Chaos ARTICLE scitation.org/journal/cha

Replicator dynamics of public goods games with global exclusion

Cite as: Chaos **32**, 073132 (2022); doi: 10.1063/5.0099697
Submitted: 18 May 2022 · Accepted: 27 June 2022 ·
Published Online: 28 July 2022

View Online Export Citation CrossMark

Xiaofeng Wang[1,2,a)] and Matjaž Perc[3,4,5,6,b)]

AFFILIATIONS

[1] Department of Automation, School of Information Science & Technology, Donghua University, Shanghai 201620, China
[2] Engineering Research Center of Digitized Textile & Apparel Technology, Donghua University, Ministry of Education, Shanghai 201620, China
[3] Faculty of Natural Sciences and Mathematics, University of Maribor, Koroška cesta 160, 2000 Maribor, Slovenia
[4] Department of Medical Research, China Medical University Hospital, China Medical University, Taichung 404332, Taiwan
[5] Alma Mater Europaea, Slovenska ulica 17, 2000 Maribor, Slovenia
[6] Complexity Science Hub Vienna, Josefstädterstraße 39, 1080 Vienna, Austria

[a)] **Author to whom correspondence should be addressed:** rokywang1987@gmail.com
[b)] matjaz.perc@gmail.com

ABSTRACT

Studies to date on the role of social exclusion in public cooperation have mainly focused on the peer or pool sanctioning types of excluding free-riders from the share of common goods. However, the exclusive behaviors are not necessarily performed by individuals or local organizations but may rather be implemented by a centralized enforcement institution at a global scale. Besides, previous modeling methods of either peer or pool exclusion often presuppose some particular forms of feedback between the individual or collective efforts and the efficiency of social exclusion and, therefore, cannot comprehensively evaluate their effects on the evolution of cooperation in the social dilemma situations. Here, we construct a general model of global exclusion by considering the successful construction of the centralized exclusive institution as an arbitrary non-decreasing and smooth function of the collective efforts made by the global excluders and then theoretically analyze its potential impacts in the replicator dynamics of the public goods game. Interestingly, we have shown that, despite the presence of both the first- and second-order free-riding problems, global exclusion can indeed lead to the emergence or even stabilization of public cooperation without the support of any other evolutionary mechanism. In addition, we have also observed rich dynamical behaviors, such as the occurrence of a global or local family of neutrally stable closed orbits revolving around a nonlinear center or the existence of stable heteroclinic cycles between defectors, cooperators as well as global excluders, which give rise to a classification of up to 21 different phases.

The overexploitation of natural resources, climate inaction, and vaccination hesitancy are all different examples of public goods dilemmas, the dilemma being that what might be best for an individual in the short term is at odds with that is best for the public goods in the long term. The public goods game is long established as the theoretical framework to study what might avert the tragedy of the common and promote cooperation, i.e., acting selflessly and at a personal cost for the greater good. Previous studies have often considered rewarding and punishment, as well as various forms of reputation and image scoring as possible strategies and mechanisms to enhance cooperation, all to different levels of success depending on the details of implementation. In our work, we consider global exclusion as an extreme form of punishment that is enforced by a centralized institution, and we perform a theoretical analysis of its merits by means of replicator dynamics. We show that global exclusion alone, without the support of other strategies or mechanisms, can promote public cooperation by means of a rich plethora of fascinating dynamical states that emerge due to the considered nonlinear evolutionary dynamics.

I. INTRODUCTION

Replicator dynamics is one of the most classical approaches to study evolutionary games in the infinite and well-mixed populations.[1] From a mathematical point of view, it is described by a set of ordinary differential equations, i.e., the replicator equations,[2]

$$\dot{x}_i = x_i\left[E(P_i) - E(\bar{P})\right], \ i = 1, 2, \ldots, n, \tag{1}$$

where $i = 1, 2, \ldots, n$ labels each of the n (≥ 2) strategies in an evolutionary game. x_i denotes the frequency of the ith strategy in a population, which satisfies $\sum_{i=1}^{n} x_i = 1$. $E(P_i)$ represents the expected payoff of strategy i, which can be used to express the average payoff of the population $E(\bar{P}) = \sum_{i=1}^{n} x_i E(P_i)$. From Eq. (1), one can find that the replicator equations assume that the difference between the expected payoff of one strategy and the mean payoff of the population determines direction and speed of natural selection, which reflects the idea that strategies performing relatively better become more abundant in the population. The replicator dynamics can be conveniently described on the simplex S_n, wherein each point denotes a population state: Each vertex of S_n corresponds to a homogeneous state in which only one strategy is present; each face of S_n represents a heterogeneous state that is defined by one or several strategies being present; and each point inside S_n refers to a fully heterogeneous state where all strategies are present. Interestingly, the simplex S_n under replicator dynamics governed by Eq. (1) has many important properties from the perspective of dynamical systems.[3,4] For instance, it is invariant in the sense that each trajectory starting from the simplex never leaves the simplex. The corners of the simplex are all fixed points because the replicator equation depicts dynamics of pure selection.

One of the most important applications of replicator dynamics is to explain the evolution of cooperation in both nature and human society.[5] The standard metaphor for studying the evolutionary game dynamics of cooperation in the context of group interaction is public goods game.[6] In such an N-player social dilemma game, each player can simultaneously decide whether or not to contribute to the common pool: Cooperators make a certain amount of contribution (= 1 without loss of generality), while defectors do not make any contribution. Then, the sum of contributions is multiplied by an enhancement factor $r \in (1, N)$, which takes the synergistic impacts of cooperation into account, and the produced public goods are distributed equally among all group members irrespective of their initial decisions on contribution. According to the above rules of the public goods game, the payoffs of defectors and cooperators in a group can thus be, respectively, given by

$$\begin{cases} P_D = \dfrac{r(N - N_D)}{N}, & \text{(2a)} \\[2ex] P_C = \dfrac{rN_C}{N} - 1, & \text{(2b)} \end{cases}$$

where N_D and $N_C = N - N_D$ represent the number of defectors and cooperators, respectively. Therefore, the expected payoffs for defectors and cooperators in an infinitely well-mixed population are

$$\begin{cases} E(P_D) = \dfrac{r(N-1)(1-x_D)}{N}, & \text{(3a)} \\[2ex] E(P_C) = \dfrac{r\left[(N-1)x_C + 1\right]}{N} - 1, & \text{(3b)} \end{cases}$$

where x_D and $x_C = 1 - x_D$ denote, respectively, the relative frequencies of defectors and cooperators in the population. Because $E(P_D) - E(P_C) = 1 - r/N > 0$ always holds for any composition of the population [see Eq. (3)], all individuals in the populations will refuse to make any contributions to the common pools eventually, which is thus the unique stable fixed point for the replicator dynamics in an infinite and well-mixed population and is also the evolutionarily stable strategy in evolutionary game theory.[7] This in turn leads to no production of any public goods and thus no collective benefits distributed to any player. However, had each group member contributed to the common pool, everyone would have received a positive net benefit $r - 1 > 0$. Apparently, the public goods game captures elegantly the essence of cooperation puzzle in social dilemma situations. In the past few decades, many mechanisms have been proposed by researchers to explain the evolution of cooperation in public goods game. For example, Hauert *et al.* have shown that the introduction of voluntary participation into the public goods game is able to induce the emergence of cooperation by cyclic dominance between cooperators, loners, and defectors.[8] Santos *et al.* have found that population structures organized by networks can promote the evolution of cooperation in the public goods game.[9] Wang *et al.* have revealed that collective risk is able to provide a way out of the tragedy of the commons.[10] Arenas *et al.* have demonstrated that the indiscriminate destruction toward public goods can result in the bursts of cooperation by forming a rock-paper-scissors dynamics.[11]

Noteworthy, a large number of studies have also focused on the impacts of costly punishment and reward in the evolutionary dynamics of public cooperation or in more realistic scenarios, such as the issue of "yield to pedestrians" in recent years.[12–14] Generally speaking, without the support of other mechanisms, such as network reciprocity, direct reciprocity, and indirect reciprocity, such a manner of negative or positive incentives cannot construct long-term cooperation due to the presence of the second-order free-rider problems in the evolutionary public goods game. However, as an important manner of negative incentives, exclusive behaviors have been found by numerous researchers to be able to overcome both the first- and second-order free-rider problems in various types of populations.[15–19] Note that previous studies on the resolution of public cooperation dilemmas by social exclusions concentrate exclusively in the impacts of peer and pool exclusions, which are executed by local excluders, either collectively or individually, within each group to enforce cooperation in that group of individuals. Nevertheless, exclusive behaviors may also act at the global scale through the construction of a centralized enforcement institution, such as the United Nations—supported by all excluders in the population—which regulates all group interactions in the population. In this paper, we aim to construct a general model of global exclusion in a stochastic manner and to study its effects on the replicator dynamics of public goods game. Interestingly, we have found that global exclusion leads to the emergence or even stabilization of cooperation and, therefore, can solve the tragedy of the commons in the public goods game. In addition, we have shown that the introduction of global exclusion also opens the gate to fascinating rich dynamical behaviors, such as the occurrence of a family of neutrally stable closed orbits or a stable heteroclinic limiting cycle between defectors, cooperators, and global excluders.

This paper is organized as follows. In Sec. II, we present the evolutionary model of public goods game with global exclusion. In Sec. III, we first classify the replicator dynamics and then provide some numerical examples to verify the theoretical analysis. Section IV is devoted to concluding remarks and further discussions.

II. EVOLUTIONARY PUBLIC GOODS GAME WITH GLOBAL EXCLUSION

Consider an infinitely well-mixed population consisting of x_D defectors, x_C cooperators, and $x_{GE} = 1 - x_D - x_C$ global excluders. In the evolutionary public goods game with global exclusion, a number of N individuals are randomly chosen from the population to constitute a group. The N group members adopt one of the following three strategies: defection, cooperation, and global exclusion. While defectors and cooperators play the same role as they do in the classical public goods game, global excluders in our model not only contribute to the joint venture but are also willing to pay a cost $c_{GE} > 0$ to construct a centralized exclusive institution for the purpose of excluding free-riders from sharing the public goods produced in the groups they belong to. Note that global excluders in our model are challenged by dual social dilemmas: (1) The first-order free-riding problem: Cooperative players, who contribute to the common pool, seem to fare worse than those who do not cooperate. (2) The second-order free-riding problem: Global exclusion seems to be an altruistic act, given that players, who do cooperate but do not make joint efforts to construct the coercive organization, are better off than global excluders.

Here, we introduce global exclusion into the evolutionary public goods game in a stochastic manner: The centralized exclusive institution is successfully constructed with a probability $p(x_{GE}) \in [0, 1]$ satisfying

$$p'(x_{GE}) \geq 0, \tag{4}$$

which takes into account the collective efforts of global excluders in the construction of the exclusive institution and, therefore, is designed as a general non-decreasing and smooth function of x_{GE}. Otherwise, the construction of exclusive institution fails. In this case, the payoffs of defectors, cooperators, and global excluders in the group, respectively, become

$$\begin{cases} P_D = \left\{1 - p(x_{GE})\left[1 - \varepsilon(-N_{GE})\right]\right\} \dfrac{r(N - N_D)}{N}, & \text{(5a)} \\ P_C = \left\{1 - p(x_{GE})\left[1 - \varepsilon(-N_{GE})\right]\right\} \dfrac{r(N_C + N_{GE})}{N} \\ \qquad + p(x_{GE})\left[1 - \varepsilon(-N_{GE})\right] r - 1, & \text{(5b)} \\ P_{GE} = \left[1 - p(x_{GE})\right] \dfrac{r(N_C + N_{GE})}{N} + p(x_{GE})\, r - 1 - c_{GE}, & \text{(5c)} \end{cases}$$

where N_D, N_C, and $N_{GE} = N - N_D - N_C$ denote the number of defectors, cooperators, and global excluders, respectively. Here, $\varepsilon(z)$ is the unit step function, which satisfies

$$\varepsilon(z) = \begin{cases} 0, & z < 0, \\ 1, & z \geq 0. \end{cases} \tag{6}$$

In our model, we assume that the exclusive institution would exclude defectors within a particular group from sharing the public goods only if global excluders do exist in this group [see Eq. (5)]. Therefore, the second term in the curly bracket of Eq. (5a) represents the probability of successful exclusion toward defectors, which should satisfy the following two conditions: (1) The exclusive institution is succeeded to be constructed. (2) The involved group does exist global excluders. Equation (5a) thus denotes the expected payoff obtained by the focal defector if defectors are not successfully excluded. Note that the focal defector obtains nothing if defectors are successfully excluded. Similarly, the first terms in Eqs. (5b) and (5c) denote the expected public goods shared, respectively, by the focal cooperator and the focal global excluder if defectors are not successfully excluded, whereas the second terms represent the expected public goods shared, respectively, by the focal cooperator and the focal global excluder otherwise.

III. ANALYSIS AND RESULTS

For replicator dynamics in an infinite and well-mixed population, the interaction groups are randomly formed in accordance with binomial sampling. Then, the probability with which a focal individual interacts with N_D defectors, N_C cooperators, and N_{GE} global excluders among its $N - 1$ co-players is given by

$$\frac{(N-1)!}{N_D! N_C! N_{GE}!} {x_D}^{N_D} {x_C}^{N_C} {x_{GE}}^{N_{GE}}, \tag{7}$$

with $N_D + N_C + N_{GE} = N - 1$. Thus, the average number of contributors among the focal individual's co-players is equal to

$$\sum_{i=0}^{N-1} \frac{(N-1)!}{(N-1-i)!i!} {x_D}^{N-1-i} (1 - x_D)^i = (N-1)(1 - x_D). \tag{8}$$

Similarly, one can also obtain the average number of defectors, cooperators as well as global excluders that the focal individual confronts are $(N-1)x_D$, $(N-1)x_C$, and $(N-1)x_{GE}$, respectively. According to the evolutionary game theoretical model described in Sec. II, the expected payoffs for defectors, cooperators, and global excluders in an infinite and well-mixed population can, therefore, be given by

$$\begin{cases} E(P_D) = r\left[1 - p(x_{GE})\right] \dfrac{(N-1)(1 - x_D)}{N} \\ \qquad + rp(x_{GE}) \dfrac{(N-1)(1 - x_{GE} - x_D)(1 - x_{GE})^{N-2}}{N}, & \text{(9a)} \\ E(P_C) = r\left[1 - p(x_{GE})\right] \dfrac{1 + (N-1)(1 - x_D)}{N} \\ \qquad + rp(x_{GE}) \left[1 - \dfrac{(N-1)x_D(1 - x_{GE})^{N-2}}{N}\right] - 1, & \text{(9b)} \\ E(P_{GE}) = r\left[1 - p(x_{GE})\right] \dfrac{1 + (N-1)(1 - x_D)}{N} \\ \qquad + rp(x_{GE}) - 1 - c_{GE}, & \text{(9c)} \end{cases}$$

respectively. The replicator equations for the public goods game with global exclusion can then be written as

$$\begin{cases} \dot{x}_D = x_D\left[E(P_D) - E(\bar{P})\right], & \text{(10a)} \\ \dot{x}_C = x_C\left[E(P_C) - E(\bar{P})\right], & \text{(10b)} \\ \dot{x}_{GE} = x_{GE}\left[E(P_{GE}) - E(\bar{P})\right], & \text{(10c)} \end{cases}$$

where $E(\bar{P}) = x_D E(P_D) + x_C E(P_C) + x_{GE} E(P_{GE})$ denotes the average population payoff.

