Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Minority control maximizes cumulative cooperation

**Qi Su**

qisu@sjtu.edu.cn

Shanghai Jiao Tong University https://orcid.org/0000-0002-6490-686X

**Sha Song**

Shanghai Jiao Tong University

**Ziyou Gao**

Beijing Jiaotong University https://orcid.org/0000-0002-6519-4681

**Jianjun Wu**

Dalian University of Technology

**Article**

**Keywords:**

**Posted Date:** November 2nd, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-11311153/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

**Additional Declarations:** There is **NO** Competing Interest.

# Minority control maximizes cumulative cooperation

Sha Song[1] Ziyou Gao[2] Jianjun Wu [3] Qi Su[1]*

[1]School of Automation and Intelligent Sensing, Shanghai Jiao Tong University, Shanghai, 200240, China

[2]School of Systems Science, Beijing Jiaotong University, Beijing, 100044, China

[3]School of Economics and Management, Dalian University of Technology, Dalian, 116024, China

## Abstract

Designing incentives that promote cooperation is a central problem in public policy. Such incentives are typically evaluated by the stable cooperation level, which implicitly assumes that they last forever. In practice, incentives are costly and must end, after which the cooperation they built may erode. Since cooperation counts for as long as it lasts, a more practically relevant question is how to spend a limited budget to maximize the cooperation accumulated over the evolutionary process. Here, we develop a framework for maximizing cumulative cooperation in evolving populations through rewards or punishments. For general social dilemmas, including the prisoner's dilemma, snowdrift game, and stag-hunt game, we show that reaching a target state at minimum cumulative cost follows a minority-control principle: reward minority cooperators or punish minority defectors. Furthermore, under a fixed control budget, to maximize the cumulative cooperation, the optimal strategy is to apply this principle to drive the population to a desired state, then maintain that state until the budget is exhausted, and finally allow natural evolution. Our results shift the focus from maximizing cooperation at a single time point to maximizing cooperation accumulated over time.

*Correspondence: qisu@sjtu.edu.cn

# Introduction

Cooperation is fundamental to the prosperity of living systems because it enhances collective welfare, yet it is difficult to sustain when individuals prioritize private interests over collective outcomes. This tension has motivated extensive research on why cooperation emerges and how it can be promoted among self-interested individuals [1–6]. Among the mechanisms explored, incentive-based approaches seek to promote cooperation by rewarding cooperative contributions or punishing defection, thereby altering the relative payoffs of cooperation and defection [7–9]. The effectiveness of rewards and punishments in promoting cooperation has also been demonstrated experimentally [10, 11]. Because the two mechanisms target different strategies, their relative effectiveness can vary with population composition, providing a basis for coordinating institutional incentives as cooperation evolves [12, 13]. These studies provide a foundation for actively regulating the evolution of cooperation through the design and coordination of rewards and punishments.

Despite these advances, most studies have focused on the eventual state of an evolving system, typically measured by the frequency of cooperation in a single time step or at equilibrium. Much less attention has been paid to cumulative cooperation throughout the evolutionary process. This distinction is important because the viability and prosperity of many systems depend not on cooperation at a single time step, but on the total amount of cooperation sustained over time. It becomes especially critical when maintaining cooperation is costly. In real-world settings, promoting cooperation often requires regulatory investment in areas such as water-demand management [14, 15], fisheries governance [16–18], and forest conservation [19–21]. Once such costs are considered, incentive mechanisms can no longer be treated as cost-free interventions that can be applied indefinitely. A cooperative state that persists only under continuous regulation requires sustained expenditure [10, 22–24], whereas real regulatory budgets are finite [25–27]. Thus, the central question shifts from whether cooperation can be promoted to how limited regulatory resources should be allocated to maximize cumulative cooperation over time.

Here, we develop a framework for studying cumulative cooperation in an evolving population under external control through rewards or punishments. We formulate two optimal control problems: how to reach a prescribed cooperation level with minimal cumulative cost, and how to maximize cumulative cooperation under a fixed cumulative cost. For general cooperation dilemmas, including the prisoner's dilemma, snowdrift game, and stag-hunt game, we show that the least-cost strategy follows a minority control principle: rewarding minority cooperators or punishing minority defectors. We further show that, under a fixed budget, the optimal strategy is to apply this minority control principle to drive the population to an appropriate cooperation level, maintain that state until the budget is exhausted, and then allow the system to evolve naturally. This strategy maximizes cumulative cooperation throughout the evolutionary process and shows that, when sustaining cooperation is costly, the primary objective should be to maximize

cooperation accumulated over time rather than cooperation at any single time step.

