



UvA-DARE (Digital Academic Repository)

Optimization of institutional incentives for cooperation in structured populations

Wang, S.; Chen, X.; Xiao, Z.; Szolnoki, A.; Vasconcelos, V.V.

DOI

[10.1098/rsif.2022.0653](https://doi.org/10.1098/rsif.2022.0653)

Publication date

2023

Document Version

Final published version

Published in

Journal of the Royal Society Interface

License

Article 25fa Dutch Copyright Act (<https://www.openaccess.nl/en/policies/open-access-in-dutch-copyright-law-taverne-amendment>)

[Link to publication](#)

Citation for published version (APA):

Wang, S., Chen, X., Xiao, Z., Szolnoki, A., & Vasconcelos, V. V. (2023). Optimization of institutional incentives for cooperation in structured populations. *Journal of the Royal Society Interface*, 20(199), Article 20220653. <https://doi.org/10.1098/rsif.2022.0653>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Research



Cite this article: Wang S, Chen X, Xiao Z, Szolnoki A, Vasconcelos VV. 2023 Optimization of institutional incentives for cooperation in structured populations. *J. R. Soc. Interface* **20**: 20220653.
<https://doi.org/10.1098/rsif.2022.0653>

Received: 5 September 2022
 Accepted: 3 January 2023

Subject Category:
 Life Sciences—Physics interface

Subject Areas:
 computational biology, evolution,
 biomathematics

Keywords:
 cooperation, Prisoner's Dilemma game,
 structured populations, institutional incentives,
 evolutionary game theory

Author for correspondence:
 Xiaojie Chen
 e-mail: xiaojiechen@uestc.edu.cn

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.6384898>.

Optimization of institutional incentives for cooperation in structured populations

Shengxian Wang^{1,2}, Xiaojie Chen¹, Zhilong Xiao^{1,3}, Attila Szolnoki⁴ and Vítor V. Vasconcelos^{5,6}

¹School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 611731, People's Republic of China

²Faculty of Science and Engineering, University of Groningen, Groningen 9747 AG, The Netherlands

³School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou 510006, People's Republic of China

⁴Institute of Technical Physics and Materials Science, Centre for Energy Research, P.O. Box 49, Budapest 1525, Hungary

⁵Computational Science Lab, Informatics Institute, University of Amsterdam, Amsterdam 1098XH, The Netherlands

⁶Institute for Advanced Study, University of Amsterdam, Amsterdam 1012 GC, The Netherlands

ORCID SW, 0000-0002-7094-511X; XC, 0000-0002-9129-2197; AS, 0000-0002-0907-0406; VVV, 0000-0002-4621-5272

The application of incentives, such as reward and punishment, is a frequently applied way for promoting cooperation among interacting individuals in structured populations. However, how to properly use the incentives is still a challenging problem for incentive-providing institutions. In particular, since the implementation of incentive is costly, to explore the optimal incentive protocol, which ensures the desired collective goal at a minimal cost, is worthy of study. In this work, we consider the positive and negative incentives for a structured population of individuals whose conflicting interactions are characterized by a Prisoner's Dilemma game. We establish an index function for quantifying the cumulative cost during the process of incentive implementation, and theoretically derive the optimal positive and negative incentive protocols for cooperation on regular networks. We find that both types of optimal incentive protocols are identical and time-invariant. Moreover, we compare the optimal rewarding and punishing schemes concerning implementation cost and provide a rigorous basis for the usage of incentives in the game-theoretical framework. We further perform computer simulations to support our theoretical results and explore their robustness for different types of population structures, including regular, random, small-world and scale-free networks.

1. Introduction

Cooperation is of vital importance in the contemporary era [1]. However, its evolution and emergence conflict with the immediate self-interest of interacting individuals [2]. Evolutionary game theory provides a common mathematical framework to depict how agents interact with each other in a dynamical process and to predict how cooperative action evolves from a population level [3]. As a representative paradigm, the Prisoner's Dilemma game has received considerable attention for studying the problem of cooperation in a population of interacting individuals [4].

When individuals play the evolutionary Prisoner's Dilemma game in a large well-mixed population, in which all are equally likely to interact, cooperation cannot emerge. However, the real-world population structures are not well-mixed but relatively complicated, and the interactions among individuals are limited to a set of neighbours in a structured population [5]. Inspired by the fast progress of network science, several interaction topologies have been tested, including small-world [6] and scale-free networks [7]. Features of these networks can influence the evolutionary dynamics of cooperation significantly [8–14]. Beside network topology, another crucial

determinant of the evolution of cooperation is the strategy update rule, which determines the microscopic update procedure [15–18]. For instance, cooperation can be favoured under the so-called death–birth (DB) strategy update rule if the benefit-to-cost ratio in the Prisoner’s Dilemma game exceeds the average degree of the specific interaction network [15–17]. In contrast, it cannot emerge under the alternative birth–death (BD) strategy update rule [15].

Overall, situations abound when cooperation cannot emerge in structured populations when no additional regulation mechanisms or moral nudges are incorporated [15,19,20]. In these unfavourable environments, prosocial incentives can be used to sustain cooperation among unrelated and competing agents [21–29]. Specifically, cooperators can be rewarded for their positive acts, or defectors are punished for their sweepingly negative impact. At the individual level, cooperation can be favoured whenever the amount of incentive exceeds the payoff difference between cooperating and defecting, no matter whether the incentive is positive or negative. However, in a given unfavourable environment for the emergence of cooperation, it is still unclear how intensive positive or negative incentive is needed to drive the population toward the desired direction. Furthermore, applying incentives is always costly [30–34]. Previous related works assume that, when used, the incentive amount is fixed at a certain value [30,32]. In those circumstances, the obtained incentive protocol is not necessarily the one with the minimal cumulative cost [33]. Therefore, finding the optimal time-varying incentive protocol is vital to ensure effective interventions that drive populations towards a productive and cooperative state at a minimal execution cost.

Our work addresses how much incentive is needed for cooperation to emerge and explores the optimal incentive protocols in a game-theoretical framework, where time-varying institutional positive or negative incentives are provided for a structured population of individuals playing the Prisoner’s Dilemma game. We establish an index function for quantifying the executing cumulative cost. We systematically survey all relevant strategy update rules, including DB, BD, imitation (IM) and pairwise-comparison (PC) updating [15–17,35]. Using optimal control theory, we obtain the dynamical incentive protocol under each strategy update rule leading to the minimal cumulative cost for the emergence of cooperation. Interestingly, we find that the optimal negative and positive incentive protocols are identical and time-invariant for each given strategy update rule. However, applying punishment can induce a lower cumulative cost than the usage of reward if the initial cooperation level is larger than the difference between the full cooperation state and the desired cooperation state. Otherwise, applying reward requires a lower cost. Beside analytical calculations, we perform computer simulations confirming that our results are valid in a broad class of population structures.