A. Classification of dynamics

The replicator dynamics of the three strategies, i.e., defection, cooperation, and global exclusion, take place in the state space $S_3 = \{(x_D, x_C, x_{GE}) : x_D, x_C, x_{GE} \geq 0, x_D + x_C + x_{GE} = 1\}$. The three homogeneous states, i.e., the pure D state $(1,0,0)$, the pure C state $(0,1,0)$, and the pure GE state $(0,0,1)$, corresponding to the three vertices of the simplex S_3, are trivial fixed points for the evolutionary game dynamics described by Eq. (8). Of the three homogeneous states, the pure D state is stable unless $c_{GE} \in [0, r-1)$ and $p(x_{GE}=0) \in \left(\frac{N(c_{GE}+1)-r}{(N-1)r}, 1\right]$, in which case it is a saddle point. Note that the stability of equilibriums in the evolutionary system can often be analyzed by linearizing about the fixed points. However, the linear stability analysis fails when the Jacobian matrix, evaluated at an equilibrium point, may have some eigenvalues with zero real parts and no eigenvalues with positive real parts, e.g., the two eigenvalues of the Jacobian matrix at the equilibrium $(1,0,0)$ when $c_{GE} \in [0, r-1)$ and $p(x_{GE}=0) = \frac{N(c_{GE}+1)-r}{(N-1)r}$: 0 and $\frac{r}{N} - 1 < 0$, in which case one requires to use the methods of advanced stability analysis (for example, see the advanced stability analysis of the homogeneous D state in the Appendix).[20] If $c_{GE} > 0$, the pure C state is a saddle point; otherwise, it is unstable. The pure GE state is unstable except when $c_{GE} = 0$ and $p(x_{GE}=1) \in \left(\frac{N-r}{(N-1)r}, 1\right]$, in which case, it is stable as well as when $c_{GE} \in (0, r-1)$ and $p(x_{GE}=1) \in \left(\frac{N(c_{GE}+1)-r}{(N-1)r}, 1\right]$, in which case, it is a saddle point. There are no interior fixed points on the edge $D-C$ of the simplex S_3 because

$$E(P_D) - E(P_C) = 1 - \frac{r}{N} > 0 \tag{11}$$

always holds for $x_D \in (0,1)$, $x_C \in (0,1)$, and $x_{GE} = 0$. This also indicates that the evolutionary trajectory on the edge $D-C$ is unidirectional from C to D. On the edge $D-GE$, there exists a unique interior fixed point U at $\left(1 - x_{GE}^-, 0, x_{GE}^-\right)$ for $x_{GE}^- \in (0,1)$ satisfying

$$E(P_D) - E(P_{GE}) = 1 + c_{GE} - \frac{r(N-1)}{N} p\left(x_{GE}^-\right) - \frac{r}{N} = 0 \tag{12}$$

as long as $c_{GE} \in [0, r-1)$, $p(x_{GE}=0) \in \left[0, \frac{N(c_{GE}+1)-r}{(N-1)r}\right)$, $p(x_{GE}=1) \in \left(\frac{N(c_{GE}+1)-r}{(N-1)r}, 1\right]$, and $p(x_{GE}) \neq p\left(x_{GE}^-\right)$ for any x_{GE} belonging to an arbitrarily small neighborhood of x_{GE}^-. The unique interior fixed point U on the edge $D-GE$ is unstable if it does exist. In addition, if $p(x_{GE}) = p\left(x_{GE}^-\right) = \frac{N(c_{GE}+1)-r}{(N-1)r}$ with $c_{GE} \in [0, r-1)$ for any x_{GE} belonging to the neighborhood of x_{GE}^- is satisfied for an $S-T$ segment, where S and T, respectively, locate at $\left(1 - x_{GE}^- + \Delta_1, 0, x_{GE}^- - \Delta_1\right)$ and $\left(1 - x_{GE}^- - \Delta_2, 0, x_{GE}^- + \Delta_2\right)$ with $\Delta_1 \in \left(0, x_{GE}^-\right)$ and $\Delta_2 \in \left(0, 1 - x_{GE}^-\right)$, each point on the segment $S-T$ is a fixed point, in which case the fixed points satisfying $x_{GE} \in \left(0, 1 - \left[\frac{N c_{GE}}{N(c_{GE}+1)-r}\right]^{\frac{1}{N-1}}\right)$ are stable, while those satisfying $x_{GE} \in \left[1 - \left[\frac{N c_{GE}}{N(c_{GE}+1)-r}\right]^{\frac{1}{N-1}}, 1\right)$ are unstable. Otherwise, there are no interior fixed points on the edge $D-GE$ of the simplex S_3. Because

$$E(P_C) - E(P_{GE}) = c_{GE} \geq 0 \tag{13}$$

for $x_D = 0$, $x_C \in (0,1)$, and $x_{GE} \in (0,1)$, the following results on the existence of fixed points on the edge $C-GE$ of the simplex S_3 can be obtained: (1) All interior points on the edge $C-GE$ are fixed points only if $c_{GE} = 0$. (2) No fixed point exists on the edge $C-GE$ if $c_{GE} > 0$. When $c_{GE} = 0$, the stability of the interior fixed points $\left(0, 1 - x_{GE}^+, x_{GE}^+\right)$ on the edge $C-GE$ is determined by the value of $p(x_{GE}=1)$: (1) All the interior fixed points are unstable if $p(x_{GE}=1) \in \left[0, \frac{N-r}{(N-1)r}\right]$. (2) If, however, $p(x_{GE}=1) \in \left(\frac{N-r}{(N-1)r}, 1\right]$, there exists a critical fixed point V satisfying

$$p\left(x_{GE}^+\right) = \frac{N-r}{(N-1)\, r\left[1 - \left(1 - x_{GE}^+\right)^{N-1}\right]}, \tag{14}$$

which divides the edge $C-GE$ into two segments with stable fixed points on the segment $V-GE$ and unstable fixed points on the segment $C-V$. Note that the three trivial fixed points on the vertices of the simplex S_3 are all saddle points when $c_{GE} \in (0, r-1)$, $p(x_{GE}=1) \in \left(\frac{N(c_{GE}+1)-r}{(N-1)r}, 1\right]$, and $p(x_{GE}=0) \in \left(\frac{N(c_{GE}+1)-r}{(N-1)r}, 1\right]$. In this case, defectors are wiped out by global excluders, cooperators are dominated by defectors, and global excluders are in turn invaded by cooperators, which leads to the formation of a heteroclinic cycle on the boundary of the simplex S_3. For the purpose of determining the stability of the heteroclinic cycle, one should consider the eigenvalues of the Jacobian matrix of the three trivial fixed points,

$$\begin{cases} \lambda_1|_{(1,0,0)} = -1 + \dfrac{r}{N} < 0,\ \lambda_2|_{(1,0,0)} \\ \qquad = -c_{GE} - 1 + \dfrac{r}{N} + \dfrac{N-1}{N} r p(x_{GE}=0) > 0, & \text{(15a)} \\ \lambda_1|_{(0,1,0)} = -c_{GE} < 0,\ \lambda_2|_{(0,1,0)} = 1 - \dfrac{r}{N} > 0, & \text{(15b)} \\ \lambda_1|_{(0,0,1)} = 1 - \dfrac{r}{N} + c_{GE} - \dfrac{N-1}{N} r p(x_{GE}=1) < 0,\ \lambda_2|_{(0,0,1)} \\ \qquad = c_{GE} > 0. & \text{(15c)} \end{cases}$$

Let $\lambda|_{(1,0,0)} = -\frac{\lambda_1|_{(1,0,0)}}{\lambda_2|_{(1,0,0)}}$, $\lambda|_{(0,1,0)} = -\frac{\lambda_1|_{(0,1,0)}}{\lambda_2|_{(0,1,0)}}$, and $\lambda|_{(0,0,1)} = -\frac{\lambda_1|_{(0,0,1)}}{\lambda_2|_{(0,0,1)}}$, we thus have

$$\lambda|_{(1,0,0)}\, \lambda|_{(0,1,0)}\, \lambda|_{(0,0,1)} = \frac{r\left[1 + (N-1)\, p(x_{GE}=1)\right] - N(c_{GE}+1)}{r\left[1 + (N-1)\, p(x_{GE}=0)\right] - N(c_{GE}+1)}. \tag{16}$$

From Eq. (16), one can find that the heteroclinic cycle is asymptotically stable when $p(x_{GE}=0) < p(x_{GE}=1)$ but is neutrally stable when $p'(x_{GE}) = 0$ for any $x_{GE} \in [0,1]$.

Now, let us turn to analyze the replicator dynamics in the interior of the simplex S_3. Letting $E(P_D) - E(P_C) = 0$ leads to

$$p(x_{GE})\left[1-(1-x_{GE})^{N-1}\right] = \frac{N-r}{(N-1)r}. \quad (17)$$

As the left-hand and right-hand side of Eq. (17) are, respectively, increased with and independent on x_{GE}, there exists one fixed point at most in the interior of the simplex S_3. By solving Eq. (17), we have

$$x_{GE}{}^{*} = 1-\left[1-\frac{N-r}{rp(x_{GE}{}^{*})(N-1)}\right]^{\frac{1}{N-1}}. \quad (18)$$

Here, one can find that the line determined by Eq. (18) and the edge $C-GE$ of the simplex S_3 intersect at the critical fixed point V, the location of which satisfies Eq. (14). Similarly, by solving $E(P_C) - E(P_{GE}) = 0$ as well as applying Eq. (18), one can get

$$x_D{}^{*} = \frac{Nc_{GE}\left[1-\frac{N-r}{rp(x_{GE}{}^{*})(N-1)}\right]^{\frac{1}{N-1}}}{rp(x_{GE}{}^{*})(N-1)-N+r}. \quad (19)$$

Substituting Eqs. (18) and (19) into $x_C{}^{*} = 1 - x_D{}^{*} - x_{GE}{}^{*}$ yields

$$x_C{}^{*} = \left[1-\frac{N-r}{rp(x_{GE}{}^{*})(N-1)}\right]^{\frac{1}{N-1}} \times\left[1-\frac{Nc_{GE}}{rp(x_{GE}{}^{*})(N-1)-N+r}\right]. \quad (20)$$

Equation (20) indicates that $x_C{}^{*} = 0$ requires

$$p(x_{GE}{}^{*}) = \frac{N(c_{GE}+1)-r}{(N-1)r}, \quad (21)$$

which is equivalent to Eq. (12). This means that the line given by Eq. (18) connects to the edge $D-GE$ of the simplex S_3 at the fixed point U, the location of which satisfies Eq. (12). From Eqs. (18)–(20), one can derive the critical condition that ensures the existence of an interior fixed point $(x_D{}^{*}, x_C{}^{*}, x_{GE}{}^{*})$ inside the simplex S_3: $p(x_{GE}{}^{*}) \in \left(\frac{N(c_{GE}+1)-r}{(N-1)r}, 1\right]$ and $c_{GE} \in (0, r-1)$. The interior fixed point Q at $(x_D{}^{*}, x_C{}^{*}, x_{GE}{}^{*})$ is unstable unless $p'(x_{GE}{}^{*}) = 0$, in which case it is neutrally stable. If $p'(x_{GE}) = 0$ is further satisfied for $x_{GE} \in [x_{GE}{}^{*} - \Delta_3, x_{GE}{}^{*} + \Delta_4]$ with $\Delta_3 \in (0, x_{GE}{}^{*})$ and $\Delta_4 \in (0, 1 - x_{GE}{}^{*})$, the interior fixed point Q becomes a neutrally stable center. In order to show the existence of such an nonlinear center, let us introduce a new variable $x = \frac{x_C}{x_C+x_D}$, which represents the fraction of cooperators among members who do not contribute to the construction of the exclusive institution. This yields

$$\dot{x} = -x(1-x)\left[E(P_D) - E(P_C)\right]. \quad (22)$$

Then, the master equations, i.e., Eq. (10), that describe the replicator dynamics of public goods game with global exclusion can be rewritten as

$$\begin{cases} \dot{x} = -x(1-x)\left\{1-\frac{r}{N}-rp(x_{GE})\left[1-\frac{(N-1)(1-x_{GE})^{N-1}+1}{N}\right]\right\}, & (23a)\\ \dot{x}_{GE} = -x_{GE}(1-x_{GE})\left\{c_{GE}+(1-x)\left[1-\frac{r}{N}-rp(x_{GE})\left(1-\frac{1}{N}\right)\right]\right\}. & (23b)\end{cases}$$

Dividing the right-hand side of Eq. (23) by $x(1-x)x_{GE}(1-x_{GE})$, which is positive for any (x_{GE}, x) on the unit square $(0,1)^2$, leads to

$$\begin{cases} \dot{x} = \frac{-1+\frac{r}{N}+rp(x_{GE}{}^{*})\left[1-\frac{(N-1)(1-x_{GE})^{N-1}+1}{N}\right]}{x_{GE}(1-x_{GE})} = s(x_{GE}), & (24a)\\ \dot{x}_{GE} = \frac{-c_{GE}-(1-x)\left[1-\frac{r}{N}-rp(x_{GE}{}^{*})\left(1-\frac{1}{N}\right)\right]}{x(1-x)} = -w(x), & (24b)\end{cases}$$

where $p(x_{GE})$ is replaced by a constant $p(x_{GE}{}^{*})$ for $x_{GE} \in [x_{GE}{}^{*} - \Delta_3, x_{GE}{}^{*} + \Delta_4]$. Such a transformation corresponds to a change in velocity and does not affect the orbits. Introducing $H(x_{GE}, x) = S(x_{GE}) + W(x)$, where $S(x_{GE})$ and $W(x)$ are primitives of $s(x_{GE})$ and $w(x)$, respectively,

$$\begin{cases} S(x_{GE}) = \left[-1+\frac{r}{N}+\frac{(N-1)r}{N}p(x_{GE}{}^{*})\right]\ln\frac{x_{GE}}{1-x_{GE}} - \frac{(N-1)r}{N}p(x_{GE}{}^{*})\int\frac{(1-x_{GE})^{N-2}}{x_{GE}}dx_{GE}, & (25a)\\ W(x) = \left[c_{GE}+1-\frac{r}{N}-\frac{(N-1)r}{N}p(x_{GE}{}^{*})\right]\ln x - c_{GE}\ln(1-x), & (25b)\end{cases}$$

where

$$\int\frac{(1-x_{GE})^{N-2}}{x_{GE}}dx_{GE} = \sum_{i=1}^{N-2}\binom{N-2}{i}\frac{(-x_{GE})^i}{i} + \ln x_{GE}. \quad (26)$$

Then, we obtain the Hamiltonian system,

$$\begin{cases} \dot{x} = \dfrac{\partial H}{\partial x_{GE}}, & (27a) \\ \dot{x}_{GE} = -\dfrac{\partial H}{\partial x}. & (27b) \end{cases}$$

Due to the fact that the system is conservative and the Hamiltonian H attains a strict maximum at $\left(1-\left(1-\frac{N-r}{rp(x_{GE}^*)(N-1)}\right)^{\frac{1}{N-1}}, 1 - \frac{Nc_{GE}}{rp(x_{GE}^*)(N-1)-N+r}\right)$, the interior fixed point (x_D^*, x_C^*, x_{GE}^*) is a center; i.e., it is neutrally stable and is surrounded by a family of closed orbits in its local neighborhood. Beyond this region, there exist unstable spirals attracted by the stable heteroclinic limit cycle if $p(x_{GE}=1) > p(x_{GE}=0)$ and $p(x_{GE}=0) \in \left[\frac{N(c_{GE}+1)-r}{(N-1)r}, 1\right]$. If, however, $p(x_{GE}=1) > p(x_{GE}=0)$ and $p(x_{GE}=0) \in \left[0, \frac{N(c_{GE}+1)-r}{(N-1)r}\right)$, the evolutionary trajectories move toward the stable fixed point D at $(1,0,0)$. Otherwise, i.e., $p'(x_{GE}) = 0$ for any $x_{GE} \in [0,1]$ (or $p(x_{GE}=0) = p(x_{GE}=1) \in \left(\frac{N(c_{GE}+1)-r}{(N-1)r}, 1\right]$), Eq. (27) shows that $G(x_{GE}) \to -\infty$ for $x_{GE} \to 0, 1$ and $L(x) \to -\infty$ for $x \to 0, 1$, which results in $H \to -\infty$ uniformly near the boundary of $[0,1]^2$. This means that the region of closed orbits is extended from the local neighborhood of the nonlinear center to the whole simplex S_3 in this case.

According to the detailed analysis of the replicator dynamics presented above, one can obtain the following 21 classes of deterministic behavior with different dynamical properties:

Phase i: $c_{GE} = 0$ and $p(x_{GE}=1) \in \left[0, \frac{N-r}{(N-1)r}\right)$; $c_{GE} = 0$ and $p(x_{GE}=1) = \frac{N-r}{(N-1)r} > p(x_{GE})$ for any $x_{GE} \in [0,1)$. There exist an infinite number of unstable fixed points on the edge $C-GE$ and a unique stable fixed point on the vertex D of the simplex S_3.