# Model

We consider a well-mixed population of $N$ individuals. In each generation, every individual adopts either cooperation, denoted by C, or defection, denoted by D, and plays a pairwise game with every other individual. The payoff matrix for each interaction is

$$\begin{array}{c} \\ \mathrm{C} \\ \mathrm{D} \end{array} \begin{array}{c} \begin{array}{cc} \mathrm{C} & \mathrm{D} \end{array} \\ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \end{array}. \tag{1}$$

Mutual cooperation gives each cooperator payoff $a$, whereas mutual defection gives each defector payoff $d$. Unilateral cooperation gives payoff $b$ to the cooperator and payoff $c$ to the defector. Different orderings of $a$, $b$, $c$, and $d$ correspond to different cooperation dilemmas, including the prisoner's dilemma with $c > a > d > b$, the snowdrift game with $c > a$ and $b > d$, and the stag-hunt game with $a > c$ and $b < d$. Let $n_{\mathrm{C}}$ denote the number of cooperators in the population. Since the population is well mixed, its state can be fully described by $n_{\mathrm{C}}$.

We introduce external regulation by allowing a controller to assign an additional payoff, either positive or negative, to individuals according to their strategies and interaction contexts. We consider two classes of control modes: strategy-based control and interaction-based control. Under strategy-based control, the additional payoff received by a focal individual $i$ depends only on its own strategy. This gives rise to two control schemes. In C-based control, individual $i$ receives an additional payoff $u_{\mathrm{r}}$ if it cooperates and receives no additional payoff if it defects. In D-based control, individual $i$ receives an additional payoff $u_{\mathrm{p}}$ if it defects and receives no additional payoff if it cooperates. Under interaction-based control, the additional payoff depends on both the focal individual's strategy and its opponent's strategy. This gives rise to four schemes. In CC-based control, individual $i$ receives $u_{\mathrm{CC}}$ when both individuals cooperate. In CD-based control, individual $i$ receives $u_{\mathrm{CD}}$ when it cooperates and its opponent defects. In DC-based control, individual $i$ receives $u_{\mathrm{DC}}$ when it defects and its opponent cooperates. In DD-based control, individual $i$ receives $u_{\mathrm{DD}}$ when both individuals defect. Together, these two control modes define six basic control schemes, whose combinations and coexistence can give rise to more complex forms of regulation.

After all pairwise interactions in a generation, each individual accumulates payoffs from both game interactions and regulation, from which we compute its average payoff. For example, under C-based control, a cooperator obtains

$$\pi_{\mathrm{C}} = \frac{a(n_{\mathrm{C}} - 1) + b(N - n_{\mathrm{C}})}{N - 1} + u_{\mathrm{r}}, \tag{2}$$

whereas a defector obtains

$$\pi_{\mathrm{D}} = \frac{cn_{\mathrm{C}} + d(N - n_{\mathrm{C}} - 1)}{N - 1}. \tag{3}$$

When $u_{\mathrm{r}} = 0$, C-based control reduces to the classical evolutionary system without external regulation. We emphasize that $u_{\mathrm{r}}$ need not be constant; rather, it may vary over time or depend on the current population state (namely the fraction of cooperators in the population). Analogous payoff expressions for the other five control schemes are provided in Methods, with illustrations shown in Fig. 1.

At the end of each generation, strategy updating occurs through pairwise imitation. A random individual $i$ is selected for updating, and another random individual $j$ is selected as a role model. Individual $i$, with payoff $\pi_i$, adopts the strategy of individual $j$, with payoff $\pi_j$, with probability

$$P^{i \to j} = \frac{1}{1 + \exp\left[-\omega(\pi_j - \pi_i)\right]}. \tag{4}$$

Here, $\omega$ is the selection intensity, which determines how strongly payoff differences affect strategy updating. With the complementary probability, individual $i$ retains its current strategy.