2. Results

We start with a structured population of N individuals who interact in a regular network of degree $k > 2$. In this graph, vertices represent interacting agents and the edges determine who interacts with whom. Each individual i plays the Prisoner’s Dilemma game with its neighbours. An agent can choose either to be a cooperator (C), which confers a benefit b to its opponent at a cost c to itself, or to be a defector (D), which

is costless and does not distribute any benefits. After playing the game with one neighbour, if institutional positive incentives are in place, the agent is rewarded an amount μ_R of incentive for choosing C. When negative institutional incentives are implemented, it is fined by an amount μ_P for choosing D. The agent collects an accumulated payoff π_i by interacting with all neighbours. We set the fitness, f_i , of individual i , chiefly the reproductive rate, as $1 - \omega + \omega\pi_i$, where ω ($0 \leq \omega \leq 1$) measures the strength of selection [36]. In this work, we concentrate on the effects of weak selection, meaning that $0 < \omega \ll 1$. Four different fitness-dependent strategy update rules are considered separately, and the details are given in the Model and methods section. In addition, in order to help readers intuitively understand the evolutionary process in networked Prisoner’s Dilemma game with institutional reward or punishment, we present an illustration figure as shown in figure 1.

2.1. Theoretical predictions for optimal incentive protocols

In this section, we, respectively, explore the incentive protocols under four alternative strategy update rules, including DB, BD, IM and PC updating. The details of these calculations are described in the electronic supplementary material, and here we only summarize the pair approximation approach for our dynamical system with positive or negative incentive in the weak selection limit.

Using the pair approximation approach (see electronic supplementary material, section S1), we get the dynamical equation of the fraction of cooperators under DB rule with positive or negative incentive as

$$\frac{dp_C}{dt} = F_{DB}(p_C, \mu_v, t) = \frac{\omega(k-2)}{k-1} [b + k(\mu_v - c)] p_C(1 - p_C) + o(\omega^2), \quad (2.1)$$

where p_C is the fraction of cooperators in the whole population and μ_v is the amount of positive incentive that one cooperator receives from an incentive-providing institution or the amount of negative incentive imposed on a defector by a central institution (note that such incentive is applied for every interaction as explained in the Model and methods section). Equation (2.1) has two equilibria, one at $p_C = 0$ and the other at $p_C = 1$. If $\mu_v > c - b/k$, the first is unstable and the second stable, indicating that cooperators prevail over defectors (further details are presented in electronic supplementary material, section S1). Notably, in the absence of incentives, i.e. $\mu_v = 0$, we get back the previously identified $b/c > k$ condition for the evolution of cooperation [15,16].

Using the Hamilton–Jacobi–Bellman (HJB) equation [37–39], in the condition of $\mu_v > c - b/k > 0$, we obtain analytically the optimal protocol $\mu_v^* = 2(ck - b)/k$ both for reward and punishment (see electronic supplementary material, section S1 for details). The optimal rewarding and optimal punishing protocols are both time-invariant and the optimal incentive levels for punishment and reward are identical, namely, $\mu_v^* = \mu_R^* = \mu_P^*$. Accordingly, with the optimal rewarding or punishing protocol, the dynamical system described by equation (2.1) can be solved and its solution is $p_C = 1/[1 + e^{-\beta_{DB}t}(1 - p_0)/p_0]$, where $\beta_{DB} = \omega(k-2)(ck - b)/(k-1)$ and $p_0 = p_C(0) > 0$ denotes the initial fraction of cooperators in the population.

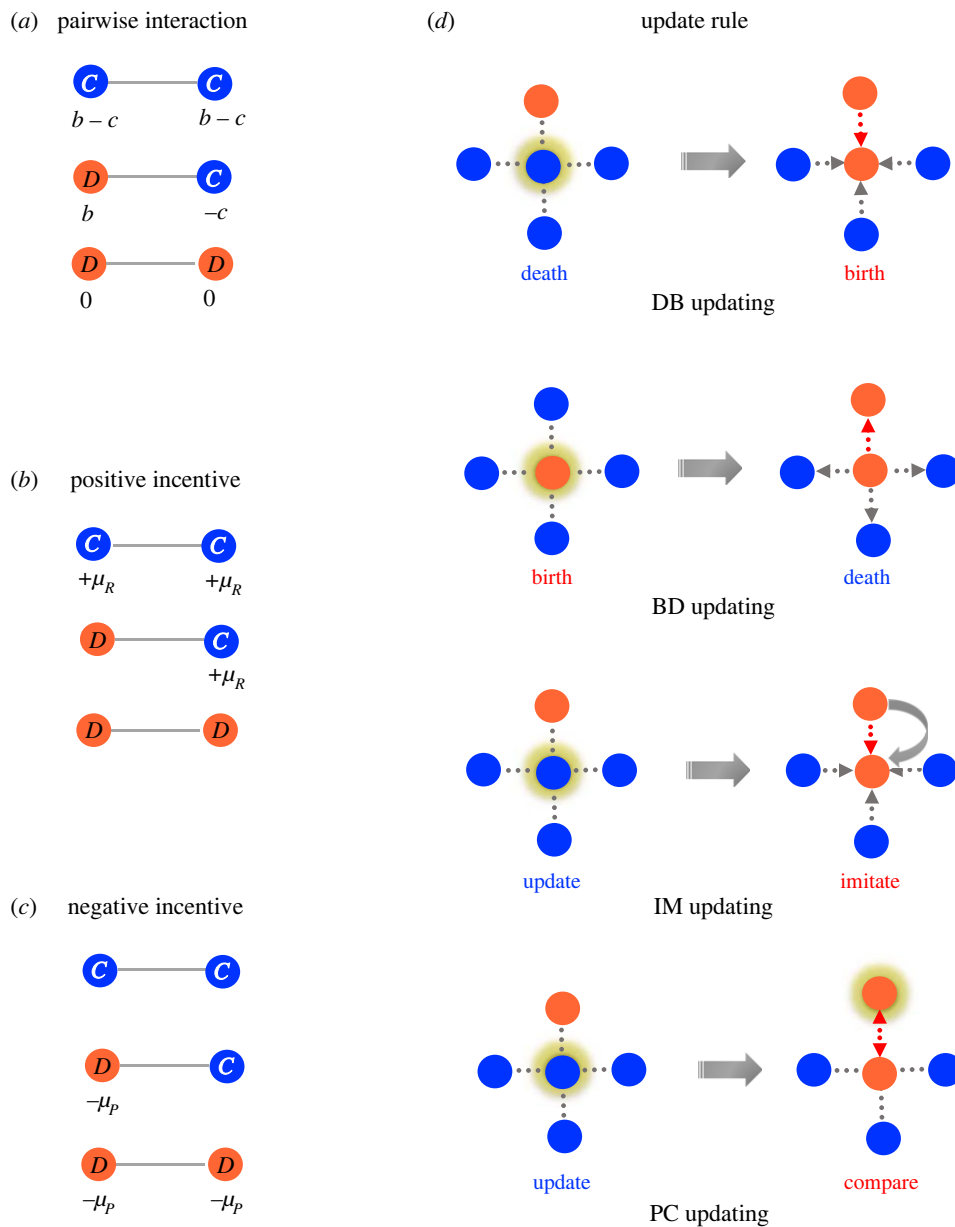


Figure 1. Evolutionary Prisoner's Dilemma game on a graph with institutional reward or punishment. (a) The pairwise interaction between two connected neighbours in a network. (b,c) How incentives are implemented for the two connected agents who played the game when positive (negative) incentives from the incentive-providing institution are applied. (d) The illustration of the four strategy update rules, depicting how agents update their strategies after obtaining payoffs from the pairwise interactions with neighbours and the incentive-providing institution.