Phase ii: $c_{GE} = 0$ and $p(x_{GE}) = \frac{N-r}{(N-1)r}$ for $x_{GE} \in \left[x_{GE}^- - \Delta_1, 1\right]$ with $\Delta_1 \in \left(0, x_{GE}^-\right)$. In this phase, the dynamical behaviors are the same as those in Phase i except that there exists an infinite number of interior stable fixed points on the segment $S-GE$ of the edge $D-GE$.

Phase iii: $c_{GE} = 0$, $p(x_{GE}=1) \in \left(\frac{N-r}{(N-1)r}, 1\right]$, $p(x_{GE}=0) \in \left[0, \frac{N-r}{(N-1)r}\right)$ and $p(x_{GE}) \neq p\left(x_{GE}^-\right) = \frac{N-r}{(N-1)r}$ satisfied in an arbitrarily small neighborhood of x_{GE}^-. The vertex D of the simplex S_3 is a stable fixed point, whereas the pure C and GE states are unstable and stable, respectively. On the edge $C-GE$, all points on the segment $C-V$ and $V-GE$ are unstable and stable fixed points, respectively. In addition, there exists an unstable fixed point U on the edge $D-GE$.

Phase iv: $c_{GE} = 0$, $p(x_{GE}=1) \in \left(\frac{N-r}{(N-1)r}, 1\right]$ and $p(x_{GE}) = \frac{N-r}{(N-1)r}$ for $x_{GE} \in \left[x_{GE}^- - \Delta_1, x_{GE}^- + \Delta_2\right]$ with $\Delta_1 \in \left(0, x_{GE}^-\right)$ and $\Delta_2 \in \left(0, 1 - x_{GE}^-\right)$. Here, the dynamical behaviors are the same as those in Phase ii except that there exists an infinite number of stable fixed points on the segment $S-T$ of the edge $D-GE$.

Phase v: $c_{GE} = 0$, $p(x_{GE}=1) \in \left(\frac{N-r}{(N-1)r}, 1\right]$ and $p(x_{GE}) = \frac{N-r}{(N-1)r}$ for $x_{GE} \in \left[0, x_{GE}^- + \Delta_2\right]$ with $\Delta_2 \in \left(0, 1 - x_{GE}^-\right)$. In this case, the dynamical behaviors are the same as those in Phase iv except that the infinite stable fixed points move from the segment $S-T$ to $D-T$ on the edge $D-GE$.

Phase vi: $c_{GE} = 0$ and $p(x_{GE}=0) \in \left(\frac{N-r}{(N-1)r}, 1\right]$; $c_{GE} = 0$, $p\left(x_{GE} = x_{GE}^- = 0\right) = \frac{N-r}{(N-1)r}$ and $p(x_{GE}) \neq p\left(x_{GE}^-\right) = \frac{N-r}{(N-1)r}$ satisfied in an arbitrarily small neighborhood of x_{GE}^-. Here, the dynamical behaviors are the same as those in Phase iii except that both the pure D and GE states become saddle points and that the unique interior unstable fixed point U disappears whereas the evolutionary trajectory becomes unidirectional from D to GE on the edge $D-GE$.

Phase vii: $c_{GE} \in (0, r-1)$ and $p(x_{GE}=1) \in \left[0, \frac{N(c_{GE}+1)-r}{(N-1)r}\right)$; $c_{GE} \in (0, r-1)$ and $p(x_{GE}=1) = \frac{N(c_{GE}+1)-r}{(N-1)r} > p(x_{GE})$ for any $x_{GE} \in [0,1)$; $c_{GE} \in [r-1, +\infty)$. In this phase, the dynamical behaviors are the same as those in Phase i except that the pure state C becomes a saddle point and that the infinite number of unstable interior fixed points on the edge $C-GE$ disappear, whereas the evolutionary trajectory becomes unidirectional from GE to C.

Phase viii: $c_{GE} \in (0, r-1)$ and $p(x_{GE}) = \frac{N(c_{GE}+1)-r}{(N-1)r}$ for $x_{GE} \in \left[x_{GE}^- - \Delta_1, 1\right]$ with $\Delta_1 \in \left(0, x_{GE}^-\right)$. In this case, the dynamical behaviors are the same as those in Phase vii except that there exists an infinite number of fixed points on the segment $S-GE$ of the edge $D-GE$, in which case the fixed points satisfying $x_{GE} \in \left(0, 1 - \left[\frac{Nc_{GE}}{N(c_{GE}+1)-r}\right]^{\frac{1}{N-1}}\right)$ are stable, while those satisfying $x_{GE} \in \left[1 - \left[\frac{Nc_{GE}}{N(c_{GE}+1)-r}\right]^{\frac{1}{N-1}}, 1\right)$ are unstable.

Phase ix: $c_{GE} \in (0, r-1)$, $p(x_{GE}=1) \in \left(\frac{N(c_{GE}+1)-r}{(N-1)r}, 1\right]$, $p(x_{GE}=0) \in \left[0, \frac{N(c_{GE}+1)-r}{(N-1)r}\right)$, $p(x_{GE}) \neq p\left(x_{GE}^-\right) = \frac{N(c_{GE}+1)-r}{(N-1)r}$ satisfied in an arbitrarily small neighborhood of x_{GE}^-, and $p'(x_{GE}^*) \neq 0$. The vertex D of the simplex S_3 is a stable fixed point, whereas both the pure C and GE state are saddle points. The evolutionary trajectory is unidirectional from GE to C on the edge $C-GE$ of the simplex S_3. Moreover, there exists a unique unstable fixed point U on the edge $D-GE$. Besides, there also emerges a unique unstable fixed point Q in the interior area of the simplex S_3.

Phase x: $c_{GE} \in (0, r-1)$, $p(x_{GE}=1) \in \left(\frac{N(c_{GE}+1)-r}{(N-1)r}, 1\right]$, $p(x_{GE}=0) \in \left[0, \frac{N(c_{GE}+1)-r}{(N-1)r}\right)$, $p(x_{GE}) \neq p\left(x_{GE}^-\right) = \frac{N(c_{GE}+1)-r}{(N-1)r}$ satisfied in an arbitrarily small neighborhood of x_{GE}^-, and $p'(x_{GE}^*) = 0$. Here, the dynamical behaviors are the same as those in Phase ix except that the unique interior fixed point Q becomes neutrally stable.

Phase xi: $c_{GE} \in (0, r-1)$, $p(x_{GE}=1) \in \left(\frac{N(c_{GE}+1)-r}{(N-1)r}, 1\right]$, $p(x_{GE}=0) \in \left[0, \frac{N(c_{GE}+1)-r}{(N-1)r}\right)$, $p(x_{GE}) \neq p\left(x_{GE}^-\right) = \frac{N(c_{GE}+1)-r}{(N-1)r}$ satisfied in an arbitrarily small neighborhood of x_{GE}^-, and $p'(x_{GE}) = 0$ satisfied for $x_{GE} \in [x_{GE}^* - \Delta_3, x_{GE}^* + \Delta_4]$ with $\Delta_3 \in (0, x_{GE}^*)$ and $\Delta_4 \in (0, 1 - x_{GE}^*)$. In this phase, the dynamical behaviors are the same as those in Phase ix except that the unique interior fixed point Q becomes a neutrally stable center; i.e., it is neutrally stable and is surrounded by a family of closed orbits in its local neighborhood.

Phase xii: $c_{GE} \in (0, r-1)$, $p(x_{GE}=1) \in \left(\frac{N(c_{GE}+1)-r}{(N-1)r}, 1\right]$, $p(x_{GE}) = \frac{N(c_{GE}+1)-r}{(N-1)r}$ for $x_{GE} \in \left[x_{GE}^- - \Delta_1, x_{GE}^- + \Delta_2\right]$ with $\Delta_1 \in \left(0, x_{GE}^-\right)$

and $\Delta_2 \in (0, 1 - x_{GE}{}^-)$, and $p'(x_{GE}{}^*) \neq 0$. In this case, the dynamical behaviors are the same as those in Phase ix except that there exists an infinite number of interior fixed points on the segment $S - T$ of the edge $D - GE$, in which case the fixed points satisfying $x_{GE} \in \left(0, 1 - \left[\frac{Nc_{GE}}{N(c_{GE}+1)-r}\right]^{\frac{1}{N-1}}\right)$ are stable, while those satisfying $x_{GE} \in \left[1 - \left[\frac{Nc_{GE}}{N(c_{GE}+1)-r}\right]^{\frac{1}{N-1}}, 1\right)$ are unstable.

Phase xiii: $c_{GE} \in (0, r-1)$, $p(x_{GE} = 1) \in \left(\frac{N(c_{GE}+1)-r}{(N-1)r}, 1\right]$, $p(x_{GE}) = \frac{N(c_{GE}+1)-r}{(N-1)r}$ for $x_{GE} \in [x_{GE}{}^- - \Delta_1, x_{GE}{}^- + \Delta_2]$ with $\Delta_1 \in (0, x_{GE}{}^-)$ and $\Delta_2 \in (0, 1 - x_{GE}{}^-)$, and $p'(x_{GE}{}^*) = 0$. Here, the dynamical behaviors are the same as those in Phase xii except that the unique interior fixed point Q becomes neutrally stable.

Phase xiv: $c_{GE} \in (0, r-1)$, $p(x_{GE} = 1) \in \left(\frac{N(c_{GE}+1)-r}{(N-1)r}, 1\right]$, $p(x_{GE}) = \frac{N(c_{GE}+1)-r}{(N-1)r}$ for $x_{GE} \in [x_{GE}{}^- - \Delta_1, x_{GE}{}^- + \Delta_2]$ with $\Delta_1 \in (0, x_{GE}{}^-)$ and $\Delta_2 \in (0, 1 - x_{GE}{}^-)$, and $p'(x_{GE}) = 0$ satisfied for $x_{GE} \in [x_{GE}{}^* - \Delta_3, x_{GE}{}^* + \Delta_4]$ with $\Delta_3 \in (0, x_{GE}{}^*)$ and $\Delta_4 \in (0, 1 - x_{GE}{}^*)$. In this phase, the dynamical behaviors are the same as those in Phase xii except that the unique interior fixed point Q becomes a neutrally stable center; i.e., it is neutrally stable and is surrounded by a family of closed orbits in its local neighborhood.

Phase xv: $c_{GE} \in (0, r-1)$, $p(x_{GE} = 1) \in \left(\frac{N(c_{GE}+1)-r}{(N-1)r}, 1\right]$, $p(x_{GE}) = \frac{N(c_{GE}+1)-r}{(N-1)r}$ for $x_{GE} \in [0, x_{GE}{}^- + \Delta_2]$ with $\Delta_2 \in (0, 1 - x_{GE}{}^-)$, and $p'(x_{GE}{}^*) \neq 0$. In this case, the dynamical behaviors are the same as those in Phase xii except that the infinite fixed points move from the segment $S - T$ to $D - T$ on the edge $D - GE$.

Phase xvi: $c_{GE} \in (0, r-1)$, $p(x_{GE} = 1) \in \left(\frac{N(c_{GE}+1)-r}{(N-1)r}, 1\right]$, $p(x_{GE}) = \frac{N(c_{GE}+1)-r}{(N-1)r}$ for $x_{GE} \in [0, x_{GE}{}^- + \Delta_2]$ with $\Delta_2 \in (0, 1 - x_{GE}{}^-)$, and $p'(x_{GE}{}^*) = 0$. Here, the dynamical behaviors are the same as those in Phase xv except that the unique interior fixed point Q becomes neutrally stable.

Phase xvii: $c_{GE} \in (0, r-1)$, $p(x_{GE} = 1) \in \left(\frac{N(c_{GE}+1)-r}{(N-1)r}, 1\right]$, $p(x_{GE}) = \frac{N(c_{GE}+1)-r}{(N-1)r}$ for $x_{GE} \in [0, x_{GE}{}^- + \Delta_2]$ with $\Delta_2 \in (0, 1 - x_{GE}{}^-)$, and $p'(x_{GE}) = 0$ satisfied for $x_{GE} \in [x_{GE}{}^* - \Delta_3, x_{GE}{}^* + \Delta_4]$ with $\Delta_3 \in (0, x_{GE}{}^*)$ and $\Delta_4 \in (0, 1 - x_{GE}{}^*)$. In this phase, the dynamical behaviors are the same as those in Phase xv except that the unique interior fixed point Q becomes a neutrally stable center; i.e., it is neutrally stable and is surrounded by a family of closed orbits in its local neighborhood.

Phase xviii: $c_{GE} \in (0, r-1)$, $p(x_{GE} = 0) \in \left(\frac{N(c_{GE}+1)-r}{(N-1)r}, 1\right]$, $p(x_{GE} = 1) > p(x_{GE} = 0)$ and $p'(x_{GE}{}^*) \neq 0$; $c_{GE} \in (0, r-1)$, $p(x_{GE} = 1) \in \left(\frac{N(c_{GE}+1)-r}{(N-1)r}, 1\right]$, $p(x_{GE} = x_{GE}{}^- = 0) = \frac{N(c_{GE}+1)-r}{(N-1)r}$ and $p(x_{GE}) \neq p(x_{GE}{}^-)$ satisfied in an arbitrarily small neighborhood of $x_{GE}{}^-$, and $p'(x_{GE}{}^*) \neq 0$. The vertices D, C, and GE of the simplex S_3 are saddle points. Furthermore, there exists a stable heteroclinic cycle $GE - C - D$ around the boundary of the simplex S_3. Additionally, there also exists a unique unstable fixed point Q in the interior area of the simplex S_3.

Phase xix: $c_{GE} \in (0, r-1)$, $p(x_{GE} = 0) \in \left(\frac{N(c_{GE}+1)-r}{(N-1)r}, 1\right]$, $p(x_{GE} = 1) > p(x_{GE} = 0)$ and $p'(x_{GE}{}^*) = 0$; $c_{GE} \in (0, r-1)$, $p(x_{GE} = 1) \in \left(\frac{N(c_{GE}+1)-r}{(N-1)r}, 1\right]$, $p(x_{GE} = x_{GE}{}^- = 0) = \frac{N(c_{GE}+1)-r}{(N-1)r}$ and $p(x_{GE}) \neq p(x_{GE}{}^-)$ satisfied in an arbitrarily small neighborhood of $x_{GE}{}^-$, and $p'(x_{GE}{}^*) = 0$. Here, the dynamical behaviors are the same as those in Phase xviii except that the unique interior fixed point Q becomes neutrally stable.

Phase xx: $c_{GE} \in (0, r-1)$, $p(x_{GE} = 0) \in \left(\frac{N(c_{GE}+1)-r}{(N-1)r}, 1\right]$, $p(x_{GE} = 1) > p(x_{GE} = 0)$, and $p'(x_{GE}) = 0$ satisfied for $x_{GE} \in [x_{GE}{}^* - \Delta_3, x_{GE}{}^* + \Delta_4]$ with $\Delta_3 \in (0, x_{GE}{}^*)$ and $\Delta_4 \in (0, 1 - x_{GE}{}^*)$; $c_{GE} \in (0, r-1)$, $p(x_{GE} = 1) \in \left(\frac{N(c_{GE}+1)-r}{(N-1)r}, 1\right]$, $p(x_{GE} = x_{GE}{}^- = 0) = \frac{N(c_{GE}+1)-r}{(N-1)r}$ and $p(x_{GE}) \neq p(x_{GE}{}^-)$ satisfied in an arbitrarily small neighborhood of $x_{GE}{}^-$, and $p'(x_{GE}) = 0$ satisfied for $x_{GE} \in [x_{GE}{}^* - \Delta_3, x_{GE}{}^* + \Delta_4]$ with $\Delta_3 \in (0, x_{GE}{}^*)$ and $\Delta_4 \in (0, 1 - x_{GE}{}^*)$. In this phase, the dynamical behaviors are the same as those in Phase xviii except that the unique interior fixed point Q becomes a neutrally stable center; i.e., it is neutrally stable and is surrounded by a family of closed orbits in its local neighborhood.

Phase xxi: $c_{GE} \in (0, r-1)$, $p(x_{GE} = 0) \in \left(\frac{N(c_{GE}+1)-r}{(N-1)r}, 1\right]$, and $p(x_{GE} = 1) = p(x_{GE} = 0)$. In this case, the dynamical behaviors are the same as those in Phase xx except that the heteroclinic cycle $GE - C - D$ becomes neutral and that the region of closed orbits is extended from the local neighborhood of the nonlinear center to the whole simplex S_3.