# Results

## Driving the system to a target state at minimum cumulative cost

A general $2 \times 2$ matrix game can be classified into four types: the harmony game (HG), prisoner's dilemma (PD), snowdrift game (SG), and stag–hunt game (SH). These games exhibit different dynamical regimes. HG and PD are dominance games, in which cooperation and defection are dominant, respectively. SG is characterized by stable coexistence, whereas SH is a coordination game with bistability [28]. HG and PD have no interior equilibrium, while SG and SH each possess an interior equilibrium. These differences require distinct control treatments.

We first consider games without an interior equilibrium, namely HG and PD. For a sufficiently large population, let $x = n_{\mathrm{C}}/N$ denote the fraction of cooperators. In the absence of control, the average payoff difference between cooperators and defectors is $\Phi(x) = K_0 + K_1 x$, where $K_0 = b - d$ and $K_1 = a - b - c + d$. The fraction of cooperators then evolves according to the following replicator equation (see Methods):

$$\dot{x} = \frac{\omega}{2} x(1 - x)\Phi(x) \equiv F(x). \tag{5}$$

For each of the six control schemes illustrated in Fig. 1, $k \in \{\mathrm{r}, \mathrm{p}, \mathrm{CC}, \mathrm{CD}, \mathrm{DC}, \mathrm{DD}\}$, the repli-

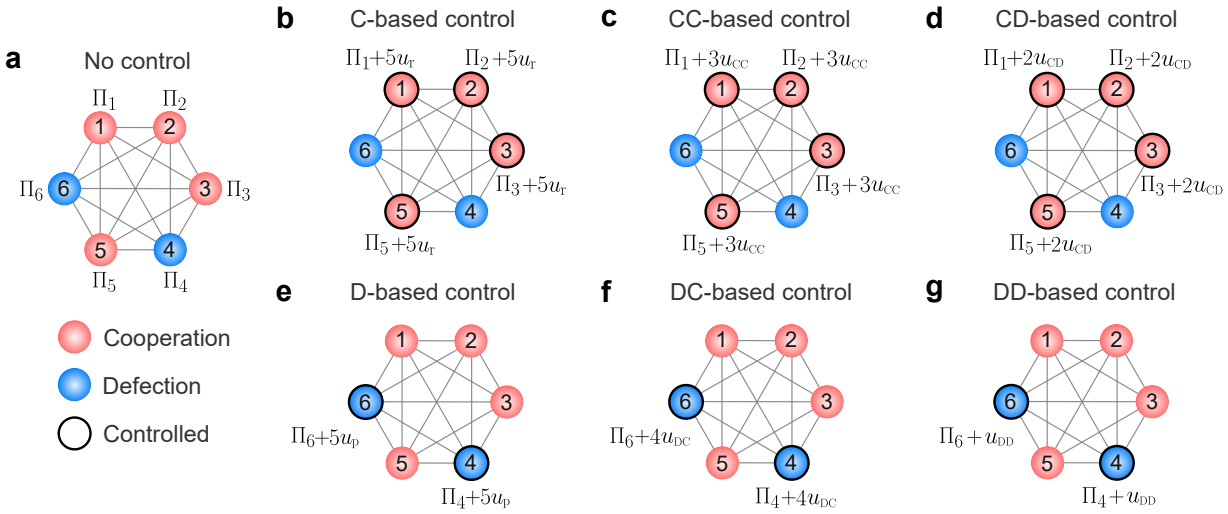



**Figure 1: Six control schemes in evolutionary games.** **a,** Game-playing interactions in the absence of external control. Red nodes denote cooperators, and blue nodes denote defectors. The payoff $\Pi_i$ shown next to each node $i$ is the baseline payoff obtained without external control. We consider two strategy-based control schemes: C-based control, in which each cooperator receives an additional payoff $u_{\mathrm{r}}$ in each interaction (**b**); and D-based control, in which each defector receives an additional payoff $u_{\mathrm{p}}$ in each interaction (**e**). We also consider four interaction-based control schemes: CC-based control, in which a cooperator receives an additional payoff $u_{\mathrm{CC}}$ when both participants cooperate (**c**); CD-based control, in which a cooperator receives an additional payoff $u_{\mathrm{CD}}$ under unilateral cooperation (**d**); DC-based control, in which a defector receives an additional payoff $u_{\mathrm{DC}}$ under unilateral defection (**f**); and DD-based control, in which a defector receives an additional payoff $u_{\mathrm{DD}}$ under mutual defection (**g**). In all other cases, no additional payoff is assigned. The panels illustrate the accumulated payoffs of selected individuals under different control schemes. For example, under C-based control, individual $i$, as a cooperator, receives a total additional payoff of $5u_{\mathrm{r}}$ from five interactions, resulting in an accumulated payoff $\Pi_i + 5u_{\mathrm{r}}$.