The cumulative cost produced by the rewarding protocol μ^*_R for the dynamical system to reach the expected terminal state $p_C(t_f)$ from the initial state p_0 is

$$J^*_R = \int_0^{t_f} \frac{(kNp_C\mu^*_R)^2}{2} dt = \frac{(kN\mu^*_R)^2}{2\beta_{DB}} \left[p_0 + \delta - 1 + \ln\left(\frac{1-p_0}{\delta}\right) \right], \quad (2.2)$$

and, similarly, the cumulative cost produced by the punishing protocol μ^*_P becomes

$$J^*_P = \int_0^{t_f} \frac{(kNp_D\mu^*_P)^2}{2} dt = \frac{(kN\mu^*_P)^2}{2\beta_{DB}} \left[p_0 + \delta - 1 + \ln\left(\frac{1-\delta}{p_0}\right) \right]. \quad (2.3)$$

We further find that $J^*_R > J^*_P$ if $p_0 > \delta$ and $J^*_R < J^*_P$ if $p_0 < \delta$ (see electronic supplementary material, section S1 for details).

For BD updating, the dynamical equation is

$$\frac{dp_C}{dt} = F_{BD}(p_C, \mu_v, t) = \frac{\omega k(k-2)}{k-1} (\mu_v - c) p_C (1 - p_C) + o(\omega^2). \quad (2.4)$$

We prove that when $\mu_v > c$, the system can reach the stable full cooperation state. Naturally, in the absence of incentives we get back the results of Ohtsuki *et al.* [15,16]. By solving the HJB equation, the optimal incentive protocol is $\mu^*_v = 2c$ both for reward and punishment. The solution of equation (2.4) is $p_C = 1/[1 + e^{-\beta_{BD}t}(1-p_0)/p_0]$, where $\beta_{BD} = \omega k(k-2)c/(k-1)$ (see electronic supplementary material, section S2). Consequently, the cumulative cost to reach the expected terminal state in the case of the optimal rewarding protocol is

$$J^*_R = \frac{(kN\mu^*_R)^2}{2\beta_{BD}} \left[p_0 + \delta - 1 + \ln\left(\frac{1-p_0}{\delta}\right) \right], \quad (2.5)$$

while for the optimal punishing protocol it is

$$J_P^* = \frac{(kN\mu_P^*)^2}{2\beta_{BD}} \left[p_0 + \delta - 1 + \ln\left(\frac{1-\delta}{p_0}\right) \right]. \quad (2.6)$$

In the case of IM updating, the dynamical equation becomes

$$\begin{aligned} \frac{dp_C}{dt} &= F_{IM}(p_C, \mu_v, t) \\ &= \frac{\omega k^2(k-2)}{(k+1)^2(k-1)} [b + (\mu_v - c)(k+2)] p_C(1-p_C) \\ &\quad + o(\omega^2), \end{aligned} \quad (2.7)$$

which indicates that the system evolves to the full cooperation state if $\mu_v > c - b/(k+2)$. The solution of HJB for the optimal incentive protocol gives $\mu_v^* = 2[c(k+2) - b]/(k+2)$ both for reward and punishment (see electronic supplementary material, section S3). The solution of equation (2.7) is $p_C = 1/[1 + e^{-\beta_{IM}t}(1-p_0)/p_0]$, where $\beta_{IM} = \omega k^2(k-2)[c(k+2) - b]/[(k+1)^2(k-1)]$. The cumulative cost required for the optimal rewarding protocol is

$$J_R^* = \frac{(kN\mu_R^*)^2}{2\beta_{IM}} \left[p_0 + \delta - 1 + \ln\left(\frac{1-p_0}{\delta}\right) \right], \quad (2.8)$$

while for the optimal punishing protocol it becomes

$$J_P^* = \frac{(kN\mu_P^*)^2}{2\beta_{IM}} \left[p_0 + \delta - 1 + \ln\left(\frac{1-\delta}{p_0}\right) \right]. \quad (2.9)$$

When PC updating is applied, the dynamical equation is given by

$$\begin{aligned} \frac{dp_C}{dt} &= F_{PC}(p_C, \mu_v, t) \\ &= \frac{\omega k(k-2)}{2(k-1)} (\mu_v - c) p_C(1-p_C) + o(\omega^2), \end{aligned} \quad (2.10)$$

from which we find that when $\mu_v > c$ is satisfied, the stable full cooperation state can be reached (see electronic supplementary material, section S4). For the optimal incentive protocol, we get $\mu_v^* = \mu_R^* = \mu_P^* = 2c$. The solution of equation (2.10) is $p_C = 1/[1 + e^{-\beta_{PC}t}(1-p_0)/p_0]$, where $\beta_{PC} = \omega k(k-2)c/[2(k-1)]$. Consequently, the cumulative cost for the optimal rewarding protocol is

$$J_R^* = \frac{(kN\mu_R^*)^2}{2\beta_{PC}} \left[p_0 + \delta - 1 + \ln\left(\frac{1-p_0}{\delta}\right) \right], \quad (2.11)$$

while it is

$$J_P^* = \frac{(kN\mu_P^*)^2}{2\beta_{PC}} \left[p_0 + \delta - 1 + \ln\left(\frac{1-\delta}{p_0}\right) \right], \quad (2.12)$$

for the optimal punishing protocol.

Note that for each update rule, the governing equation in the weak selection limit always has two equilibrium points, which are the full defection and full cooperation states, respectively. Accordingly, we can obtain the minimal amounts of incentive needed for the evolution of cooperation, as summarized in figure 2. Independently of the applied update rule, the optimal incentive protocols are time-invariant and equal both for reward and punishment. Hence, $\mu_v^* = \mu_R^* = \mu_P^*$ for each update rule. The optimal protocols are summarized in figure 2. We further present the cumulative cost for the optimal reward and punishment protocols for each update rule in figure 2.

2.2. Numerical calculations and computer simulations for optimal incentive protocols

In the following, we validate our analytical results for μ_v^* ($v \in \{R, P\}$) by means of numerical calculations and computer simulations. Figure 3 illustrates the fraction of cooperators p_C as a function of time t for the optimal μ_v^* and two other incentive schemes under four strategy update rules we considered. First, we note that our system always reaches the expected terminal state $p_C(t_f) = 0.99$ when we launch the evolution from a random $p_0 = 0.5$ state for each update rule. However, reaching this state requires significantly different costs for the incentive-providing institution, and this value is the lowest for μ_v^* , no matter whether we apply reward or punishment. Notably, the optimal incentive protocol does not produce the fastest relaxation.

Moreover, we present the results of Monte Carlo simulations when positive or negative incentives are applied for each update rule (electronic supplementary material, figures S1–S4). We can find that for different update rules the usage of optimal incentive protocol does not result in the fastest relaxation to the desired cooperation state, but it always leads to the smallest cumulative cost for the institution. We must stress that our findings are not limited to regular networks, but they remain valid for a broad range of interaction graphs from irregular random networks to small-world networks and scale-free networks. Our numerical calculations and simulation results coincide with the analytical predictions, and they illustrate that the cumulative cost is the lowest under the optimal incentive protocol.