B. Numerical examples

In order to verify the results derived by theoretical analysis above, we provide three classes of numerical examples by considering the following normalized logistic function of $p(x_{GE})$:

$$p(x_{GE}) = p_{\min} + \frac{l(x_{GE}) - l(0)}{l(1) - l(0)} (p_{\max} - p_{\min}), \tag{28}$$

where $p_{\min}$ and $p_{\max}$, respectively, set the lower and upper bounds that $p(x_{GE})$ can achieve for $x_{GE} \in [0, 1]$ and, therefore, should satisfy $0 \leq p_{\min} \leq p_{\max} \leq 1$. $l(x_{GE})$ denotes the logistic function,

$$l(x_{GE}) = \frac{1}{1 + \exp[s(h - x_{GE})]}, \tag{29}$$

where $h \in [0, 1]$ defines the position of the inflection point: $h \to 0$ leads to the decrease of $p'(x_{GE})$ with x_{GE}, while $h \to 1$ the increase of $p'(x_{GE})$ with x_{GE} (see Fig. 1); $s > 0$ governs the steepness of the function at the inflection point: $s \to 0$ results in the linear dependence of $p(x_{GE})$ on x_{GE}, whereas $s \to +\infty$ the step function of $p(x_{GE})$ with x_{GE}, whose transition point is at $x_{GE} = h$ (see Fig. 1). Depending on the parameters $p_{\min}$, $p_{\max}$, h, and s determining the shape of $p(x_{GE})$ in Eq. (28), one can classify the numerical examples into the following three typical categories:

Class i. Consider the special case that the construction of the centralized institution is independent on the collective efforts of global excluders,[21] i.e., $p(x_{GE}) = p$, which can be obtained by setting $p_{\min} = p_{\max} = p \in [0, 1]$ in Eq. (28).

Class ii. Consider the limiting case that the construction of the centralized institution replies discounted or synergistically on the

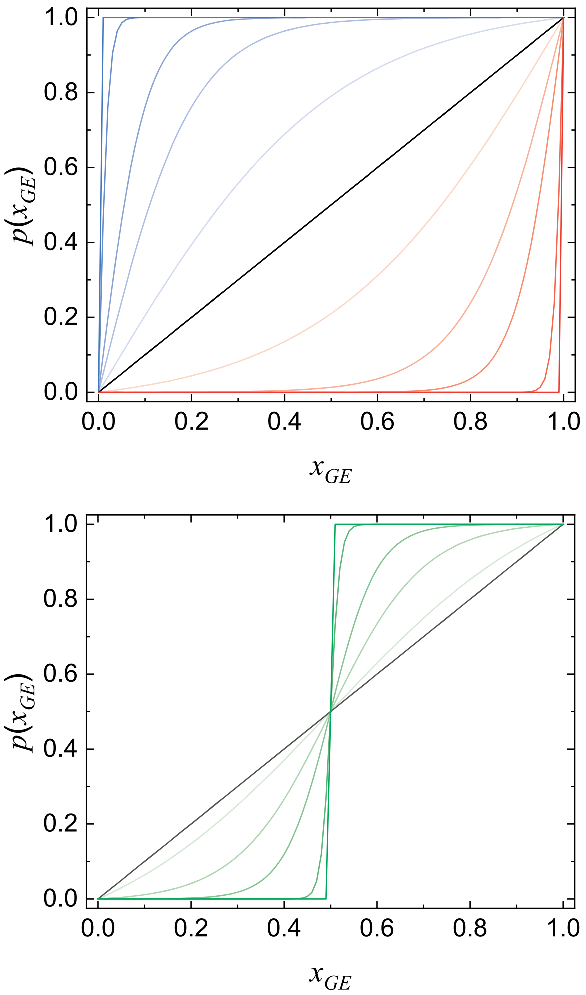


FIG. 1. Construction probability of the centralized exclusive institution $p(x_{GE})$, defined by a normalized logistic function [see Eq. (28)] at three different positions of the inflection points h: $h = 0$ (upper left part of top panel), $h = 1$ (lower right part of top panel), and $h = 0.5$ (bottom panel), in dependence on the frequency of global excluders x_{GE}. For the different classes of function $p(x_{GE})$ resulted from the three values of h, the increment of steepness s leads to the increasing opacity of the curves. Parameter settings: $p_{\min} = 0$ and $p_{\max} = 1$.

joint efforts provided by the global excluders,[22] which can be realized by letting $h = 0$ and $p_{\min} \neq p_{\max}$, or $h = 1$ and $p_{\min} \neq p_{\max}$ in Eq. (28), respectively (see the upper part of Fig. 1).

Class iii. Consider the case that the construction of the centralized institution is sigmoidally dependent on the group efforts by the global excluders,[23] which requires to satisfy $h \in (0, 1)$ and $p_{\min} \neq p_{\max}$ in Eq. (28) (see the bottom part of Fig. 1). Particularly, if $p_{\min} = 0$, $p_{\max} = 1$ and $s \to +\infty$ is further satisfied, the construction of the centralized institution replies deterministically on the threshold parameter h:[24] If $x_{GE} < h$, the construction of the centralized institution fails; If $x_{GE} > h$, the construction of the centralized institution succeeds.

For $p(x_{GE})$ belonging to Class i, the theoretical analysis predicts that the dynamical behaviors of the evolutionary system may be in Phase i or Phase vii if $c_{GE} \in [0, r-1)$ and $p(x_{GE}) = p \in \left[0, \frac{N(c_{GE}+1)-r}{(N-1)r}\right)$ or $c_{GE} \in [r-1, +\infty)$, or in Phase ii if $c_{GE} = 0$ and $p(x_{GE}) = p = \frac{N-r}{(N-1)r}$, or in Phase vi if $c_{GE} = 0$ and $p(x_{GE}) = p \in \left(\frac{N-r}{(N-1)r}, 1\right]$, or in Phase viii if $c_{GE} \in (0, r-1)$ and $p(x_{GE}) = p = \frac{N(c_{GE}+1)-r}{(N-1)r}$, or in Phase xxi if $c_{GE} \in (0, r-1)$ and $p(x_{GE}) = p \in \left(\frac{N(c_{GE}+1)-r}{(N-1)r}, 1\right]$. The full $c_{GE} - p$ phase diagram obtained by numerical simulations confirms our theoretical predictions, as shown in the top left panel of Fig. 2. Particularly, if the global exclusion is costless, i.e., $c_{GE} = 0$, and the construction probability of the centralized exclusive institution is large enough, i.e., $p \in \left(\frac{N-r}{(N-1)r}, 1\right]$, the pure state of global excluders as well as a mixture of cooperators and global excluders along the $V - GE$ segment on the $C - GE$ edge of the simplex S_3 are stable, but the homogeneous state of defectors, that of cooperators as well as a mixture of cooperators and global excluders along the $C - V$ segment on the $C - GE$ edge of the simplex S_3 are all unstable, which means that the first-order free-riding problem is solved, but the second-order free-riding problem may still remain in this case (see the top right panel of Fig. 2). A moderate increment of the cost of global exclusion, i.e., $c_{GE} \in \left(0, \frac{p(N-1)+1}{N} r - 1\right)$, leads to the destabilization of the pure state of global excluders, which in turn results in the formation of a global family of neutral stable cycles moving anti-clockwise around a nonlinear center (see the bottom left panel of Fig. 2). This means that the emergence of cooperation can be realized in this case. A further increment of the cost of global exclusion to $c_{GE} = \frac{p(N-1)+1}{N} r - 1$ results in the transition of the evolutionary system from Phase xxi to Phase viii, wherein the fixed points along the $D - U$ segment and the $U - GE$ segment on the $D - GE$ edge of the simplex S_3 are stable and unstable, respectively (see the bottom right panel of Fig. 2). In other cases, the homogeneous state of defectors is the unique stable fixed point.

For $p(x_{GE})$ belonging to Class ii or Class iii satisfying $s \neq +\infty$, we have $h = 0$ and $p_{\min} \neq p_{\max}$, or $h = 1$ and $p_{\min} \neq p_{\max}$, or $h \in (0, 1)$ and $p_{\min} \neq p_{\max}$, which leaves s, $p_{\min}$, and $p_{\max}$ as the three remaining variables in Eq. (28). From a mathematical viewpoint, the qualitative difference among these three cases lies in the high order derivative of Eq. (28) with x_{GE}. For instance, $h = 0$ leads to the second-order derivative $p''(x_{GE}) < 0$; $h = 1$ results in $p''(x_{GE}) > 0$; if $h \in (0, 1)$, one can obtain $p''(x_{GE}) > 0$ for $x_{GE} \in [0, \bar{x}_{GE})$ and $p''(x_{GE}) < 0$ for $x_{GE} \in (\bar{x}_{GE}, 1]$, where $\bar{x}_{GE} = h$ denotes the inflection point satisfying $p''(\bar{x}_{GE}) = 0$. However, according to the theoretical analysis of the replicator dynamics in Sec. III A, the higher order derivatives of Eq. (28) with x_{GE}, e.g., the second-order derivative $p''(x_{GE})$, the position of the inflection point, i.e., h, as well as the steepness of the function at the inflection point, i.e., s ($\neq +\infty$), do not qualitatively affect the results of such deterministic dynamics, which means that one can consider all three cases at the same time. In all these cases, one can find that $p'(\bar{x}_{GE}) > 0$ is always satisfied, which leads to the theoretical predictions that the evolutionary system may exhibit the deterministic behaviors in Phase i if $c_{GE} = 0$ and $p_{\max} \in \left[0, \frac{N-r}{(N-1)r}\right]$, or in Phase iii if $c_{GE} = 0$, $p_{\max} \in \left(\frac{N-r}{(N-1)r}, 1\right]$ and

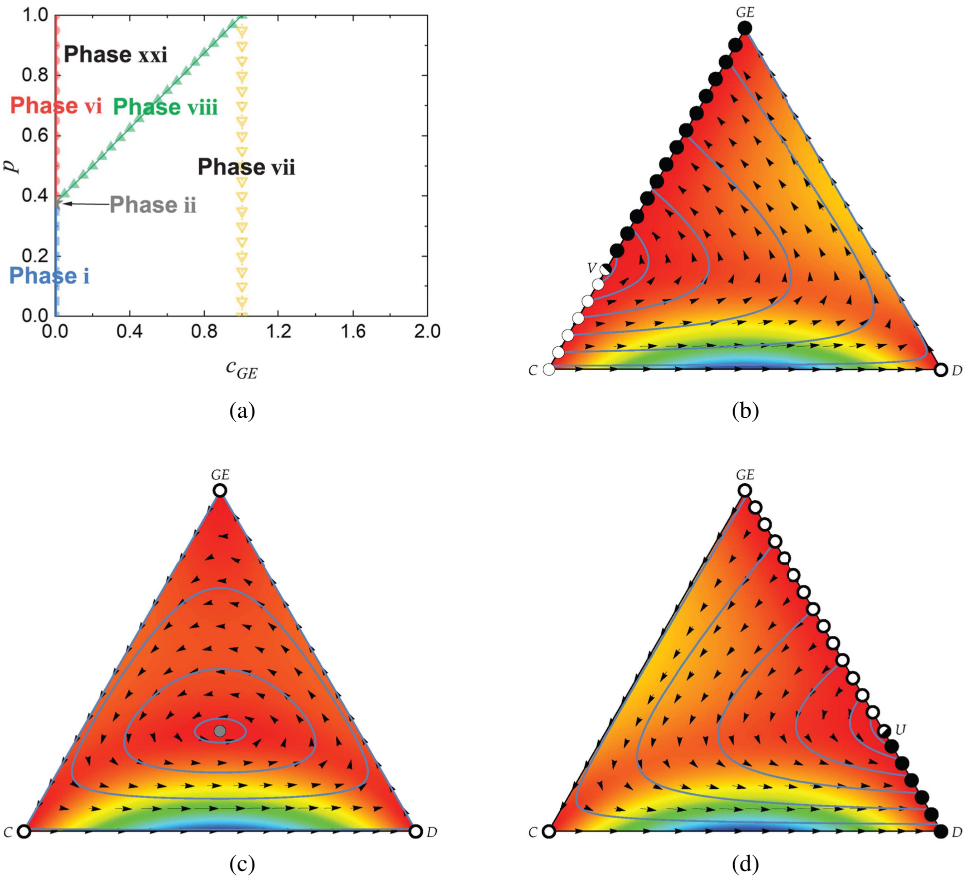


FIG. 2. Dynamics of public goods game with global exclusion for the construction probability of the centralized exclusive institution $p\,(x_{GE})$ belonging to Class i. Top left panel: Full $c_{GE} - p$ phase diagram as obtained by numerical simulations (filled symbols), which is in good agreement with that obtained by theoretical analysis (solid lines). The green filled triangles corresponding to Phase viii coincide with the green solid line predicted by the equation $p = \frac{N(c_{GE}+1)-r}{(N-1)r}$. The yellow hollow inverted triangles, which fits perfectly with the value predicted by the equation $c_{GE} = r - 1$ (see the vertical dotted line), denotes the critical cost of global exclusion c_{GE} beyond which the homogeneous state of defectors is the unique stable fixed point irrespective of the value of p. The remaining panels: Phase portraits for $c_{GE} = 0$ (top right panel), $c_{GE} = 0.1$ (bottom left panel), and $c_{GE} = 0.2$ (bottom right panel) at $p = 0.5$ corresponding respectively to Phase vi, Phase xxi, and Phase viii in the simplex S_3. The arrows show the directions of natural selection, and white (black) circles denote unstable (stable) fixed points. The gray circle represents an nonlinear center [i.e., the neutrally stable interior fixed point Q locating at $(x_D{}^*, x_C{}^*, x_{GE}{}^*) \approx (0.3535, 0.3535, 0.293)$ in the bottom left panel], while white circles with thicker edges saddle points. The half-black and half-white circle V (U) in the top right panel (bottom right panel), locating at $\left(x_D{}^+, x_C{}^+, x_{GE}{}^+\right) \approx (0, 0.7071, 0.2929)$ ($\left(x_D{}^-, x_C{}^-, x_{GE}{}^-\right) \approx (0.7071, 0, 0.2929)$), refers to a critical fixed point, which divides the edge $C - GE$ ($D - GE$) into two segments with stable fixed points on the segment $V - GE$ ($D - U$) and unstable fixed points on the segment $C - V$ ($U - GE$). Blue corresponds to fast dynamics and red to slow dynamics close to the fixed points of the system. The graphical outputs of the top right, bottom left, and bottom right panels are based on the Dynamo software.[25] Parameter settings: $N = 5$ and $r = 2$. For this special case, the emergence or even stabilization of public cooperation can be realized as long as the constant construction probability of the centralized exclusive institution and the cost of global exclusion are sufficiently large and low, respectively.

$p_{\min} \in \left[0, \frac{N-r}{(N-1)r}\right)$, or in Phase vi if $c_{GE} = 0$ and $p_{\min} \in \left[\frac{N-r}{(N-1)r}, 1\right]$, or in Phase vii if $c_{GE} \in (0, r-1)$ and $p_{\max} \in \left[0, \frac{N(c_{GE}+1)-r}{(N-1)r}\right]$, or if $c_{GE} \in [r-1, +\infty)$, or in Phase ix if $c_{GE} \in (0, r-1)$, $p_{\max} \in \left(\frac{N(c_{GE}+1)-r}{(N-1)r}, 1\right]$ and $p_{\min} \in \left[0, \frac{N(c_{GE}+1)-r}{(N-1)r}\right)$, or in Phase xviii if $c_{GE} \in (0, r-1)$ and $p_{\min} \in \left[\frac{N(c_{GE}+1)-r}{(N-1)r}, 1\right]$. The full $p_{\max} - p_{\min}$ phase diagrams obtained by numerical simulations confirm our theoretical predictions, as shown in Fig. 3. Interestingly, we do find the existence of unstable spirals approaching the heteroclinic limiting cycles on the boundary of the simplex S_3 in Phase xviii, as shown by a numerical example given in Fig. 4.

