stable equilibrium point: $(0, \beta)$.

In the main text, we provide numerical validation based on these theoretical predictions.

APPENDIX B: COEVOLUTIONARY DYNAMICS WITH INSTITUTIONAL REWARD

When considering additional institutional reward, our coupled system can be reformulated as

$$\varepsilon\dot{x} = x(1-x)\left[\binom{N-1}{n_{pg}-1}x^{n_{pg}-1}(1-x)^{N-n_{pg}}rb + b\delta\frac{1-(1-x)^N}{x} - ba\right],$$
$$\dot{a} = (a-\alpha)(\beta-a)[(1-x)u - x].$$

Through analysis, we ascertain that the system invariably exhibits four corner equilibria: $(x, a) = (0, \alpha), (0, \beta), (1, \alpha)$, and $(1, \beta)$. An interior equilibrium $(\frac{u}{1+u}, \bar{a}^*)$ is present when $\bar{a}^*$ satisfies the equation $\binom{N-1}{n_{pg}-1}(\frac{u}{1+u})^{n_{pg}-1}(\frac{1}{1+u})^{N-n_{pg}}r + \delta\sum_{i=0}^{N-1}(\frac{1}{u+1})^i = \bar{a}^*$ and $\alpha < \bar{a}^* < \beta$. Additionally, the boundary equilibrium points (x^*, α) and $(\bar{x}^*, \beta)$ are determined by the equations $\binom{N-1}{n_{pg}-1}(x^*)^{n_{pg}-1}(1-x^*)^{N-n_{pg}}r + \delta\sum_{i=0}^{N-1}(x^*)^i = \alpha$ and $\binom{N-1}{n_{pg}-1}(\bar{x}^*)^{n_{pg}-1}(1-\bar{x}^*)^{N-n_{pg}}r + \delta\sum_{i=0}^{N-1}(\bar{x}^*)^i = \beta$, respectively. It is extremely challenging to analytically investigate the number of boundary equilibria due to the high complexity of the aforementioned high-order nonlinear equations. Next, we present the Jacobian matrix of the aforementioned equilibrium points:

For $(x, a) = (0, \alpha)$, the Jacobian matrix is

$$J = \begin{bmatrix} \frac{bN\delta - b\alpha}{\varepsilon} & 0 \\ 0 & u(\beta - \alpha) \end{bmatrix},$$

thus the fixed equilibrium is unstable.

For $(x, a) = (0, \beta)$, the Jacobian matrix is

$$J = \begin{bmatrix} \frac{bN\delta - b\beta}{\varepsilon} & 0 \\ 0 & u(\alpha - \beta) \end{bmatrix},$$

thus the fixed equilibrium is stable when $N\delta < \beta$.

For $(x, a) = (1, \alpha)$, the Jacobian matrix is

$$J = \begin{bmatrix} \frac{b\alpha - b\delta}{\varepsilon} & 0 \\ 0 & \alpha - \beta \end{bmatrix},$$

thus the fixed equilibrium is stable when $\delta > \alpha$.

For $(x, a) = (1, \beta)$, the Jacobian matrix is

$$J = \begin{bmatrix} \frac{b\beta - b\delta}{\varepsilon} & 0 \\ 0 & \beta - \alpha \end{bmatrix},$$

thus the fixed equilibrium is unstable.

For $(x, a) = (x^*, \alpha)$, the Jacobian matrix is

$$J = \begin{bmatrix} a_{11} & -\frac{x^*(1-x^*)b}{\varepsilon} \\ 0 & (\beta - \alpha)(u - ux^* - x^*) \end{bmatrix},$$

where $a_{11} = \frac{1}{\varepsilon}\{\binom{N-1}{n_{pg}-1}(x^*)^{n_{pg}-1}(1-x^*)^{N-n_{pg}} rb[n_{pg} - 1 - (N-1)x^*] - b\delta x^* \sum_{i=1}^{N-1} i(1-x^*)^i\}$, thus the fixed equilibrium is stable when $a_{11} < 0$ and $u < \frac{x^*}{1-x^*}$.