2.3. Comparison between optimal reward and punishment protocols

Through theoretical analysis presented in the electronic supplementary material, we can conclude that the execution of optimal punishing scheme requires a lower cumulative cost, compared with the optimal rewarding scheme for each update rule, if the initial cooperation level is larger than the difference between the full and desired cooperation states. Otherwise, the usage of optimal rewarding scheme requires a lower cumulative cost. In order to verify such theoretical prediction and have an intuitive comparison, we show the cumulative cost values induced by the optimal punishing and rewarding protocols when the initial fraction p_0 of cooperators is adjustable. By assuming the optimal protocols of incentives, the requested cumulative cost can be determined by numerically integrating or by using Monte Carlo simulations.

Our numerical results are summarized in figure 4 for each update rule. From the top row of figure 4, we observe that when the initial cooperation level is larger than the difference between the full cooperation and desired terminal states (with a set tolerance for defection, δ), the institution needs to spend less cumulative cost to reach the expected cooperation state by means of punishing. On the contrary, from the bottom of figure 4, we see that when the initial cooperation level is less than the difference between the full cooperation and desired terminal states, the institution needs to spend less cumulative cost to reach the expected cooperation state by means of rewarding. However, in these different cases, no matter whether the optimal rewarding or punishing protocol is applied, the corresponding cumulative cost value decreases as the initial cooperation level p_0

update rule	minimal μ_R/μ_P	μ_R^*/μ_P^*	J_R^*	J_P^*
DB updating	$c - \frac{b}{k}$	$\frac{2(kc-b)}{k}$	$\frac{(kN\mu_R^*)^2}{2\beta_{DB}} [p_0 + \delta - 1 + \ln(\frac{1-p_0}{\delta})]$	$\frac{(kN\mu_P^*)^2}{2\beta_{DB}} [p_0 + \delta - 1 + \ln(\frac{1-\delta}{p_0})]$
BD updating	c	$2c$	$\frac{(kN\mu_R^*)^2}{2\beta_{BD}} [p_0 + \delta - 1 + \ln(\frac{1-p_0}{\delta})]$	$\frac{(kN\mu_P^*)^2}{2\beta_{BD}} [p_0 + \delta - 1 + \ln(\frac{1-\delta}{p_0})]$
IM updating	$c - \frac{b}{k+2}$	$\frac{2(k+2)c-2b}{k+2}$	$\frac{(kN\mu_R^*)^2}{2\beta_{IM}} [p_0 + \delta - 1 + \ln(\frac{1-p_0}{\delta})]$	$\frac{(kN\mu_P^*)^2}{2\beta_{IM}} [p_0 + \delta - 1 + \ln(\frac{1-\delta}{p_0})]$
PC updating	c	$2c$	$\frac{(kN\mu_R^*)^2}{2\beta_{PC}} [p_0 + \delta - 1 + \ln(\frac{1-p_0}{\delta})]$	$\frac{(kN\mu_P^*)^2}{2\beta_{PC}} [p_0 + \delta - 1 + \ln(\frac{1-\delta}{p_0})]$

Figure 2. Optimization of institutional incentives for cooperation for different strategy update rules. Here, the minimal μ_R (μ_P) means the minimal amount of positive (negative) incentive needed for the evolution of cooperation for different strategy update rules. μ_R^* (μ_P^*) represents the optimal positive (negative) incentive protocol for different strategy update rules. J_R^* (J_P^*) means the cumulative cost produced by the optimal rewarding (punishing) protocol μ_R^* (μ_P^*) for the dynamical system to reach the expected terminal state $1 - \delta$ from the initial state p_0 . In addition, N denotes the population size and k the degree of the regular network. b represents the benefit of cooperation and c the cost of cooperation. The parameters $\beta_{DB} = \omega(k-2)(ck-b)/(k-1)$ under DB updating, $\beta_{BD} = \omega k(k-2)c/(k-1)$ under BD updating, $\beta_{IM} = \omega k^2(k-2)[c(k+2)-b]/[(k+1)^2(k-1)]$ under IM updating, and $\beta_{PC} = \omega k(k-2)c/[2(k-1)]$ under PC updating.

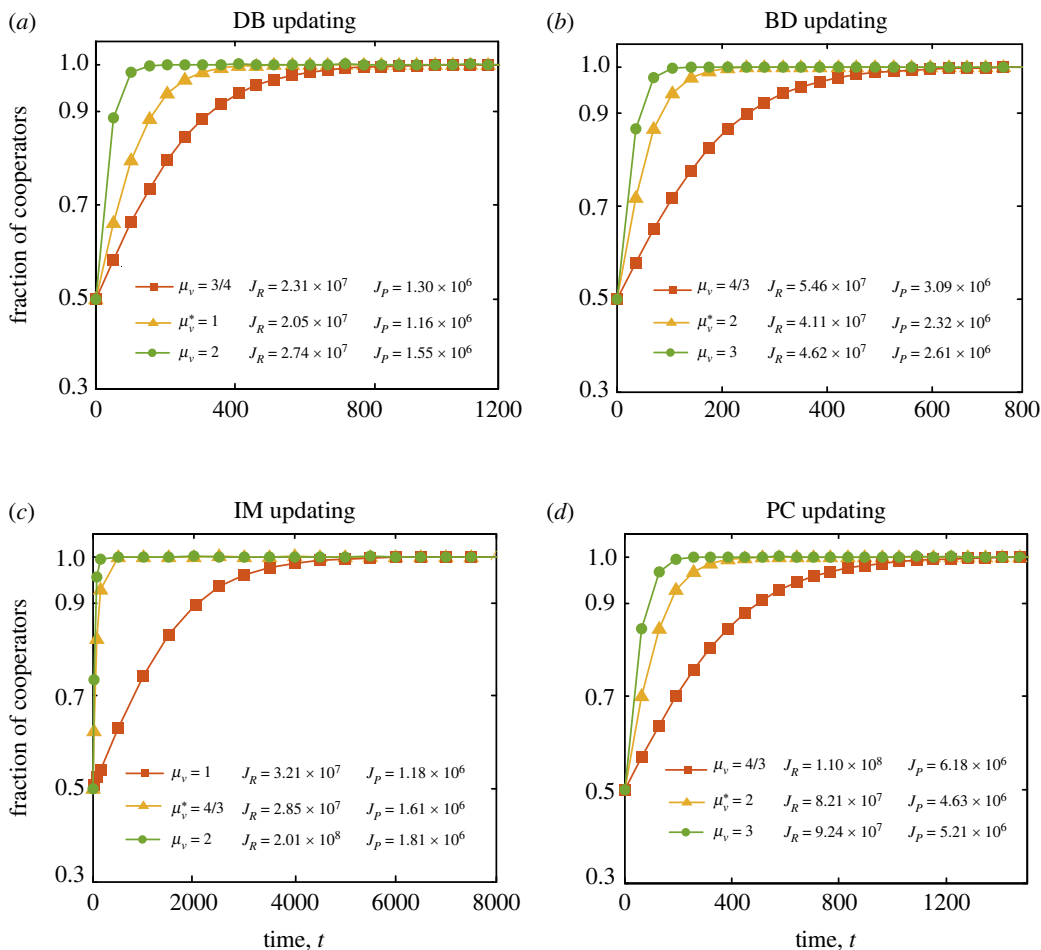


Figure 3. Time evolution of the fraction of cooperators for positive and negative incentives under different strategy update rules. Each panel shows the results derived from numerical calculations based on the obtained dynamical equation at different levels of incentives for reward (R) or punishment (P). The optimal incentive level is marked by asterisk (*). For comparison we have also marked the J_R and J_P amounts of cumulative cost to reach the desired terminal state for each incentive protocol. Parameters: $N = 100$, $b = 2$, $c = 1$, $\delta = 0.01$, $\omega = 0.01$, $p_0 = 0.5$ and $k = 4$.

increases. In addition, our simulation results presented in figure S5 of the electronic supplementary material also support our theoretical analysis, which provides a rigorous

basis for the usage of incentives under different initial conditions in the context of evolutionary Prisoner's Dilemmas game in structured populations.