For the limiting case $s \to +\infty$ in Class ii or Class iii, $p\,(x_{GE})$ becomes a step function,

$$p\,(x_{GE}) = \begin{cases} p_{\min}, & x_{GE} \in [0, h), \\ (p_{\min} + p_{\max})/2, & x_{GE} = h, \\ p_{\max}, & x_{GE} \in (h, 1], \end{cases} \tag{30}$$

where h denotes the position of the step transition, and $p_{\min}$ and $p_{\max}$, satisfying $0 \le p_{\min} < p_{\max} \le 1$, represent the constant

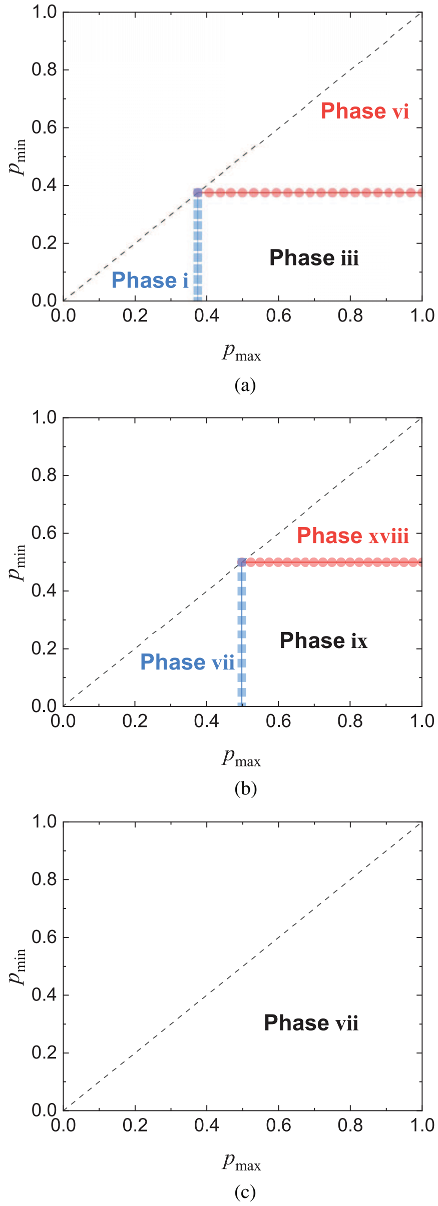


FIG. 3. Full $p_{\max} - p_{\min}$ phase diagrams of $c_{GE} = 0$ (top panel), $c_{GE} = 0.2 \in (0, r-1)$ (middle panel), and $c_{GE} = 1.5 \in [r-1, +\infty)$ (bottom panel) for $p\,(x_{GE})$ belonging to Class ii or Class iii satisfying $s \neq +\infty$ and $p_{\max} > p_{\min}$, as obtained by numerical simulations (filled symbols) as well as theoretical analysis (solid lines). Note that the phase separation here is meaningful only if the parameter region locates below the dashed line given by the equation $p_{\max} = p_{\min}$. The blue filled squares and red filled circles, which coincide perfectly with the blue vertical and red horizontal lines, respectively, given by the equations $p_{\max} = \frac{N(c_{GE}+1)-r}{(N-1)r}$ and $p_{\min} = \frac{N(c_{GE}+1)-r}{(N-1)r}$, denote phase boundaries between different phases in both top and middle panels. Parameter settings: $N = 5$ and $r = 2$. For this general case, the public cooperation can be maintained by either stabilizing itself or self-organizing the infinitely well-mixed population into the dynamics of global cyclic dominance between defectors, cooperators, and global excluders.

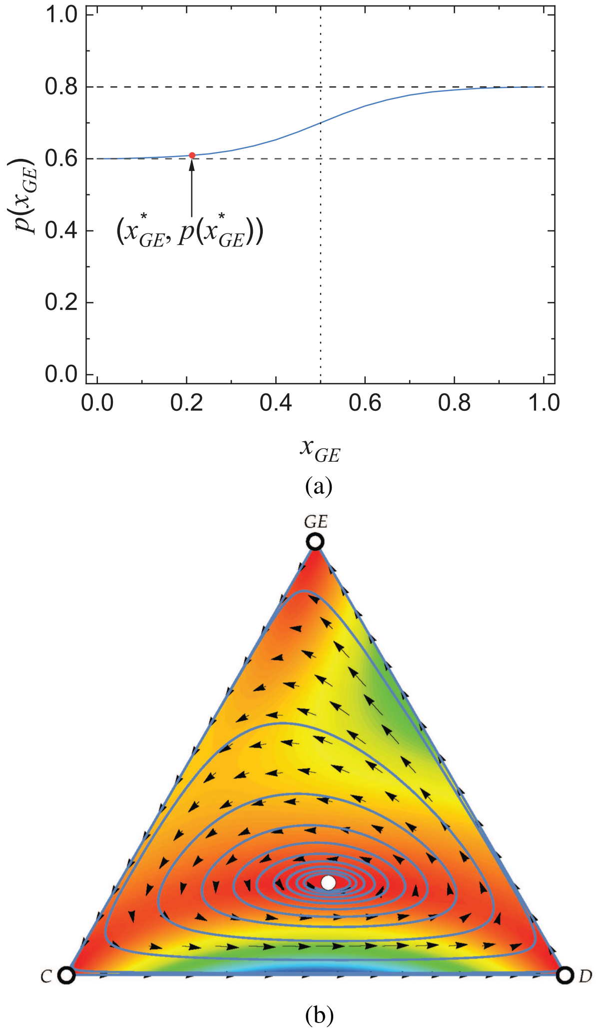


FIG. 4. A numerical example of the construction probability for the centralized exclusive institution $p\,(x_{GE})$ when $p_{\max} = 0.8$, $p_{\min} = 0.6$, $h = 0.5$, and $s = 10$, which belongs to the normalized logistic functions of Class iii resulting in the dynamical behaviors of Phase xviii. Top panel: Numerical example of $p\,(x_{GE})$ in dependence on the frequency of global excluders x_{GE}. The red solid circle marks the two-dimensional coordinate $(x_{GE}{}^*, p\,(x_{GE}{}^*)) \approx (0.2124, 0.6095)$ satisfying Eq. (18). The dotted vertical line given by the equation $x_{GE} = h$ defines the position of the inflection point for the function $p\,(x_{GE})$, while the top and bottom dashed horizontal lines correspond to the bounds of the function $p\,(x_{GE})$: $p\,(x_{GE}) = p_{\max}$ (upper bound) and $p\,(x_{GE}) = p_{\min}$ (lower bound). Bottom panel: Phase portrait in the simplex S_3 for the numerical example of $p\,(x_{GE})$ given in the top panel. The white circle represents an unstable fixed point [i.e., the interior unstable fixed point Q locating at $(x_D{}^*, x_C{}^*, x_{GE}{}^*) \approx (0.4199, 0.3677, 0.2124)$], while the ones with thicker edges the saddle points. Note that starting from any interior point except the interior unstable fixed point Q inside the simplex S_3, all unstable spirals will move anti-clockwise toward the stable heteroclinic limiting cycles on the boundary of the simplex S_3 in Phase xviii. The arrows show the directions of natural selection. Blue corresponds to fast dynamics and red to slow dynamics close to the fixed points of the system. The graphical output of the bottom panel is produced from an adapted version of the Dynamo software.[25] Parameter settings: $N = 5$, $r = 2$, and $c_{GE} = 0.2$.

values of the function $p(x_{GE})$ for $x_{GE}\in[0,h)$ and $x_{GE}\in(h,1]$, respectively. In this case, the theoretical analysis predicts that the dynamical behaviors of the evolutionary system may be in Phase i if $h\in[0,1)$, $c_{GE}=0$, and $p_{\max}\in\left[0,\frac{N-r}{(N-1)r}\right)$, or if $h=1$, $c_{GE}=0$, and $\frac{p_{\min}+p_{\max}}{2}\in\left[0,\frac{N-r}{(N-1)r}\right]$, or in Phase ii if $h\in[0,1)$, $c_{GE}=0$, and $p_{\max}=\frac{N-r}{(N-1)r}$, or in Phase iii if $h\in(0,1)$, $c_{GE}=0$, $p_{\max}\in\left(\frac{N-r}{(N-1)r},1\right]$, $p_{\min}\in\left[0,\frac{N-r}{(N-1)r}\right)$, and $p\left(x_{GE}=x_{GE}^{-}=h\right)=\frac{p_{\min}+p_{\max}}{2}=\frac{N-r}{(N-1)r}$, or in Phase v if $h\in(0,1)$, $c_{GE}=0$, $p_{\max}\in\left(\frac{N-r}{(N-1)r},1\right]$ and $p_{\min}=\frac{N-r}{(N-1)r}$, or if $h=1$, $c_{GE}=0$, $\frac{p_{\min}+p_{\max}}{2}\in\left(\frac{N-r}{(N-1)r},1\right]$ and $p_{\min}=\frac{N-r}{(N-1)r}$, or in Phase vi if $h\in(0,1]$, $c_{GE}=0$ and $p_{\min}\in\left(\frac{N-r}{(N-1)r},1\right]$, or if $h=0$, $c_{GE}=0$ and $\frac{p_{\min}+p_{\max}}{2}\in\left[\frac{N-r}{(N-1)r},1\right]$, or in Phase vii if $h\in[0,1)$, $c_{GE}\in(0,r-1)$ and $p_{\max}\in\left[0,\frac{N(c_{GE}+1)-r}{(N-1)r}\right)$, or if $h=1$, $c_{GE}\in(0,r-1)$ and $\frac{p_{\min}+p_{\max}}{2}\in\left[0,\frac{N(c_{GE}+1)-r}{(N-1)r}\right]$, or if $c_{GE}\in[r-1,+\infty)$, or in Phase viii if $h\in[0,1)$, $c_{GE}\in(0,r-1)$ and $p_{\max}=\frac{N(c_{GE}+1)-r}{(N-1)r}$, or in Phase xi if $h\in(0,x_{GE}{}^{*})$, $c_{GE}\in(0,r-1)$, $p_{\max}\in\left(\frac{N(c_{GE}+1)-r}{(N-1)r},1\right]$, $x_{GE}{}^{*}=1-\left[1-\frac{N-r}{rp_{\max}(N-1)}\right]^{\frac{1}{N-1}}$, $p_{\min}\in\left[0,\frac{N(c_{GE}+1)-r}{(N-1)r}\right)$ and $p\left(x_{GE}=x_{GE}^{-}=h\right)=\frac{p_{\min}+p_{\max}}{2}=\frac{N(c_{GE}+1)-r}{(N-1)r}$, or in Phase xvii if $h\in(0,x_{GE}{}^{*})$, $c_{GE}\in(0,r-1)$, $p_{\max}\in\left(\frac{N(c_{GE}+1)-r}{(N-1)r},1\right]$, $x_{GE}{}^{*}=1-\left[1-\frac{N-r}{rp_{\max}(N-1)}\right]^{\frac{1}{N-1}}$ and $p_{\min}=\frac{N(c_{GE}+1)-r}{(N-1)r}$, or in Phase xviii if $h=x_{GE}{}^{*}$, $c_{GE}\in(0,r-1)$, $p_{\min}\in\left(\frac{N(c_{GE}+1)-r}{(N-1)r},1\right]$ and $\frac{p_{\min}+p_{\max}}{2}=\frac{N-r}{(N-1)r\left[1-(1-x_{GE}{}^{*})^{N-1}\right]}\in\left(p_{\min},1\right)$, or in Phase xx if $h\in(0,x_{GE}{}^{*})$, $c_{GE}\in(0,r-1)$, $p_{\min}\in\left(\frac{N(c_{GE}+1)-r}{(N-1)r},1\right]$ and $p_{\max}=\frac{N-r}{(N-1)r\left[1-(1-x_{GE}{}^{*})^{N-1}\right]}\in\left(p_{\min},1\right]$, or if $h\in(x_{GE}{}^{*},1)$, $c_{GE}\in(0,r-1)$ and $p_{\min}=\frac{N-r}{(N-1)r\left[1-(1-x_{GE}{}^{*})^{N-1}\right]}\in\left(\frac{N(c_{GE}+1)-r}{(N-1)r},1\right]$, or if $h=0$, $c_{GE}\in(0,r-1)$, $\frac{p_{\min}+p_{\max}}{2}\in\left[\frac{N(c_{GE}+1)-r}{(N-1)r},1\right]$, or if $h=1$, $c_{GE}\in(0,r-1)$, $p_{\min}\in\left(\frac{N(c_{GE}+1)-r}{(N-1)r},1\right]$.