For $(x, a) = (\bar{x}^*, \beta)$, the Jacobian matrix is

$$J = \begin{bmatrix} a_{11} & -\frac{\bar{x}^*(1-\bar{x}^*)b}{\varepsilon} \\ 0 & (\alpha - \beta)(u - u\bar{x}^* - \bar{x}^*) \end{bmatrix},$$

where $a_{11} = \frac{1}{\varepsilon}\{\binom{N-1}{n_{pg}-1}(\bar{x}^*)^{n_{pg}-1}(1-\bar{x}^*)^{N-n_{pg}} rb[n_{pg} - 1 - (N-1)\bar{x}^*] - b\delta \bar{x}^* \sum_{i=1}^{N-1} i(1-\bar{x}^*)^i\}$, thus the fixed equilibrium is stable when $a_{11} < 0$ and $u > \frac{\bar{x}^*}{1-\bar{x}^*}$.

For $(x, a) = (\frac{u}{1+u}, \bar{a}^*)$, the Jacobian matrix is

$$J = \begin{bmatrix} a_{11} & -\frac{ub}{(1+u)^2\varepsilon} \\ (\bar{a}^* - \alpha)(\beta - \bar{a}^*)(-u-1) & 0 \end{bmatrix},$$

where $a_{11} = \frac{1}{\varepsilon}\{\binom{N-1}{n_{pg}-1}(\frac{u}{1+u})^{n_{pg}-1}(\frac{1}{1+u})^{N-n_{pg}} rb[n_{pg} - 1 - (N-1)\frac{u}{1+u}] - b\delta \frac{u}{1+u} \sum_{i=1}^{N-1} i(\frac{1}{1+u})^i\}$, thus the fixed equilibrium is unstable since $\det J < 0$.

Here, our primary focus is on when the optimal state is stable, and when the worst state is unstable. Then, based on the stability of the optimal states of coupled system II, we present the following two conclusions:

Case 1. When $\delta > \alpha$, the optimal state $(1, \alpha)$ of the coupled system is stable.

(i) If $\delta > \frac{\beta}{N}$, the worst state of the coupled system II is unstable.

(ii) If $\delta < \frac{\beta}{N}$, the worst state of the coupled system II is stable.

Case 2. When $\delta < \alpha$, the optimal state $(1, \alpha)$ of the coupled system is unstable.

(i) If $\delta > \frac{\beta}{N}$, the worst state of the coupled system II is unstable.

(ii) If $\delta < \frac{\beta}{N}$, the worst state of the coupled system II is stable.

Therefore, to achieve the optimal state and escape the worst state in coupled system II, it is necessary to ensure that $\delta > \max\{\alpha, \frac{\beta}{N}\}$.

In accordance with the theoretical analyses, we present the corresponding numerical results in the main text.

[1] M. Perc, J. J. Jordan, D. G. Rand, Z. Wang, S. Boccaletti, and A. Szolnoki, Statistical physics of human cooperation, Phys. Rep. **687**, 1 (2017).

[2] M. Perc and A. Szolnoki, Coevolutionary games—a mini review, Biosystems **99**, 109 (2010).

[3] M. Perc, J. Gómez-Gardenes, A. Szolnoki, L. M. Floría, and Y. Moreno, Evolutionary dynamics of group interactions on structured populations: a review, J. R. Soc. Interface **10**, 20120997 (2013).

[4] G. Hardin, Extensions of "the tragedy of the commons," Science **280**, 682 (1998).

[5] D. Feeny, F. Berkes, B. J. McCay, and J. M. Acheson, The tragedy of the commons: twenty-two years later, Human Ecol. **18**, 1 (1990).

[6] A. Szolnoki and M. Perc, Second-order free-riding on antisocial punishment restores the effectiveness of prosocial punishment, Phys. Rev. X **7**, 041027 (2017).

[7] Y.-H. Lin and J. S. Weitz, Spatial interactions and oscillatory tragedies of the commons, Phys. Rev. Lett. **122**, 148102 (2019).

[8] A. Szolnoki and M. Perc, Conformity enhances network reciprocity in evolutionary social dilemmas, J. R. Soc. Interface. **12**, 20141299 (2015).

[9] C. Xia, X. Li, Z. Wang, and M. Perc, Doubly effects of information sharing on interdependent network reciprocity, New J. Phys. **20**, 075005 (2018).

[10] C. Xia, Z. Hu, and D. Zhao, Costly reputation building still promotes the collective trust within the networked population, New J. Phys. **24**, 083041 (2022).