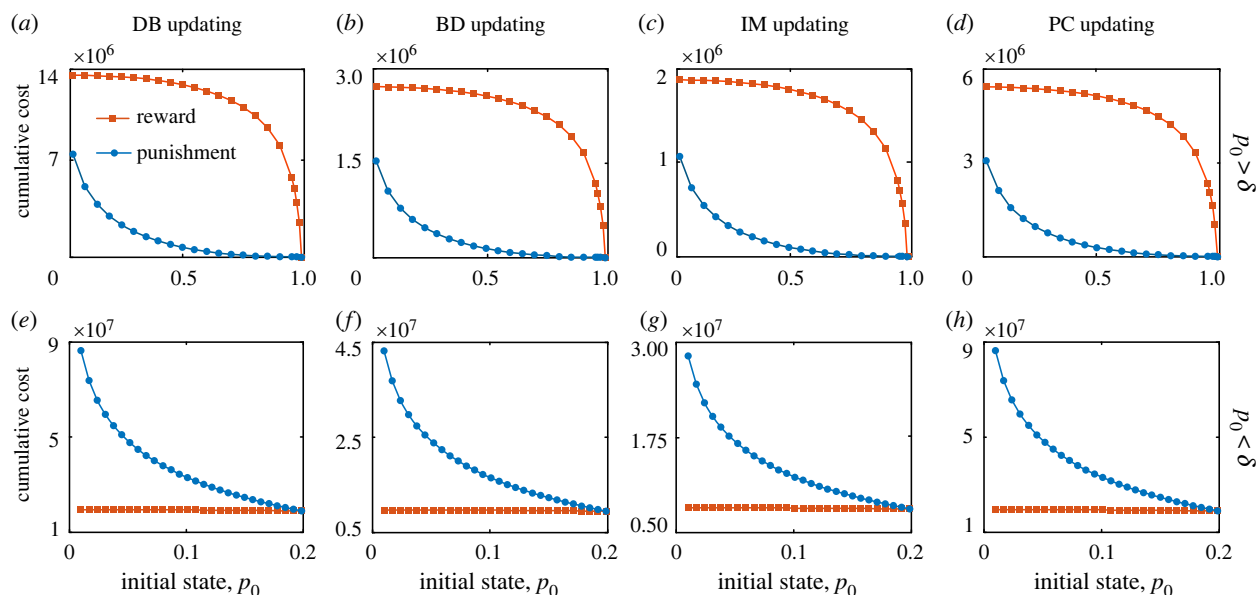


Figure 4. Cumulative cost needed for reaching the expected terminal state in dependence of the p_0 initial portion of cooperators for the optimal rewarding and punishing protocols for different strategy update rules. Each column of panels represents a strategy update rule as indicated. Top row represents the results of numerical calculations based on the obtained dynamical equation in the condition of $p_0 > \delta = 0.01$, while bottom row represents the results of numerical calculations in the condition of $p_0 < \delta = 0.2$. Other parameters: $N = 100$, $b = 2$, $c = 1$, $\omega = 0.01$ and $k = 4$.

3. Discussion

In human society, prosocial incentives are an essential means of avoiding the ‘tragedy of the commons’ [21,22,40]. For incentive-providing institutions, however, the choice of incentives is mainly based on two aspects. One of them is knowing how much incentive is needed to promote the evolution of cooperation (potentially under different update rules) in structured populations. The other is whether the applied incentive scheme is an optimal time-dependent protocol requiring the minimal cost for the institution.

In order to investigate the above-mentioned tasks, we establish a game-theoretical framework. Namely, we consider the positive or negative incentive into the networked Prisoner’s Dilemma game with four different strategy update rules, respectively. For a given update rule, we obtain the theoretical conditions of the minimal amounts of incentives needed for the evolution of cooperation. By establishing an index function for quantifying the executing cost, we derive the optimal positive and negative incentive protocols for each strategy update rule, respectively, by means of the approach of HJB equation. We find that these optimal incentive protocols are time-invariant for all the considered update rules. In addition, the optimal incentive protocols are identical both for negative and positive incentives. However, applying the punishing scheme requires a lower cumulative cost for the incentive-providing institution when the initial cooperation level is relatively high; otherwise, applying the rewarding scheme is cheaper. We further perform computer simulations, which confirm that our results are valid in different types of interaction topologies described by regular, random, small-world and scale-free networks and thus demonstrate the general robustness of the findings.

In this work, we have quantified how much incentive is needed for the evolution of cooperation under each update rule we considered. However, when prosocial incentives are provided, the game structure may be changed. In particular, when the incentive amount $\mu_v > c$, the Prisoner’s Dilemma

game will be transformed into the harmony game, where cooperators dominate defectors naturally [9]. Interestingly, we note that if $\mu_v > c - b/k$ under DB updating and if $\mu_v > c - b/(k+2)$ under IM updating, cooperation is favoured. This implies that under DB and IM rule, the μ_v value can be smaller than the cost c , which guarantees that the game structure is not changed. For BD and PC updating, however, the incentive amount needed for the evolution of cooperation must completely outweigh the cost of cooperation. Hence the effectiveness of interventions might shed some light on which update type can best capture population behaviours.

Here, we stress that the obtained optimal incentive protocols are time-invariant by solving the optimal control problems we formulated. More strikingly, the obtained optimal negative incentive protocol for each update rule is the same as the optimal positive incentive protocol. Thus, these optimal incentive protocols are state-independent. Then the institution does not need to monitor the population state from time to time for optimal incentive implementation and can save monitoring costs since monitoring is generally costly. In addition, the optimal incentive level for DB and IM updating can guarantee that the dilemma faced by individuals is still a Prisoner’s Dilemma. Accordingly, our work reflects that the incentive-based control protocols we obtained under DB and IM updating are simple and effective for promoting the evolution of cooperation in structured populations.

Incentives can be used as controlling tools to regulate the decision-making behaviours of individuals [28]. However, there are significant preference differences in the usage of punishment and reward for the evolution of cooperation from the perspectives of individuals and incentive-providing institutions [41]. It has been suggested that punishment is often not preferred since the usage of punishment leads to a low total income or a low average payoff in repeated games [23,24]. By contrast, incentive-providing institutions prefer to use punishments more frequently. This is because punishment incurs a lower cost of implementing incentives

for the promotion of cooperation [30,41]. Here, we consider the top-down-like incentive mechanism under which cooperators can be rewarded or defectors can be punished directly by the external centralized institution, which has existed and works stably. In this framework, we strictly compare the rewarding and punishing schemes concerning implementation cost and show how the best choice depends on the initial cooperation level, providing a rigorous basis for the usage of incentives in the context of the Prisoner's Dilemma game and where it might apply in human society.