Due to the dissatisfaction of Eq. (30) with the general assumption made in Sec. II that $p(x_{GE})$ should be smooth for $x_{GE}\in[0,1]$, the evolutionary system may exhibit deterministic behaviors that do not belong to any phase classified in Subsection III A: It may be in Phase iii′, i.e., Phase iii but without the presence of the unique unstable fixed point U on the edge $D-GE$, if $h=0$, $c_{GE}=0$, $p_{\max}\in\left(\frac{N-r}{(N-1)r},1\right]$ and $\frac{p_{\min}+p_{\max}}{2}\in\left[0,\frac{N-r}{(N-1)r}\right)$, or if $h\in(0,1)$, $c_{GE}=0$, $p_{\max}\in\left(\frac{N-r}{(N-1)r},1\right]$, $p_{\min}\in\left[0,\frac{N-r}{(N-1)r}\right)$ and $p\left(x_{GE}=h\right)=\frac{p_{\min}+p_{\max}}{2}\neq\frac{N-r}{(N-1)r}$, or if $h=1$, $c_{GE}=0$, $\frac{p_{\min}+p_{\max}}{2}\in\left(\frac{N-r}{(N-1)r},1\right]$ and $p_{\min}\in\left[0,\frac{N-r}{(N-1)r}\right)$, or in Phase xi′, i.e., Phase xi but without the presence of the unique unstable fixed point U on the edge $D-GE$, if $h=0$, $c_{GE}\in(0,r-1)$, $p_{\max}\in\left(\frac{N(c_{GE}+1)-r}{(N-1)r},1\right]$ and $\frac{p_{\min}+p_{\max}}{2}\in\left[0,\frac{N(c_{GE}+1)-r}{(N-1)r}\right)$, or if $h\in(0,x_{GE}{}^{*})$, $c_{GE}\in(0,r-1)$, $p_{\max}\in\left(\frac{N(c_{GE}+1)-r}{(N-1)r},1\right]$, $x_{GE}{}^{*}=1-\left[1-\frac{N-r}{rp_{\max}(N-1)}\right]^{\frac{1}{N-1}}$, $p_{\min}\in\left[0,\frac{N(c_{GE}+1)-r}{(N-1)r}\right)$ and $p\left(x_{GE}=h\right)=\frac{p_{\min}+p_{\max}}{2}\neq\frac{N(c_{GE}+1)-r}{(N-1)r}$, or in Phase xi″, i.e., Phase xi but without the presence of the unique interior fixed point Q and the family of local closed orbits, if $h\in(x_{GE}{}^{*},1)$, $c_{GE}\in(0,r-1)$, $p_{\max}\in\left(\frac{N(c_{GE}+1)-r}{(N-1)r},1\right]$, $x_{GE}{}^{*}=1-\left[1-\frac{N-r}{rp_{\max}(N-1)}\right]^{\frac{1}{N-1}}$, $p_{\min}\in\left[0,\frac{N(c_{GE}+1)-r}{(N-1)r}\right)$ and $p\left(x_{GE}=h\right)=\frac{p_{\min}+p_{\max}}{2}=\frac{N(c_{GE}+1)-r}{(N-1)r}$, or if $h=x_{GE}{}^{*}$, $c_{GE}\in(0,r-1)$, $p_{\max}\in\left(\frac{N(c_{GE}+1)-r}{(N-1)r},1\right]$, $p_{\min}\in\left[0,\frac{N(c_{GE}+1)-r}{(N-1)r}\right)$ and $p\left(x_{GE}=h\right)=\frac{p_{\min}+p_{\max}}{2}=\frac{N(c_{GE}+1)-r}{(N-1)r}$, or in Phase xi‴, i.e., Phase xi but without the presence of the unique unstable fixed point U on the edge $D-GE$, the unique interior fixed point Q and the family of local closed orbits, if $h\in(x_{GE}{}^{*},1)$, $c_{GE}\in(0,r-1)$, $p_{\max}\in\left(\frac{N(c_{GE}+1)-r}{(N-1)r},1\right]$, $x_{GE}{}^{*}=1-\left[1-\frac{N-r}{rp_{\max}(N-1)}\right]^{\frac{1}{N-1}}$, $p_{\min}\in\left[0,\frac{N(c_{GE}+1)-r}{(N-1)r}\right)$ and $p\left(x_{GE}=h\right)=\frac{p_{\min}+p_{\max}}{2}\neq\frac{N(c_{GE}+1)-r}{(N-1)r}$, or if $h=x_{GE}{}^{*}$, $c_{GE}\in(0,r-1)$, $p_{\max}\in\left(\frac{N(c_{GE}+1)-r}{(N-1)r},1\right]$, $p_{\min}\in\left[0,\frac{N(c_{GE}+1)-r}{(N-1)r}\right)$, $x_{GE}{}^{*}\neq1-\left[1-\frac{2(N-r)}{r\left(p_{\min}+p_{\max}\right)(N-1)}\right]^{\frac{1}{N-1}}$ and $p\left(x_{GE}=h\right)=\frac{p_{\min}+p_{\max}}{2}\neq\frac{N(c_{GE}+1)-r}{(N-1)r}$, or in Phase xi⁗, i.e., Phase xi but without the presence of the unique unstable fixed point U on the edge $D-GE$ and the family of local closed orbits, if $h=x_{GE}{}^{*}$, $p_{\max}\in\left(\frac{N(c_{GE}+1)-r}{(N-1)r},1\right]$, $p_{\min}\in\left[0,\frac{N(c_{GE}+1)-r}{(N-1)r}\right)$ and $x_{GE}{}^{*}=1-\left[1-\frac{2(N-r)}{r\left(p_{\min}+p_{\max}\right)(N-1)}\right]^{\frac{1}{N-1}}$, or in Phase xvii′, i.e., Phase xvii but without the presence of the unique interior fixed point Q and the family of local closed orbits, if $h\in(x_{GE}{}^{*},1)$, $c_{GE}\in(0,r-1)$, $p_{\max}\in\left(\frac{N(c_{GE}+1)-r}{(N-1)r},1\right]$, $x_{GE}{}^{*}=1-\left[1-\frac{N-r}{rp_{\max}(N-1)}\right]^{\frac{1}{N-1}}$ and $p_{\min}=\frac{N(c_{GE}+1)-r}{(N-1)r}$, or if $h=x_{GE}{}^{*}$, $c_{GE}\in(0,r-1)$, $p_{\max}\in\left(\frac{N(c_{GE}+1)-r}{(N-1)r},1\right]$, $p_{\min}=\frac{N(c_{GE}+1)-r}{(N-1)r}$ and $x_{GE}{}^{*}\neq1-\left[1-\frac{2(N-r)}{r\left(p_{\min}+p_{\max}\right)(N-1)}\right]^{\frac{1}{N-1}}$, or in Phase xvii″, i.e., Phase xvii but without the presence of the family of local closed orbits, if $h=x_{GE}{}^{*}$, $c_{GE}\in(0,r-1)$, $p_{\max}\in\left(\frac{N(c_{GE}+1)-r}{(N-1)r},1\right]$, $p_{\min}=\frac{N(c_{GE}+1)-r}{(N-1)r}$ and $x_{GE}{}^{*}=1-\left[1-\frac{2(N-r)}{r\left(p_{\min}+p_{\max}\right)(N-1)}\right]^{\frac{1}{N-1}}$, or in Phase xviii′, i.e., Phase xviii but without the presence of the unique interior fixed point Q, if $h=x_{GE}{}^{*}$, $c_{GE}\in(0,r-1)$, $p_{\min}\in\left(\frac{N(c_{GE}+1)-r}{(N-1)r},1\right]$ and $x_{GE}{}^{*}\neq1-\left[1-\frac{2(N-r)}{r\left(p_{\min}+p_{\max}\right)(N-1)}\right]^{\frac{1}{N-1}}$, or in Phase xx′, i.e., Phase xx but without the presence of the unique interior fixed point Q and the family of local closed orbits, if $h\in(0,x_{GE}{}^{*})$, $c_{GE}\in(0,r-1)$, $p_{\min}\in\left(\frac{N(c_{GE}+1)-r}{(N-1)r},1\right]$ and $x_{GE}{}^{*}\neq1-\left[1-\frac{N-r}{rp_{\max}(N-1)}\right]^{\frac{1}{N-1}}$, or if $h\in(x_{GE}{}^{*},1)$, $c_{GE}\in(0,r-1)$, $p_{\min}\in\left(\frac{N(c_{GE}+1)-r}{(N-1)r},1\right]$ and $x_{GE}{}^{*}\neq1-\left[1-\frac{N-r}{rp_{\min}(N-1)}\right]^{\frac{1}{N-1}}$, or in Phase ix′, i.e., Phase ix but without the presence of the unique unstable fixed point U on the

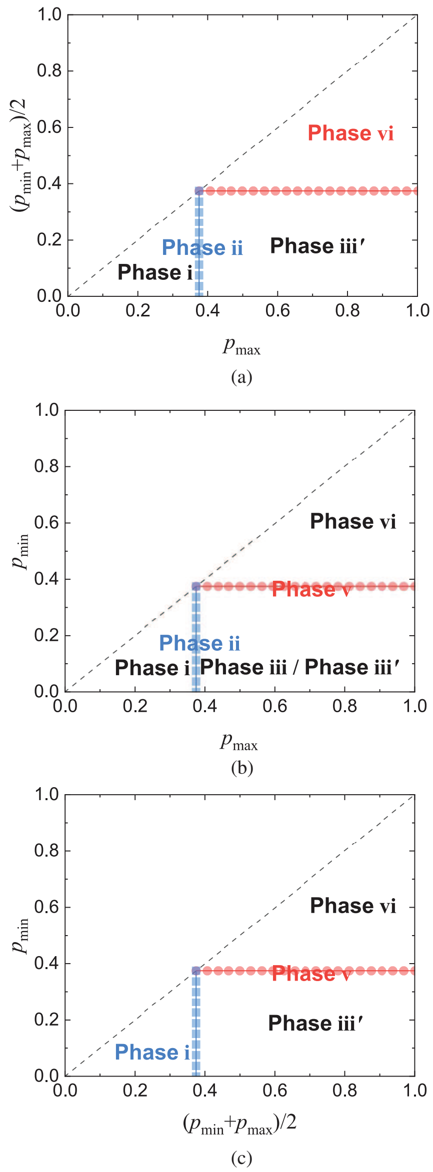


FIG. 5. Full $p_{max} - (p_{min} + p_{max})/2$, $p_{max} - p_{min}$ and $(p_{min} + p_{max})/2 - p_{min}$ phase diagrams of $h = 0$ (top panel), $h = 0.5 \in (0, 1)$ (middle panel), and $h = 1$ (bottom panel) at $c_{GE} = 0$ for $p\,(x_{GE})$ belonging to Class ii or Class iii satisfying $s \to +\infty$ and $p_{max} > p_{min}$, as obtained by numerical simulations (filled symbols) as well as theoretical analysis (solid lines). Note that the phase separation in the top, middle, and bottom panels is meaningful only if the parameter region locates below the dashed line given by the equations $p_{max} = \frac{p_{min}+p_{max}}{2}$, $p_{max} = p_{min}$, and $\frac{p_{min}+p_{max}}{2} = p_{min}$, respectively. The blue filled squares and red filled circles in all three panels, which coincide perfectly with the blue vertical and red horizontal lines respectively given by the equations p_{max} or $\frac{p_{min}+p_{max}}{2} = \frac{N-r}{(N-1)r}$ and $\frac{p_{min}+p_{max}}{2}$ or $p_{min} = \frac{N-r}{(N-1)r}$, denote phase boundaries between different phases. Parameter settings: $N = 5$ and $r = 2$. For this limiting case, when global exclusion is costless, there emerges a new phase, i.e., Phase iii′, in which the evolutionary system exhibits a novel behavior: the separation of attraction basins for different stable fixed points by the transition point h in the simplex S_3.

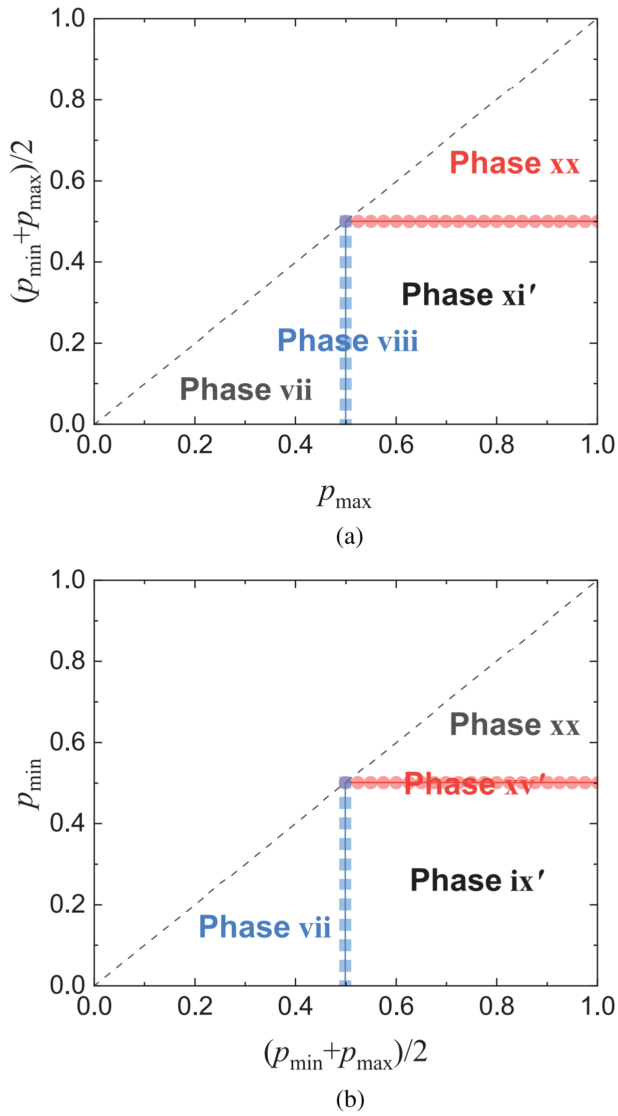


FIG. 6. Full $p_{max} - (p_{min} + p_{max})/2$ and $(p_{min} + p_{max})/2 - p_{min}$ phase diagrams of $h = 0$ (top panel) and $h = 1$ (bottom panel) at $c_{GE} = 0.2$ for $p\,(x_{GE})$ belonging to Class ii satisfying $s \to +\infty$ and $p_{max} > p_{min}$, as obtained by numerical simulations (filled symbols) as well as theoretical analysis (solid lines). Note that the phase separation in the phase diagrams is meaningful only if the parameter region locates below the dashed line given by the equations $p_{max} = \frac{p_{min}+p_{max}}{2}$ and $\frac{p_{min}+p_{max}}{2} = p_{min}$ for $h = 0$ (top panel) and $h = 1$ (bottom panel), respectively. The blue filled squares and red filled circles in both panels, which coincide perfectly with the blue vertical and red horizontal lines respectively given by the equations p_{max} or $\frac{p_{min}+p_{max}}{2} = \frac{N(c_{GE}+1)-r}{(N-1)r}$ and $\frac{p_{min}+p_{max}}{2}$ or $p_{min} = \frac{N(c_{GE}+1)-r}{(N-1)r}$, denote phase boundaries between different phases. Parameter settings: $N = 5$ and $r = 2$. For this limiting case, when the cost of global exclusion is moderate, the evolutionary system may stay in three newly emerging phases, including Phase ix′, Phase xi′, and Phase xv′, in which it exhibits novel behaviors. For instance, in Phase xi′, there co-occur a family of closed orbits in the local neighborhood of an nonlinear center Q as well as other evolutionary trajectories moving toward the stable fixed point D in the simplex S_3.

edge $D-GE$ and the unique interior fixed point Q, if $h=1$, $c_{GE} \in (0, r-1)$, $\frac{p_{\min}+p_{\max}}{2} \in \left(\frac{N(c_{GE}+1)-r}{(N-1)r}, 1\right]$ and $p_{\min} \in \left[0, \frac{N(c_{GE}+1)-r}{(N-1)r}\right)$, or in Phase xv′, i.e., Phase xv but without the presence of the unique interior fixed point Q, if $h=1$, $c_{GE} \in (0, r-1)$, $\frac{p_{\min}+p_{\max}}{2} \in \left(\frac{N(c_{GE}+1)-r}{(N-1)r}, 1\right]$ and $p_{\min} = \frac{N(c_{GE}+1)-r}{(N-1)r}$. The full $p_{\max} - (p_{\min}+p_{\max})/2$, $p_{\max}-p_{\min}$ and $(p_{\min}+p_{\max})/2 - p_{\min}$ phase diagrams at $c_{GE}=0$ and $c_{GE} \in (0, r-1)$ obtained respectively for $h=0$, $h \in (0,1)$ and $h=1$ via numerical simulations confirm our theoretical predictions in this limiting case, i.e., $s \to +\infty$, as shown in Figs. 5–7. Note that, whenever $c_{GE} \in [r-1, +\infty)$, the dynamical behaviors of the evolutionary system for $s \to +\infty$ also inevitably fall into the category of Phase vii. Due to the discontinuity of the construction probability for the centralized exclusive institution $p\,(x_{GE})$, one can observe some intriguing behaviors of the evolutionary system, such as the

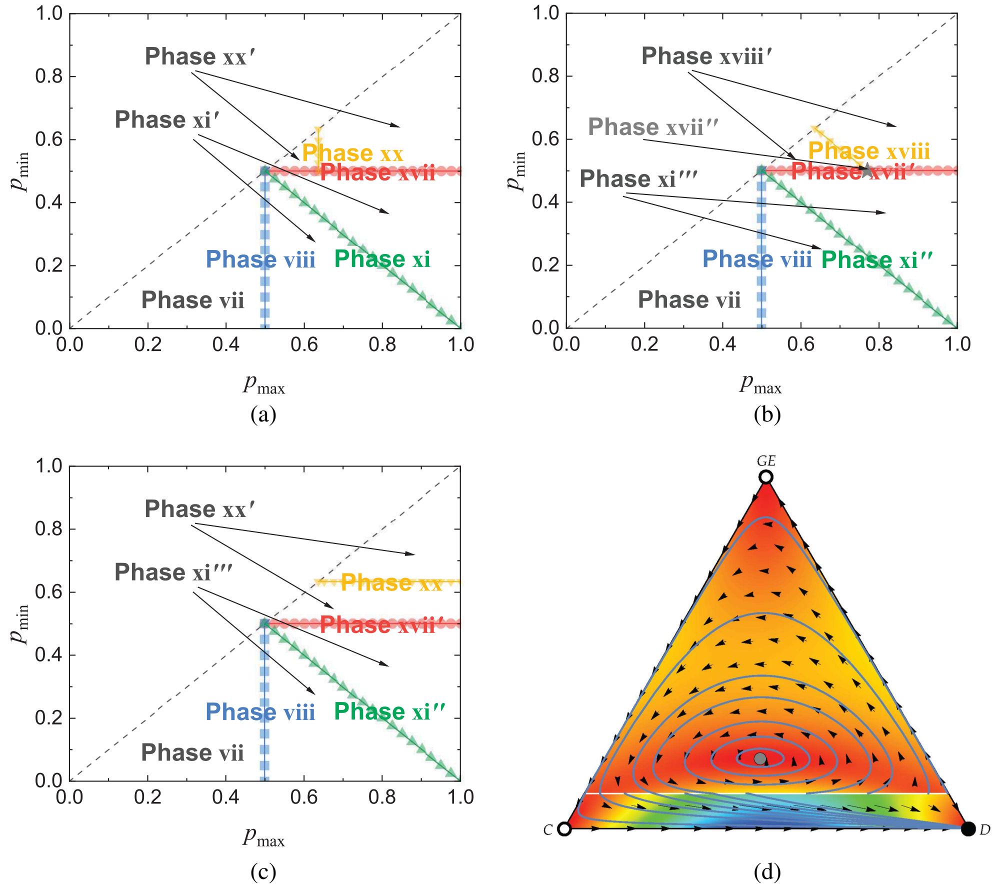
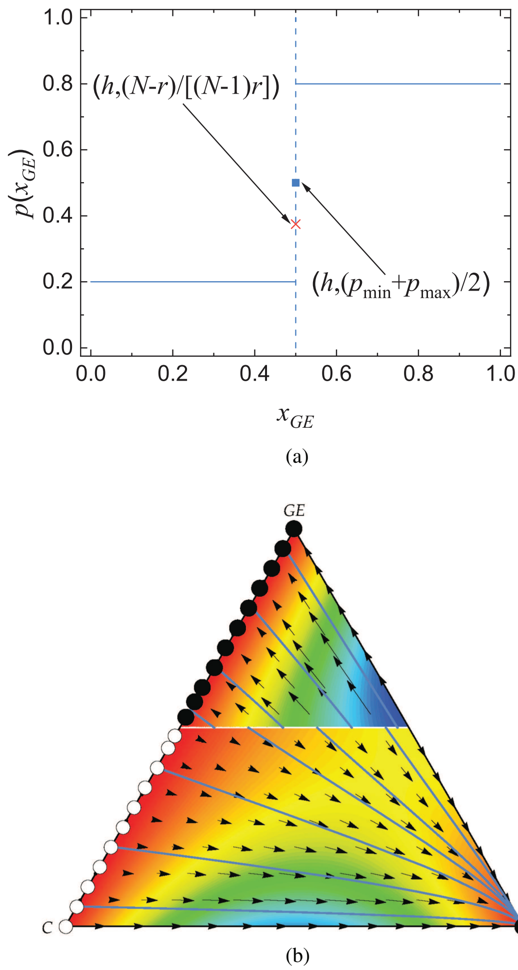