[11] F. Fu, C. Hauert, M. A. Nowak, and L. Wang, Reputation-based partner choice promotes cooperation in social networks, Phys. Rev. E **78**, 026117 (2008).

[12] M. H. Duong and T. A. Han, Cost efficiency of institutional incentives for promoting cooperation in finite populations, Proc. R. Soc. A **477**, 20210568 (2021).

[13] A. Szolnoki and G. Szabó, Cooperation enhanced by inhomogeneous activity of teaching for evolutionary Prisoner's Dilemma games, Europhys. Lett. **77**, 30004 (2007).

[14] Z. Wang, M. Jusup, L. Shi, J.-H. Lee, Y. Iwasa, and S. Boccaletti, Exploiting a cognitive bias promotes cooperation in social dilemma experiments, Nat. Commun. **9**, 2954 (2018).

[15] M. Perc and A. Szolnoki, Social diversity and promotion of cooperation in the spatial Prisoner's Dilemma game, Phys. Rev. E **77**, 011904 (2008).

[16] J. M. Pacheco, F. C. Santos, M. O. Souza, and B. Skyrms, Evolutionary dynamics of collective action in *N*-person stag hunt dilemmas, Proc. R. Soc. B **276**, 315 (2009).

[17] Q. Luo, L. Liu, and X. Chen, Evolutionary dynamics of cooperation in the *N*-person stag hunt game, Physica D **424**, 132943 (2021).

[18] Y. Zhang, J. Wang, C. Ding, and C. Xia, Impact of individual difference and investment heterogeneity on the collective cooperation in the spatial public goods game, Knowledge-Based Syst. **136**, 150 (2017).

[19] T. A. Han, L. M. Pereira, and T. Lenaerts, Avoiding or restricting defectors in public goods games? J. R. Soc. Interface **12**, 20141203 (2015).

[20] J. Quan, X. Li, and X. Wang, The evolution of cooperation in spatial public goods game with conditional peer exclusion, Chaos **29**, 103137 (2019).

[21] J. Quan, X. Chen, W. Yang, and X. Wang, Cooperation dynamics in spatial public goods games with graded punishment mechanism, Nonlin. Dyn. **111**, 8837 (2023).

[22] A. Szolnoki and X. Chen, Tactical cooperation of defectors in a multi-stage public goods game, Chaos Solitons Fractals **155**, 111696 (2022).

[23] M. Milinski, R. D. Sommerfeld, H.-J. Krambeck, F. A. Reed, and J. Marotzke, The collective-risk social dilemma and the