Extensions of our work are plentiful, especially in analytical terms. We would like to point out that we use the pair approximation approach to obtain our analytical results and this approach is used mainly for regular networks [15]. Indeed, this approach can be extended for other types of complex networks when the properties of these networks are additionally considered [42,43]. Along this line, hence it is worth analytically investigating the low-cost incentive policy for the evolution of cooperation by fully considering spatial properties of the underlying population structures. Furthermore, our theoretical analysis could be extended by resorting to methods relying on calculating the coalescence times of random walks [11]. These methods might be particularly constructive when interventions at the topology level are at stake since they create a tighter link between the outcome and the network. In addition, we obtain our theoretical results in the limit of weak selection. Numerical analysis in other contexts has shown that due to relevant factors from the environment, the intensity of selection can change the game dynamics no matter the population is well-mixed or structured, and plays a crucial role in the determination of cost-efficient institutional incentive [44–47]. Hence, a natural question arising here is whether our theoretical results are still valid when selection is not weak [48]. Indeed, analytical calculations for strong selection remain tractable for some structures with high symmetry. Thus, there is potential to identify the theoretical conditions of how much incentive is needed to promote cooperation and to explore the optimal incentive protocols for these population structures.

In addition, our work focuses on minimizing the incentive costs up to a set level of cooperation in a population, irrespectively of how long that takes. Further research is required for situations in which the rate at which the transition happens can be important and a time-varying protocol might be the solution. We consider our research in the framework of the Prisoner's Dilemma game, which is a paradigm for studying the evolution of cooperation. There are other prototypical two-person dilemmas, e.g. snowdrift game [49] and stag-hunt game [50]. A promising extension of this work is to consider these mentioned games for future study as well as group interactions or higher-order interactions [51–55]. Furthermore, we design the cost-efficient incentive protocols by considering the global information (i.e. the fraction of cooperators in the whole population), but the local neighbourhood properties on a structured network, such as how many cooperators are there in a neighbourhood, affect the final evolutionary outcomes [56–58]. Hence, it would be important to take into account this local information for optimal incentive protocols with minimal cost in the future work. Finally, other prosocial behaviours, such as honesty [59] or trust and trustworthiness [60], are also fundamental for cooperation, and hence it is a meaningful extension to study the optimization problems of incentives for promoting the evolution of these behaviours.

4. Model and methods

4.1. Prisoner's Dilemma game

We consider that a population of individuals are distributed on the nodes of an interaction graph. At each round, each individual plays the evolutionary Prisoner's Dilemma game with its neighbours and can choose to cooperate (C) or defect (D). We consider the payoff matrix for the game as

$$\begin{array}{cc} & C & D \\ C & (b-c, -c) & \\ D & (b, 0) & \end{array}, \quad (4.1)$$

where b represents the benefit of cooperation and c ($0 < c < b$) represents the cost of cooperation. After engaging in the pairwise interactions with all the adjacent neighbours, each individual collects its payoff based on the payoff matrix.

4.2. Institutional incentives

Furthermore, prosocial incentives provided by incentive-providing institutions can be used to reward cooperators or punish defectors after they play the game with their neighbours. Here, we consider both types of incentives, i.e. positive and negative incentives, respectively. If positive incentives are used, then a cooperator in the game is rewarded with a μ_R amount received from a central institution when interacting with a neighbour [24,28]. If negative incentives are used, then a defector is fined by a μ_P amount for each interaction [23]. Consequently, when positive incentives are used, the modified payoff matrix becomes

$$\begin{array}{cc} & C & D \\ C & (b-c+\mu_R, -c+\mu_R) & \\ D & (b, 0) & \end{array}, \quad (4.2)$$

and when negative incentives are used, the modified payoff matrix becomes

$$\begin{array}{cc} & C & D \\ C & (b-c, -c) & \\ D & (b-\mu_P, -\mu_P) & \end{array}. \quad (4.3)$$

As a result, based on the above payoff matrices each individual collects its total payoff, which is derived from the pairwise interactions with neighbours and the incentive-providing institution.

4.3. Strategy update rules

According to the evolutionary selection principle, players update their strategies from time to time, but the way how to do it may influence the evolutionary outcome significantly [9,15]. In agreement with previous works [15,16], we here consider four major strategy update rules, describing DB, BD, IM and PC updating. Specifically, for DB updating, at each time step a random individual from the entire population is chosen to die; subsequently the neighbours compete for the empty site with probability proportional to their fitness. For BD updating, at each time step an individual is chosen for reproduction from the entire population with probability proportional to fitness; the offspring of this individual replaces a randomly selected neighbour. For IM updating, at each time step a random individual from the entire population is chosen to update its strategy; it will

either stay with its own strategy or imitate one of the neighbours' strategies with probability proportional to their fitness. For PC updating, at each time step a random individual is chosen to update its strategy, and it compares its own fitness with a randomly chosen neighbour. The focal individual either keeps its current strategy or adopts the neighbour's strategy with a probability that depends on the fitness difference.

4.4. Optimizing incentives

Since providing incentives is costly for institutions, it has a paramount importance to find the optimal μ_v^* , which requires the minimal effort but is still capable of supporting cooperation effectively. To reach this goal, we first establish an index function for quantifying the executing cumulative cost, which is expressed as

$$J_v = \int_{t_0}^{t_f} \frac{(kNp_i\mu_v)^2}{2} dt, \quad (4.4)$$

where $p_i = p_C$ if $v = R$, which means that positive incentives are applied, otherwise $p_i = p_D$, which means that negative incentives are applied. Here t_0 is the initial time and is set to 0 in this work, and t_f is the terminal time for the system. Based on the above description, we then explore the optimal incentive protocol during the evolutionary period between 0 and t_f by using optimal control theory [37–39]. It is a crucial assumption that t_f is not fixed, but we monitor the evolution until the fraction of cooperators reaches the target level $p_C(t_f)$ at t_f . Here, we suppose that $p_C(t_f) = 1 - \delta > p_0$, where p_0 ($p_0 > 0$) is the initial cooperation level and δ is the parameter determining the expected cooperation level at t_f , satisfying $0 \leq \delta < 1 - p_0$.