FIG. 7. Full $p_{\max}-p_{\min}$ phase diagrams of $h \in (0, x_{GE}{}^*)$ (top left panel), $h = x_{GE}{}^*$ (top right panel), and $h \in (x_{GE}{}^*, 1)$ (bottom left panel) at $c_{GE}=0.2$ and $x_{GE}{}^*=0.2$ for $p\,(x_{GE})$ belonging to Class iii satisfying $s \to +\infty$ and $p_{\max} > p_{\min}$, as obtained by numerical simulations (filled symbols) as well as theoretical analysis (solid lines). Note that the phase separation in the phase diagrams is meaningful only if the parameter region locates below the dashed line given by the equation $p_{\max}=p_{\min}$. The blue filled squares, red filled circles, green filled triangles, and the yellow filled inverted triangles in all phase diagrams, which coincide perfectly with the blue vertical, red horizontal, green diagonal, and yellow lines, respectively, given by the equations $p_{\max} = \frac{N(c_{GE}+1)-r}{(N-1)r}$, $p_{\min} = \frac{N(c_{GE}+1)-r}{(N-1)r}$, $\frac{p_{\min}+p_{\max}}{2} = \frac{N(c_{GE}+1)-r}{(N-1)r}$ and $p_{\max}$, $\frac{p_{\min}+p_{\max}}{2}$ or $p_{\min} = \frac{N-r}{(N-1)r\left[1-(1-x_{GE}{}^*)^{N-1}\right]}$, denote phase boundaries between different phases. For this limiting case, much more new phases, including Phase xi′, Phase xi″, Phase xi‴, Phase xvii′, Phase xvii″, Phase xviii′, and Phase xx′, emerge in the evolutionary system. The bottom right panel shows a phase portrait for a numerical example of the construction probability for the centralized exclusive institution $p\,(x_{GE})$ when $p_{\max}=625/984 \approx 0.6352$, $p_{\min}=0.2$, $h=0.1$, and $s \to +\infty$ leading to the dynamical behaviors of Phase xi′. The white circles with thicker edges represent saddle points, while the filled black (gray) one the stable fixed point [the nonlinear center, i.e., the interior fixed point Q locating at $(x_D{}^*, x_C{}^*, x_{GE}{}^*) \approx (0.3844, 0.4156, 0.2)$]. Because $\frac{p_{\min}+p_{\max}}{2} = 4109/9840 \approx 0.4176 < \frac{N(c_{GE}+1)-r}{(N-1)r} = 0.5$, instead of any interior fixed point existing on the edge $D-GE$, there emerges a discontinuous transition for the evolutionary speed along the white solid line given by $x_{GE}=h=0.1$. Besides a family of closed orbits revolving anti-clockwise around the nonlinear center Q in its local neighborhood, all other evolutionary trajectories would converge to the stable fixed point D eventually. The arrows show the directions of natural selection. Blue corresponds to fast dynamics and red to slow dynamics close to the fixed points of the system. The graphical output of the bottom right panel is produced from an adapted version of the Dynamo software.[25] Parameter settings: $N=5$ and $r=2$.



(a)

(b)

FIG. 8. A numerical example of the construction probability for the centralized exclusive institution $p\,(x_{GE})$ when $p_{\max} = 0.8$, $p_{\min} = 0.2$, $h = 0.5$, and $s \to +\infty$, which becomes the Heaviside step function resulting in the dynamical behaviors of Phase iii′. Top panel: Numerical example of $p\,(x_{GE})$ in dependence on the frequency of global excluders x_{GE}. The blue filled square marking the two-dimensional coordinate $(h, (p_{\min} + p_{\max})/2) = (0.5, 0.5)$ does not coincide with the red filled cross locating at $(h, (N - r)/[(N - 1)\,r]) = (0.5, 0.375)$, which indicates that the function value of $p\,(x_{GE})$ at $h = 0.5$ does not satisfy Eq. (12). The dashed vertical line given by the equation $x_{GE} = h = 0.5$ defines the transition point for the Heaviside step function $p\,(x_{GE})$. Bottom panel: Phase portrait in the simplex S_3 for the numerical example of $p\,(x_{GE})$ given in the top panel. The white circles represent unstable fixed points, while the filled ones the stable fixed points. Due to the non-overlap between the blue filled square and the red filled cross in the upper panel, i.e., $\frac{p_{\min}+p_{\max}}{2} = 0.5 > \frac{N-r}{(N-1)r} = 0.375$, instead of any interior fixed point emerging on the edge $D - GE$, there exists a white solid line given by $x_{GE} = h = 0.5$ dividing the simplex S_3 into two regions of attraction basins: (1) All evolutionary trajectories originating from the population states belonging to the parameter region above the white solid line (including itself) would move toward the stable fixed points on the edge $C - GE$. (2) The population starting from the states belonging to the parameter area below the white solid line would evolve to the stable fixed point D. The arrows show the directions of natural selection. Blue corresponds to fast dynamics and red to slow dynamics close to the fixed points of the system. The graphical output of the bottom panel is based on the Dynamo software.[25] Parameter settings: $N = 5$, $r = 2$, and $c_{GE} = 0$.

separation of attraction basins for different stable fixed points by a line in the simplex S_3 determined by the position of the step transition $p\,(x_{GE}) = h$, instead of unstable fixed points typically (see Fig. 8), or the local destruction of the global cyclic dominance between defectors, cooperators, and global excluders, as well as the concurrent introduction of the stable fixed point at the vertex D of the simplex S_3 (see the bottom right panel of Fig. 7).

IV. CONCLUSION AND DISCUSSIONS

In sum, we have constructed an evolutionary model of public goods game with global exclusion and have theoretically studied the replicator dynamics among defectors, cooperators as well as global excluders in an infinite and well-mixed population. While defectors and cooperators in our model play the same roles as they do in the classic public goods game, global excluders refer to such a type of individuals who contribute not only to the common pool for the purpose of producing public goods but also to the collective construction of a centralized exclusive institution so as to coercively exclude free-riders from sharing the common goods. By analyzing the dynamical behaviors of public goods game with global exclusion, we have shown that global exclusion is able to induce the emergence or even stabilization of cooperation in spite of its facing both first- and second-order free-riding problems. Despite being somewhat an unrealistic condition, i.e., $c_{GE} = 0$ and $p\,(x_{GE} = 1) \in \left(\frac{N-r}{(N-1)r}, 1\right]$, for the stabilization of cooperation, we do find that the emergence of cooperation, achieved by the spontaneous formation of cyclic dominance between defectors, cooperators, and global excluders, can be observed in more realistic as well as considerably larger parameter regions, e.g., $c_{GE} \in (0, r-1)$ and $p\,(x_{GE}) = p\,(x_{GE}{}^*) \in \left(\frac{N(c_{GE}+1)-r}{(N-1)r}, 1\right]$ satisfied for $x_{GE} \in [x_{GE}{}^* - \Delta_3, x_{GE}{}^* + \Delta_4]$ with $\Delta_3 \in (0, x_{GE}{}^*)$ and $\Delta_4 \in (0, 1 - x_{GE}{}^*)$, the conclusion of which thus deserves to be further testified by empirical studies in the future. Besides, the evolutionary model of public goods game with global exclusion, despite its simplicity, can give rise to rich, i.e., 21 classes in total, dynamical behaviors. For instance, we have demonstrated the co-occurrence of a stable heteroclinic cycle on the boundary of the simplex S_3 and a family of closed orbits in the local neighborhood of the nonlinear center Q^* in Phase xx.

Note that, different from many previous works, we introduce the extension of model, i.e., global exclusion, into the public goods game by designing $p\,(x_{GE})$ as a general non-decreasing function of x_{GE} but not by prescribing some specific class of functions, which allows us to consider its arbitrary function forms, and, therefore, to provide a more general theoretical framework for studying the impacts of global exclusion in the evolutionary dynamics of public cooperation. Herein, we would like to point out that such a manner of model construction is necessary in the sense that only in this way can we reveal the full picture for the replicator dynamics of public goods game with global exclusion. Interestingly, it has been shown that the dynamical behaviors of the evolutionary system do not qualitatively change with the specific shape of $p\,(x_{GE})$ but are merely dependent on the values of $p\,(x_{GE})$ at some key positions of x_{GE} as well as in their neighborhoods (see the critical conditions for the classification of the replicator dynamics in Subsection III A) if all other conditions being equal. Besides, thanks also to the general way of model construction for global exclusion and the following theoretical analysis of its replicator dynamics in our work, we can make some predictions for many relevant studies on social dilemma games with exclusion. For example, we argue that the emergence of a family of closed orbits in the whole neighborhood of the nonlinear center claimed in Ref. 19 needs rather strict condition, i.e., requirement of the deterministic exclusion toward defectors for any particular groups involving any pool excluders, which is also an implicit assumption made in this study. Otherwise, one cannot observe the oscillatory coexistence between defectors, cooperators, and pool excluders in the whole neighborhood of the nonlinear center. Note that peer excluders in Ref. 15, unlike pool and global excluders, can stabilize itself in rather reasonable and larger parameter regions owing to their equal performance of evolution with that of second-order free-riders, i.e., cooperators, whenever defectors are absent in the population. On the other hand, from a broader perspective, institutional punishment, similar to global exclusion in our model, represents another form of negative incentives acting at a global scale. In one study, the institutional punishment is assumed to be successfully constructed with certainty and to be independent on the composition of the population, which leads to the observation that cooperation can completely dominate the infinite and well-mixed population even in the classic public goods game as long as the negative incentive is sufficiently large.[21] In another study, the construction of the punishment institution depends deterministically on a threshold number of punishers: If the number of punishers in the finite and well-mixed population is no less than such a threshold, the construction of the punishment institution succeeds; otherwise, the construction of the punishment institution fails.[24] In contrast with both studies mentioned above, we consider any forms of positive or no feedback between the state of the evolutionary system, i.e., x_{GE}, and that of the centralized institution determined by $p\,(x_{GE})$, which thus can make our model cover more cases in reality.

Furthermore, some other relevant studies have also considered the coupled effects between global exclusion and prior agreement with posterior compensations not only in the general model of public goods game but also in more specific scenarios, such as the issue of AI safety.[26,27] In general, regardless of the underlying backgrounds, one can observe a positive interplay between these two mechanisms, which indicates the potential synergistic impacts of global exclusion with other evolutionary rules in the resolution of public cooperation dilemma. Moreover, considering the fact that there are numerous ways to implement institutional incentives, such as institutional rewards, punishment as well as exclusion, it is thus interesting to ask which manner is more cost-efficient than other ones. Actually, such optimization problems have already been considered and studied in the context of both deterministic and stochastic evolutionary dynamics of social dilemma games with institutional rewards and punishment as the two alternative ways of incentives.[28,29] In parallel, a natural question arising here can be that whether the global exclusion proposed in this study can be more cost-efficient than both institutional rewards and punishment. Generally speaking, one cannot make a direct comparison of cost-efficiency between global exclusion and institutional rewards as well as punishment because global exclusion results in not only the reduction of payoff for defectors and the increment of payoff for cooperators in mixed groups but also the different degree of its efficiency in any different heterogeneous groups, which means that it becomes very difficult to compare the performance of the three forms of institutional incentives under the condition of the same efficiency ratio of corresponding incentives.[29] Note that only in this way can we unbiasedly evaluate the cost-efficiency of global exclusion, institutional rewards, and punishment. In future studies, it deserves to make much more efforts to further investigate this topic. Besides, numerous empirical studies have shown the existence of antisocial incentives, such as antisocial punishment and antisocial rewards.[30,31] Motivated by the experimental results, several theoretical works have studied the evolutionary dynamics of cooperation when both prosocial and antisocial incentives are present and have shown that antisocial incentives will, at least, weaken the positive effects of prosocial incentives on the evolution of public cooperation in the well-mixed population.[32–34] However, whenever the interaction range of individuals is limited, network reciprocity can restore (or even promote) the effectiveness of prosocial incentives by escaping from the adverse effects of antisocial incentives.[35,36] Thus, we expect that similar results can be obtained if antisocial global exclusion is also presented in our model. Finally, we also expect global exclusion can play a positive role in other types of multi-player social dilemma games, such as the multi-player snowdrift game and the multi-player stag-hunt game, as well as in structured populations, as it does in the present model. Works along this line are in progress.