prevention of simulated dangerous climate change, Proc. Natl. Acad. Sci. USA **105**, 2291 (2008).
[24] F. C. Santos and J. M. Pacheco, Risk of collective failure provides an escape from the tragedy of the commons, Proc. Natl. Acad. Sci. USA **108**, 10421 (2011).
[25] C. Hilbe, M. Abou Chakra, P. M. Altrock, and A. Traulsen, The evolution of strategic timing in collective-risk dilemmas, PLoS ONE **8**, e66490 (2013).
[26] X. Chen, A. Szolnoki, and M. Perc, Risk-driven migration and the collective-risk social dilemma, Phys. Rev. E **86**, 036101 (2012).
[27] X. Chen, A. Szolnoki, and M. Perc, Averting group failures in collective-risk social dilemmas, Europhys. Lett. **99**, 68003 (2012).
[28] X. Chen, Y. Zhang, T.-Z. Huang, and M. Perc, Solving the collective-risk social dilemma with risky assets in well-mixed and structured populations, Phys. Rev. E **90**, 052823 (2014).
[29] J. Quan, X. Zhang, W. Chen, and X. Wang, Cooperation dynamics in collective risk games with endogenous endowments, Chaos **33**, 073107 (2023).
[30] Z. Wang, M. Jusup, H. Guo, L. Shi, S. Geček, M. Anand, M. Perc, C. T. Bauch, J. Kurths, S. Boccaletti *et al.*, Communicating sentiment and outlook reverses inaction against collective risks, Proc. Natl. Acad. Sci. USA **117**, 17650 (2020).
[31] J. Wang, F. Fu, T. Wu, and L. Wang, Emergence of social cooperation in threshold public goods games with collective risk, Phys. Rev. E **80**, 016101 (2009).
[32] L. Liu and X. Chen, Evolution of public cooperation in a risky society with heterogeneous assets, Front. Phys. **5**, 67 (2018).
[33] V. V. Vasconcelos, F. C. Santos, J. M. Pacheco, and S. A. Levin, Climate policies under wealth inequality, Proc. Natl. Acad. Sci. USA **111**, 2212 (2014).
[34] N. He, X. Chen, and A. Szolnoki, Central governance based on monitoring and reporting solves the collective-risk social dilemma, Appl. Math. Comput. **347**, 334 (2019).
[35] W. Sun, L. Liu, X. Chen, A. Szolnoki, and V. V. Vasconcelos, Combination of institutional incentives for cooperative governance of risky commons, iScience **24**, 102844 (2021).
[36] V. V. Vasconcelos, F. C. Santos, and J. M. Pacheco, A bottom-up institutional approach to cooperative governance of risky commons, Nat. Clim. Change **3**, 797 (2013).
[37] S. Hua and L. Liu, Governance of risky public goods under the threat of ostracism, Physica D **454**, 133836 (2023).
[38] S. Hua, Z. Hui, and L. Liu, Evolution of conditional cooperation in collective-risk social dilemma with repeated group interactions, Proc. R. Soc. B **290**, 20230949 (2023).
[39] 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, Proc. Natl. Acad. Sci. USA **113**, E7518 (2016).
[40] A. R. Tilman, J. B. Plotkin, and E. Akçay, Evolutionary games with environmental feedbacks, Nat. Commun. **11**, 915 (2020).
[41] Z. Wang, Z. Song, C. Shen, and S. Hu, Emergence of punishment in social dilemma with environmental feedback, Proc. AAAI Conf. Artif. Intell. **37**, 11708 (2023).
[42] L. Shu and F. Fu, Determinants of successful mitigation in coupled social-climate dynamics, Proc. R. Soc. A **479**, 20230679 (2023).
[43] S. Hua and L. Liu, Coevolutionary dynamics of population and institutional rewards in public goods games, Expert Syst. Appl. **237**, 121579 (2024).
[44] M. R. Arefin and J. Tanimoto, Imitation and aspiration dynamics bring different evolutionary outcomes in feedback-evolving games, Proc. R. Soc. A **477**, 20210240 (2021).
[45] L. Liu, X. Chen, and A. Szolnoki, Coevolutionary dynamics via adaptive feedback in collective-risk social dilemma game, eLife **12**, e82954 (2023).
[46] J. Tanimoto, A simple scaling of the effectiveness of supporting mutual cooperation in donor-recipient games by various reciprocity mechanisms, BioSystems **96**, 29 (2009).
[47] J. Tanimoto and H. Sagara, Relationship between dilemma occurrence and the existence of a weakly dominant strategy in a two-player symmetric game, BioSystems **90**, 105 (2007).
[48] Z. Wang, S. Kokubo, M. Jusup, and J. Tanimoto, Universal scaling for the dilemma strength in evolutionary games, Phys. Life Rev. **14**, 1 (2015).
[49] J. Tanimoto, *Fundamentals of Evolutionary Game Theory and Its Applications* (Springer, Japan, 2015).
[50] H. Ito and J. Tanimoto, Scaling the phase-planes of social dilemma strengths shows game-class changes in the five rules governing the evolution of cooperation, R. Soc. Open Sci. **5**, 181085 (2018).
[51] A. R. Góis, F. P. Santos, J. M. Pacheco, and F. C. Santos, Reward and punishment in climate change dilemmas, Sci. Rep. **9**, 16193 (2019).
[52] C. Hilbe, Š. Šimsa, K. Chatterjee, and M. A. Nowak, Evolution of cooperation in stochastic games, Nature (London) **559**, 246 (2018).
[53] A. Traulsen, J. C. Claussen, and C. Hauert, Coevolutionary dynamics: From finite to infinite populations, Phys. Rev. Lett. **95**, 238701 (2005).
[54] C. Hilbe, Local replicator dynamics: a simple link between deterministic and stochastic models of evolutionary game theory, Bull. Math. Biol. **73**, 2068 (2011).