Accordingly, we formulate the optimal control problem for reward or punishment given as

$$\begin{aligned} \min \quad & J_v = \int_0^{t_f} \frac{(kNp_i\mu_v)^2}{2} dt, \\ \text{s.t.} \quad & \begin{cases} \frac{dp_C}{dt} = F_j(p_C, \mu_v, t), j = \text{DB, BD, IM, or PC,} \\ p_C(0) = p_0, \\ p_C(t_f) = 1 - \delta. \end{cases} \end{aligned} \quad (4.5)$$

References

- Hauert S, Mitri S, Keller L, Floreano D. 2010 Evolving cooperation: from biology to engineering. In *The horizons of evolutionary robotics* (eds PA Vargas, EA Di Paolo, I Harvey, P Husbands), pp. 203–218. Boston, MA: MIT Press.
- Perc M, Jordan JJ, Rand DG, Wang Z, Boccaletti S, Szolnoki A. 2017 Statistical physics of human cooperation. *Phys. Rep.* **687**, 1–51. (doi:10.1016/j.physrep.2017.05.004)
- Hofbauer J, Sigmund K. 1998 *Evolutionary games and population dynamics*. Cambridge, UK: Cambridge University Press.
- Nowak MA, May RM. 1992 Evolutionary games and spatial chaos. *Nature* **359**, 826–829. (doi:10.1038/359826a0)
- Erdős P, Rényi A. 1959 On random graphs I. *Publ. Math. Debr.* **6**, 290–297. (doi:10.1109/ICSMC.2006.384625)
- Watts DJ, Strogatz SH. 1998 Collective dynamics of 'small-world' networks. *Nature* **393**, 440–442. (doi:10.1038/30918)
- Barabási AL, Albert R. 1999 Emergence of scaling in random networks. *Science* **286**, 509–512. (doi:10.1126/science.286.5439.509)
- Santos FC, Pacheco JM. 2005 Scale-free networks provide a unifying framework for the emergence of cooperation. *Phys. Rev. Lett.* **95**, 098104. (doi:10.1103/PhysRevLett.95.098104)
- Szabó G, Fáth G. 2007 Evolutionary games on graphs. *Phys. Rep.* **446**, 97–216. (doi:10.1016/j.physrep.2007.04.004)
- Tarnita CE, Ohtsuki H, Antal T, Fu F, Nowak MA. 2009 Strategy selection in structured populations. *J. Theor. Biol.* **259**, 570–581. (doi:10.1016/j.jtbi.2009.03.035)
- Allen B, Lippner G, Chen YT, Fotouhi B, Momeni N, Yau ST, Nowak MA. 2017 Evolutionary dynamics on any population structure. *Nature* **544**, 227–230. (doi:10.1038/nature21723)
- Li A, Zhou L, Su Q, Cornelius SP, Liu YY, Wang L, Levin SA. 2020 Evolution of cooperation on temporal networks. *Nat. Commun.* **11**, 2259. (doi:10.1038/s41467-020-16088-w)
- Su Q, McAvoy A, Mori Y, Plotkin JB. 2022 Evolution of prosocial behaviours in multilayer populations. *Nat. Hum. Behav.* **6**, 338–348. (doi:10.1038/s41562-021-01241-2)

Here the cost function J_v characterizes the cumulative cost on average during the period $[0, t_f]$ for the dynamical system to reach the terminal state $1 - \delta$ from the initial state p_0 . Thus the quantity $\min J_v$ can work as the objective of calculating the optimal incentive protocol μ_v^* with the minimal executing cost. The details of solving the optimal control problems can be found in sections S1–S4 of the electronic supplementary material.

4.5. Monte Carlo simulations

During a full Monte Carlo step, on average each player has a chance to update its strategy. The applied four different update rules are specified above. Besides, we have tested alternative interaction topologies, including regular networks generated by using a two-dimensional square lattice of size $N = L \times L$ [4] and scale-free networks obtained by using preferential-attachment model [7] starting from $m_0 = 6$, where at every time step each new node is connected to $m = 2$ existing nodes fulfilling the standard power-law distribution. Alternatively, Erdős–Rényi random graph model [5] and small-world networks of Watts–Strogatz model with $p = 0.1$ rewiring parameter [6] are considered. These simulation results are summarized in figures S1–S5 of the electronic supplementary material.

Data accessibility. The data are provided in the electronic supplementary material [61].

Authors' contributions. S.W.: conceptualization, formal analysis, investigation, writing—original draft, writing—review and editing; X.C.: conceptualization, formal analysis, funding acquisition, supervision, writing—original draft, writing—review and editing; Z.X.: data curation, formal analysis, investigation; A.S.: conceptualization, formal analysis, writing—original draft, writing—review and editing; V.V.V.: conceptualization, formal analysis, writing—original draft, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. This research was supported by the National Natural Science Foundation of China (grant nos. 61976048 and 62036002) and the Fundamental Research Funds of the Central Universities of China. S.W. acknowledges the support from China Scholarship Council (grant no. 202006070122). A.S. was supported by the National Research, Development and Innovation Office (NKFIH) under grant no. K142948. V.V.V. acknowledges funding from the Computational Science Lab—Informatics Institute of the University of Amsterdam.