ACKNOWLEDGMENTS

X.W. acknowledges funding from the National Natural Science Foundation of China (Grant No. 61903077), The Science and Technology Commission of Shanghai Municipality (2JC1401400), and from the Shanghai Sailing Program (Grant No. 19YF1402500). M.P. was supported by the Slovenian Research Agency (Grant Nos. P1-0403 and J1-9112). We thank Arne Traulsen for kindly providing an early version of the Dynamo software, which is modified to produce phase portraits presented in this paper.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Xiaofeng Wang: Conceptualization (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Software (equal); Validation (equal); Visualization (equal); Writing – original draft (equal); Writing – review and editing (equal). **Matjaž Perc:** Conceptualization (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Project administration (equal); Supervision (equal); Validation (equal); Writing – original draft (equal); Writing – review and editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

APPENDIX: ADVANCED STABILITY ANALYSIS OF THE HOMOGENEOUS D STATE

Because $x_C = 1 - x_{GE} - x_D$, Eq. (10) can be simplified as

$$\begin{cases} \dot{x}_{GE} = x_{GE}\left\{(1-x_{GE})\left[E(P_{GE}) - E(P_C)\right] - x_D\left[E(P_D) - E(P_C)\right]\right\}, & \text{(A1a)} \\ \dot{x}_D = x_D\left\{(1-x_D)\left[E(P_D) - E(P_C)\right] - x_{GE}\left[E(P_{GE}) - E(P_C)\right]\right\}, & \text{(A1b)} \end{cases}$$

where

$$E(P_{GE}) - E(P_C) = rp(x_{GE})\frac{(N-1)x_D(1-x_{GE})^{N-2}}{N} - c_{GE} \tag{A2}$$

and

$$E(P_D) - E(P_C) = 1 - \frac{r}{N} - rp(x_{GE})\left(1 - \frac{(N-1)(1-x_{GE})^{N-1}+1}{N}\right). \tag{A3}$$

Let

$$\begin{cases} f(x_{GE}, x_D) = x_{GE}\left\{(1-x_{GE})\left[E(P_{GE}) - E(P_C)\right] - x_D\left[E(P_D) - E(P_C)\right]\right\}, & \text{(A4a)} \\ g(x_{GE}, x_D) = x_D\left\{(1-x_D)\left[E(P_D) - E(P_C)\right] - x_{GE}\left[E(P_{GE}) - E(P_C)\right]\right\}, & \text{(A4b)} \end{cases}$$

one can then obtain the Jacobian matrix of the evolutionary system,

$$J = \begin{bmatrix} \frac{\partial f(x_{GE},x_D)}{\partial x_{GE}} & \frac{\partial f(x_{GE},x_D)}{\partial x_D} \\ \frac{\partial g(x_{GE},x_D)}{\partial x_{GE}} & \frac{\partial g(x_{GE},x_D)}{\partial x_D} \end{bmatrix}, \tag{A5}$$

where

$$\begin{cases} \frac{\partial f(x_{GE},x_D)}{\partial x_{GE}} = -(1-2x_{GE})c_{GE} - x_D\left(1-\frac{r}{N}\right) + \frac{N-1}{N}x_D r\left[p(x_{GE}) + x_{GE}p'(x_{GE})\right], & \text{(A6a)} \\ \frac{\partial f(x_{GE},x_D)}{\partial x_D} = x_{GE}\left(\frac{N-1}{N}rp(x_{GE}) - 1 + \frac{r}{N}\right), & \text{(A6b)} \\ \frac{\partial g(x_{GE},x_D)}{\partial x_{GE}} = x_D\begin{Bmatrix} c_{GE} - \frac{N-1}{N}(1-x_{GE})^{N-3}rp(x_{GE})\left[(N-1)(1-x_{GE}-x_D) + x_D\right] \\ +\frac{N-1}{N}rp'(x_{GE})\left[(1-x_{GE})^{N-2}(1-x_{GE}-x_D) - (1-x_D)\right] \end{Bmatrix}, & \text{(A6c)} \\ \frac{\partial g(x_{GE},x_D)}{\partial x_D} = (1-2x_D)\left(1-\frac{r}{N}\right) + x_{GE}c_{GE} + \frac{N-1}{N}rp(x_{GE})\begin{bmatrix} 2x_D - 1 \\ +(1-x_{GE})^{N-2}(1-x_{GE}-2x_D) \end{bmatrix}. & \text{(A6d)} \end{cases}$$

For the pure D state, the elements of Jacobian matrix is thus given by

$$\begin{cases} \left.\frac{\partial f(x_{GE},x_D)}{\partial x_{GE}}\right|_{(1,0,0)} = -c_{GE} - 1 + \frac{r}{N} + \frac{N-1}{N}rp(x_{GE}=0), & \text{(A7a)} \\ \left.\frac{\partial f(x_{GE},x_D)}{\partial x_D}\right|_{(1,0,0)} = 0, & \text{(A7b)} \\ \left.\frac{\partial g(x_{GE},x_D)}{\partial x_{GE}}\right|_{(1,0,0)} = c_{GE} - \frac{N-1}{N}rp(x_{GE}=0), & \text{(A7c)} \\ \left.\frac{\partial g(x_{GE},x_D)}{\partial x_D}\right|_{(1,0,0)} = -1 + \frac{r}{N}. & \text{(A7d)} \end{cases}$$

Therefore, the two eigenvalues of the Jacobian matrix at $(1, 0, 0)$ are equal to

$$\begin{cases} \lambda_1|_{(1,0,0)} = -1 + \frac{r}{N} < 0, & \text{(A8a)} \\ \lambda_2|_{(1,0,0)} = -c_{GE} - 1 + \frac{r}{N} + \frac{N-1}{N}rp(x_{GE}=0), & \text{(A8b)} \end{cases}$$

respectively. From Eq. (A8), one can find that the linearization method fails to determine the stability of the pure D state if $c_{GE} \in [0, r-1)$ and $p(x_{GE}=0) = \frac{N(c_{GE}+1)-r}{(N-1)r}$, or if $c_{GE} = r-1$ and $p(x_{GE}=0) = 1$. In the following, we aim to analyze the stability of the homogeneous D state by applying the center manifold theorem in these two critical cases.[37]

For the first case, i.e., $c_{GE} \in [0, r-1)$ and $p\left(x_{GE}=0\right) = \frac{N(c_{GE}+1)-r}{(N-1)r}$, the Jacobian matrix at the fixed point $(1,0,0)$ becomes

$$J|_{(1,0,0)} = \begin{bmatrix} 0 & 0 \\ \frac{r}{N}-1 & \frac{r}{N}-1 \end{bmatrix}, \tag{A9}$$

the eigenvalues of which are 0 and $\frac{r}{N}-1$, respectively. Let M be a matrix whose columns are the eigenvectors of $J|_{(1,0,0)}$; i.e.,

$$M = \begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix}. \tag{A10}$$

Take $T = M^{-1} = M$, we have

$$TJ|_{(1,0,0)}T^{-1} = \begin{bmatrix} 0 & 0 \\ 0 & \frac{r}{N}-1 \end{bmatrix}. \tag{A11}$$

The change of variables

$$\begin{bmatrix} y \\ z' \end{bmatrix} = T\begin{bmatrix} x_{GE} \\ x_D \end{bmatrix} = \begin{bmatrix} -x_{GE} \\ x_{GE}+x_D \end{bmatrix} \tag{A12}$$

puts the system into the form

$$\begin{cases} \dot{y} = y\left\{\begin{aligned}&(1+y)\left[rp(-y)\frac{(N-1)(z'+y)(1+y)^{N-2}}{N} - c_{GE}\right] \\ &-(z'+y)\left[1-\frac{r}{N} - rp(-y)\left(1-\frac{(N-1)(1+y)^{N-1}+1}{N}\right)\right]\end{aligned}\right\}, & \text{(A13a)} \\ \dot{z}' = (z'-1)\left\{\begin{aligned}&y\left[rp(-y)\frac{(N-1)(z'+y)(1+y)^{N-2}}{N} - c_{GE}\right] \\ &-(z'+y)\left[1-\frac{r}{N} - rp(-y)\left(1-\frac{(N-1)(1+y)^{N-1}+1}{N}\right)\right]\end{aligned}\right\}. & \text{(A13b)} \end{cases}$$

Let $z' = z+1$, the above equation set becomes

$$\begin{cases} \dot{y} = y\left\{\begin{aligned}&(1+y)\left[rp(-y)\frac{(N-1)(z+1+y)(1+y)^{N-2}}{N} - c_{GE}\right] \\ &-(z+1+y)\left[1-\frac{r}{N} - rp(-y)\left(1-\frac{(N-1)(1+y)^{N-1}+1}{N}\right)\right]\end{aligned}\right\} = 0\times y + g_1(y,z), & \text{(A14a)} \\ \dot{z} = z\left\{\begin{aligned}&y\left[rp(-y)\frac{(N-1)(z+1+y)(1+y)^{N-2}}{N} - c_{GE}\right] \\ &-(z+1+y)\left[1-\frac{r}{N} - rp(-y)\left(1-\frac{(N-1)(1+y)^{N-1}+1}{N}\right)\right]\end{aligned}\right\} = \left(\frac{r}{N}-1\right)z + g_2(y,z), & \text{(A14b)} \end{cases}$$

where

$$\begin{cases} g_1(y,z) = y\left\{\begin{aligned}&(1+y)\left[rp(-y)\frac{(N-1)(z+1+y)(1+y)^{N-2}}{N} - c_{GE}\right] \\ &-(z+1+y)\left[1-\frac{r}{N} - rp(-y)\left(1-\frac{(N-1)(1+y)^{N-1}+1}{N}\right)\right]\end{aligned}\right\}, & \text{(A15a)} \\ g_2(y,z) = z\left\{\begin{aligned}&1-\frac{r}{N}+y\left[rp(-y)\frac{(N-1)(z+1+y)(1+y)^{N-2}}{N} - c_{GE}\right] \\ &-(z+1+y)\left[1-\frac{r}{N} - rp(-y)\left(1-\frac{(N-1)(1+y)^{N-1}+1}{N}\right)\right]\end{aligned}\right\}. & \text{(A15b)} \end{cases}$$

From Eq. (A15), on can find that

$$\begin{cases} g_1(0,0) = 0;\ \frac{\partial g_1}{\partial y}(0,0) = 0;\ \frac{\partial g_1}{\partial z}(0,0) = 0, & \text{(A16a)} \\ g_2(0,0) = 0;\ \frac{\partial g_2}{\partial y}(0,0) = 0;\ \frac{\partial g_2}{\partial z}(0,0) = 0. & \text{(A16b)} \end{cases}$$

According to above equation set, the center manifold theorem tells us that there exists a constant $\delta > 0$ and a continuously differentiable function $h(y)$, defined for all $\|y\| \leq \delta$, such that $z = h(y)$ is a center manifold for the evolutionary system described by Eq. (A14). In this case, the motion of the system in the center manifold is described by the following differential equation:

$$\dot{y} = y \left\{ \begin{array}{l} (1+y)\left[rp(-y)\,\frac{(N-1)(h(y)+1+y)(1+y)^{N-2}}{N} - c_{GE}\right] \\ -(h(y)+1+y)\left[1 - \frac{r}{N} - rp(-y)\left(1 - \frac{(N-1)(1+y)^{N-1}+1}{N}\right)\right] \end{array} \right\}. \tag{A17}$$

Then, we have the center manifold equation,

$$\begin{aligned} &\frac{\partial h(y)}{\partial y} y \left\{ \begin{array}{l} (1+y)\left[rp(-y)\,\frac{(N-1)(h(y)+1+y)(1+y)^{N-2}}{N} - c_{GE}\right] \\ -(h(y)+1+y)\left[1 - \frac{r}{N} - rp(-y)\left(1 - \frac{(N-1)(1+y)^{N-1}+1}{N}\right)\right] \end{array} \right\} \\ &- h(y) \left\{ \begin{array}{l} y\left[rp(-y)\,\frac{(N-1)(h(y)+1+y)(1+y)^{N-2}}{N} - c_{GE}\right] \\ -(h(y)+1+y)\left[1 - \frac{r}{N} - rp(-y)\left(1 - \frac{(N-1)(1+y)^{N-1}+1}{N}\right)\right] \end{array} \right\} = 0 \end{aligned} \tag{A18}$$

with boundary conditions

$$h(0) = 0;\ \frac{\partial h}{\partial y}(0) = 0. \tag{A19}$$

Let us start to try $h(y) = O(|y^2|) \approx 0$. Then, the reduced system is

$$\dot{y} = y(1+y)\left[-c_{GE} - 1 + \frac{r}{N} + rp(-y)\left(1 - \frac{1}{N}\right)\right] + O(|y^3|). \tag{A20}$$

Because $p(-y) \geq p(0) = \frac{N(c_{GE}+1)-r}{(N-1)r}$ holds, the origin of the reduced system is unstable. Consequently, the origin of the full system is unstable, which indicates that the homogeneous D state is unstable in this case. In a similar manner, one can also show that the pure D state at $c_{GE} = r - 1$ and $p(x_{GE} = 0) = 1$ is stable.

REFERENCES

[1] P. Taylor and L. Jonker, "Evolutionary stable strategies and game dynamics," Math. Biosci. **40**, 145 (1978).
[2] P. Schuster and K. Sigmund, "Replicator dynamics," J. Theor. Biol. **100**, 533 (1983).
[3] J. Hofbauer and K. Sigmund, *Evolutionary Games and Population Dynamics* (Cambridge University Press, Cambridge, UK, 1998).
[4] W. H. Sandholm, *Population Games and Evolutionary Dynamics* (MIT University Press, Cambridge, MA, 2010).
[5] K. Sigmund, *The Calculus of Selfishness* (Princeton University Press, Princeton, NJ, 2010).
[6] M. Perc, J. Gómez-Gardeñes, 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).
[7] C. Hauert, N. Haiden, and K. Sigmund, "The dynamics of public goods," Discrete Contin. Dyn. Syst. B **4**, 575 (2004).
[8] C. Hauert, S. De Monte, J. Hofbauer, and K. Sigmund, "Volunteering as red queen mechanism for cooperation in public goods game," Science **296**, 1129 (2002).
[9] F. C. Santos, M. D. Santos, and J. M. Pacheco, "Social diversity promotes the emergence of cooperation in public goods games," Nature **454**, 213 (2008).
[10] J. Wang, F. Fu, T. Wu, and L. Wang, "Emergence of social cooperation in threshold public good games with collective risk," Phys. Rev. E **80**, 016101 (2009).
[11] A. Arenas, J. Camacho, J. A. Cuesta, and R. Requejo, "The joker effect: Cooperation driven by destructive agents," J. Theor. Biol. **279**, 113 (2011).
[12] K. Sigmund, "Punish or perish? Retailation and collaboration among humans," Trends Ecol. Evol. **22**, 593 (2007).
[13] 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).
[14] L. Chen, J. Sun, K. Li, and Q. Li, "Research on the effectiveness of monitoring mechanism for "yield to pedestrian" based on system dynamics," Physica A **591**, 126804 (2022).
[15] T. Sasaki and S. Uchida, "The evolution of cooperation by social exclusion," Proc. R. Soc. B **280**, 20122498 (2013).
[16] K. Li, R. Cong, T. Wu, and L. Wang, "Social exclusion in finite populations," Phys. Rev. E **91**, 042810 (2015).
[17] K. Li, R. Cong, and L. Wang, "Cooperation induced by random sequential exclusion," EPL **114**, 58001 (2016).
[18] A. Szolnoki and X. Chen, "Alliance formation with exclusion in the spatial public goods game," Phys. Rev. E **95**, 052316 (2017).
[19] L. Liu, X. Chen, and M. Perc, "Evolutionary dynamics of cooperation in the public goods game with pool exclusion strategies," Nonlinear Dyn. **97**, 749 (2019).
[20] H. K. Khalil, *Nonlinear Systems* (Prentice Hall, Englewood Cliffs, NJ, 2002).
[21] T. Sasaki, A. Brännström, U. Dieckmann, and K. Sigmund, "The take-it-or-leave-it option allows small penalties to overcome social dilemmas," Proc. Natl. Acad. Sci. U.S.A. **109**, 1165 (2012).
[22] C. Hauert, F. Michor, M. A. Nowak, and M. Doebeli, "Synergy and discounting of cooperation in social dilemmas," J. Theor. Biol. **239**, 195 (2006).
[23] M. Archetti and I. Scheuring, "Coexistence of cooperation and defection in public goods games," Evolution **65**, 1140 (2011).
[24] 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).
[25] F. Franchetti and W. H. Sandholm, "An introduction to Dynamo: Diagrams for evolutionary game dynamics," Biol. Theory **8**, 167 (2013).
[26] T. A. Han, L. M. Pereira, and T. Lenaerts, "Avoiding or restricting defectors in public goods games?," J. R. Soc. Interface **12**, 20141203 (2015).
[27] T. A. Han, T. Lenaerts, F. C. Santos, and L. M. Pereira, "Voluntary safety commitments provide an escape from over-regulation in AI development," Technol. Soc. **68**, 101843 (2022).

[28]S. Wang, X. Chen, and A. Szolnoki, "Exploring optimal institutional incentives for public cooperation," Commun. Nonlinear Sci. Numer. Simulat. **79**, 104914 (2019).
[29]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).
[30]B. Herrmann, C. Thoni, and S. Gachter, "Antisocial punishment across societies," Science **319**, 1362 (2008).
[31]R.-W. Wang, B.-F. Sun, and Q. Zheng, "Diffusive coevolution and mutualism maintenance mechanisms in a fig-fig wasp system," Ecology **91**, 1308 (2010).
[32]D. G. Rand and M. A. Nowak, "The evolution of antisocial punishment in optional public goods games," Nat. Commun. **2**, 434 (2011).
[33]M. Salahshour, "Evolution of prosocial punishment in unstructured and structured populations and in the presence of antisocial punishment," PLoS One **16**, e0254860 (2021).
[34]M. dos Santos, "The evolution of anti-social rewarding and its countermeasures in public goods games," Proc. R. Soc. B **282**, 20141994 (2015).
[35]A. Szolnoki and M. Perc, "Second-order free-riding on antisocial punishment restores the effectiveness of prosocial punishment," Phys. Rev. X **7**, 041027 (2017).
[36]A. Szolnoki and M. Perc, "Antisocial pool rewarding does not deter public cooperation," Proc. R. Soc. B **282**, 20151975 (2015).
[37]S. Wiggins, *Introduction to Applied Nonlinear Dynamical Systems and Chaos* (Springer, New York, 2003).