14. Su Q, Allen B, Plotkin JB. 2022 Evolution of cooperation with asymmetric social interactions. *Proc. Natl Acad. Sci. USA* **119**, e2113468118. (doi:10.1073/pnas.2113468118)
15. Ohtsuki H, Hauert C, Lieberman E, Nowak MA. 2006 A simple rule for the evolution of cooperation on graphs and social networks. *Nature* **441**, 502–505. (doi:10.1038/nature04605)
16. Ohtsuki H, Nowak MA. 2006 The replicator equation on graphs. *J. Theor. Biol.* **243**, 86–97. (doi:10.1016/j.jtbi.2006.06.004)
17. Nowak MA, Tarnita CE, Antal T. 2010 Evolutionary dynamics in structured populations. *Phil. Trans. R. Soc. B* **365**, 19–30. (doi:10.1098/rstb.2009.0215)
18. Zhou L, Wu B, Du J, Wang L. 2021 Aspiration dynamics generate robust predictions in heterogeneous populations. *Nat. Commun.* **12**, 3250. (doi:10.1038/s41467-021-23548-4)
19. Capraro V, Perc M. 2021 Mathematical foundations of moral preferences. *J. R. Soc. Interface* **18**, 20200880. (doi:10.1098/rsif.2020.0880)
20. Capraro V, Halpern JY, Perc M. In press. From outcome-based to language-based preferences. *J. Econ. Lit.*
21. Henrich J. 2006 Cooperation, punishment, and the evolution of human institutions. *Science* **312**, 60–61. (doi:10.1126/science.1126398)
22. Güreker Ö, Irlenbusch B, Rockenbach B. 2006 The competitive advantage of sanctioning institutions. *Science* **312**, 108–111. (doi:10.1126/science.1123633)
23. Dreber A, Rand DG, Fudenberg D, Nowak MA. 2008 Winners don't punish. *Nature* **452**, 348–351. (doi:10.1038/nature06723)
24. Rand DG, Dreber A, Ellingsen T, Fudenberg D, Nowak MA. 2009 Positive interactions promote public cooperation. *Science* **325**, 1272–1275. (doi:10.1126/science.1177418)
25. Sigmund K, De Silva H, Traulsen A, Hauert C. 2010 Social learning promotes institutions for governing the commons. *Nature* **466**, 861–863. (doi:10.1038/nature09203)
26. Han TA, Pereira LM, Lenaerts T. 2015 Avoiding or restricting defectors in public goods games? *J. R. Soc. Interface* **12**, 20141203. (doi:10.1098/rsif.2014.1203)
27. Mann RP, Helbing D. 2017 Optimal incentives for collective intelligence. *Proc. Natl Acad. Sci. USA* **114**, 5077–5082. (doi:10.1073/pnas.1618722114)
28. Riehl J, Ramazi P, Cao M. 2018 Incentive-based control of asynchronous best-response dynamics on binary decision networks. *IEEE Trans. Control Netw. Syst.* **6**, 727–736. (doi:10.1109/TCNS.2018.2873166)
29. Vasconcelos VV, Dannenberg A, Levin SA. 2022 Punishment institutions selected and sustained through voting and learning. *Nat. Sustain.* **5**, 578–585. (doi:10.1038/s41893-022-00877-w)
30. Sasaki T, Brännström Å, Dieckmann U, Sigmund K. 2012 The take-it-or-leave-it option allows small penalties to overcome social dilemmas. *Proc. Natl Acad. Sci. USA* **109**, 1165–1169. (doi:10.1073/pnas.1115219109)
31. Vasconcelos VV, Santos FC, Pacheco JM. 2013 A bottom-up institutional approach to cooperative governance of risky commons. *Nat. Clim. Change* **3**, 797–801. (doi:10.1038/NCLIMATE1927)
32. Chen X, Sasaki T, Brännström Å, Dieckmann U. 2015 First carrot, then stick: how the adaptive hybridization of incentives promotes cooperation. *J. R. Soc. Interface* **12**, 20140935. (doi:10.1098/rsif.2014.0935)
33. Wang S, Chen X, Szolnoki A. 2019 Exploring optimal institutional incentives for public cooperation. *Commun. Nonlinear Sci. Numer. Simul.* **79**, 104914. (doi:10.1016/j.cnsns.2019.104914)
34. Duong MH, Han TA. 2021 Cost efficiency of institutional incentives for promoting cooperation in finite populations. *Proc. R. Soc. A* **477**, 20210568. (doi:10.1098/rspa.2021.0568)
35. Szabó G, Töke C. 1998 Evolutionary prisoner's dilemma game on a square lattice. *Phys. Rev. E* **58**, 69–73. (doi:10.1103/PhysRevE.58.69)
36. Nowak MA, Sasaki A, Taylor C, Fudenberg D. 2004 Emergence of cooperation and evolutionary stability in finite populations. *Nature* **428**, 646–650. (doi:10.1038/nature02414)
37. Evans LC. 2005 *An introduction to mathematical optimal control theory*. Berkeley, USA: University of California Press.
38. Geering HP. 2007 *Optimal control with engineering applications*. Berlin, Germany: Springer.
39. Lenhart S, Workman JT. 2007 *Optimal control applied to biological models*. Boca Raton, USA: Chapman and Hall/CRC.
40. Ostrom E. 1990 *Governing the commons: the evolution of institutions for collective action*. New York, NY: Cambridge University Press.
41. Gächter S. 2012 Carrot or stick? *Nature* **483**, 39–40. (doi:10.1038/483039a)
42. Morita S. 2008 Extended pair approximation of evolutionary game on complex networks. *Prog. Theor. Phys.* **19**, 29–38. (doi:10.1143/PTP.119.29)
43. Overton CE, Broom M, Hadjichrysanthou C, Sharkey KJ. 2019 Methods for approximating stochastic evolutionary dynamics on graphs. *J. Theor. Biol.* **468**, 45–59. (doi:10.1016/j.jtbi.2019.02.009)
44. Pinheiro FL, Santos FC, Pacheco JM. 2012 How selection pressure changes the nature of social dilemmas in structured populations. *New J. Phys.* **14**, 073035. (doi:10.1088/1367-2630/14/7/073035)
45. Zisis I, Di Guida S, Han TA, Kirchsteiger G, Lenaerts T. 2015 Generosity motivated by acceptance-evolutionary analysis of an anticipation game. *Sci. Rep.* **5**, 18076. (doi:10.1038/srep18076)
46. McAvoy A, Rao A, Hauert C. 2021 Intriguing effects of selection intensity on the evolution of prosocial behaviors. *PLoS Comput. Biol.* **17**, e1009611. (doi:10.1371/journal.pcbi.1009611)
47. Han TA, Tran-Thanh L. 2018 Cost-effective external interference for promoting the evolution of cooperation. *Sci. Rep.* **8**, 15997. (doi:10.1038/s41598-018-34435-2)
48. Ibsen-Jensen R, Chatterjee K, Nowak MA. 2015 Computational complexity of ecological and evolutionary spatial dynamics. *Proc. Natl Acad. Sci. USA* **112**, 15 636–15 641. (doi:10.1073/pnas.1511366112)
49. Doebeli M, Hauert C. 2005 Models of cooperation based on the Prisoner's Dilemma and the Snowdrift game. *Ecol. Lett.* **8**, 748–766. (doi:10.1111/j.1461-0248.2005.00773.x)
50. Skyrms B. 2004 *The stag hunt and the evolution of social structure*. Cambridge, UK: Cambridge University Press.
51. Perc M, Gómez-Gardeñes J, Szolnoki A, Floría M, Moreno Y. 2013 Evolutionary dynamics of group interactions on structured populations: a review. *J. R. Soc. Interface* **10**, 20120997. (doi:10.1098/rsif.2012.0997)
52. Li A, Wu B, Wang L. 2014 Cooperation with both synergistic and local interactions can be worse than each alone. *Sci. Rep.* **4**, 5536. (doi:10.1038/srep05536)
53. Li A, Broom M, Du J, Wang L. 2016 Evolutionary dynamics of general group interactions in structured populations. *Phys. Rev. E* **93**, 022407. (doi:10.1103/PhysRevE.93.022407)
54. Grilli J, Barabás G, Michalska-Smith MJ, Allesina S. 2017 Higher-order interactions stabilize dynamics in competitive network models. *Nature* **548**, 210–213. (doi:10.1038/nature23273)
55. Alvarez-Rodriguez U, Battiston F, Ferraz de Arruda G, Moreno Y, Perc M, Latora V. 2021 Evolutionary dynamics of higher-order interactions in social networks. *Nat. Hum. Behav.* **5**, 586–595. (doi:10.1038/s41562-020-01024-1)
56. Han TA, Lynch S, Tran-Thanh L, Santos FC. 2018 Fostering cooperation in structured populations through local and global interference strategies. In *Proc. of the 27th Int. Joint Conf. on Artificial Intelligence and the 23rd European Conf. on Artificial Intelligence, Stockholm, Sweden, 13–19 July*, pp. 289–295. International Joint Conferences on Artificial Intelligence Organization.
57. Cimpanu T, Perret C, Han TA. 2021 Cost-efficient interventions for promoting fairness in the ultimatum game. *Knowl.-Based Syst.* **233**, 107545. (doi:10.1016/j.knsys.2021.107545)
58. Cimpanu T, Han TA, Santos FC. 2019 Exogenous rewards for promoting cooperation in scale-free networks. In *Proc. of 2019 Conf. on Artificial Life, Newcastle upon Tyne, UK, 29 July–2 August*, pp. 316–323. MIT Press.
59. Capraro V, Perc M, Vilone D. 2020 Lying on networks: the role of structure and topology in promoting honesty. *Phys. Rev. E* **101**, 032305. (doi:10.1103/PhysRevE.101.032305)
60. Kumar A, Capraro V, Perc M. 2020 The evolution of trust and trustworthiness. *J. R. Soc. Interface* **17**, 20200491. (doi:10.1098/rsif.2020.0491)
61. Wang S, Chen X, Xiao Z, Szolnoki A, Vasconcelos VV. 2023 Optimization of institutional incentives for cooperation in structured populations. Figshare. (doi:10.6084/m9.figshare.c.6384898)