cator equation under control can be written in the following unified form (see Methods):

$$\dot{x} = \frac{\omega}{2} x(1-x)\left[\Phi(x) + \Lambda_k(x) u_k\right]. \tag{6}$$

Here, $u_k$ denotes the control intensity under scheme $k$, while $\Lambda_k(x)$ quantifies how this control modifies the payoff difference between cooperators and defectors. For example, under C-based control, Eqs. (2) and (3) yield $\pi_{\mathrm{C}} - \pi_{\mathrm{D}} = \Phi(x) + u_{\mathrm{r}}$, implying that $\Lambda_{\mathrm{r}}(x) = 1$. The corresponding expressions for the other five control schemes are summarized in Table S1.

We treat control as costly and define an instantaneous cost for each control scheme. For scheme $k$, the cost depends not only on the control intensity $u_k$, but also on the extent to which the control is applied. We characterize the latter by $\Gamma_k(x)$, which accounts for both the fraction of individuals receiving the additional payoff $u_k$ and the fraction of their interactions affected by the corresponding control scheme. Consider C-based control as an example. Cooperators

constitute a fraction $x$ of the population, and each cooperator receives the additional payoff $u_{\mathrm{r}}$ in every interaction. It therefore follows that $\Gamma_{\mathrm{r}}(x) = x$. The corresponding expressions for the other five control schemes are summarized in Table S1. Motivated by the quadratic cost function commonly adopted in the control of complex networks [29–31], we define the instantaneous cost under scheme $k$ as

$$U_k(x) = \left(\Gamma_k(x)u_k\right)^2. \tag{7}$$

Given the controlled replicator dynamics in Eq. (6), we determine, for each control scheme $k$, the control input $u_k$ that drives the system from an initial state $x_0$ at time $t_0$ to a prescribed target state $x_f$, while minimizing the cumulative control cost. We define the terminal time $t_f$ as the first time at which the target state is reached. Because $t_f$ is not prescribed in advance, it is treated as a free terminal time. The cumulative cost functional is given by $J_k = \int_{t_0}^{t_f} U_k(x(t))\,dt$. The resulting optimal-control problem is formulated in detail in the Methods.

For each control scheme $k$, we use dynamic programming to obtain two candidate solutions: zero control, $u_{k_1} = 0$, and active control, $u_{k_2} = -2\Phi(x)/\Lambda_k(x)$ (see Supplementary Information section 2.1). Unlike zero control, the active control is state-dependent and varies with the fraction of cooperators $x$. We next determine the conditions under which each candidate control is optimal. For games without an interior equilibrium, $\Phi(x)$ has the same sign throughout the state space: $\Phi(x) > 0$ for HG, whereas $\Phi(x) < 0$ for PD. In the absence of control, Eq. (5) implies that $\dot{x} > 0$ when $\Phi(x) > 0$, such that $x$ increases monotonically. Conversely, $\dot{x} < 0$ when $\Phi(x) < 0$, such that $x$ decreases monotonically. Therefore, when $\Phi(x)$ and $x_f - x_0$ have the same sign, the uncontrolled dynamics naturally drive the system from $x_0$ toward the target state $x_f$. Because the target can be reached without incurring any control cost, the optimal control is $u_k^* = 0$. By contrast, when $\Phi(x)$ and $x_f - x_0$ have opposite signs, the uncontrolled dynamics drive the system away from $x_f$, and active control is required to reverse the direction of evolution. Combining these two cases, the optimal control under scheme $k$ is

$$u_k^*(x) = \begin{cases} 0, & \text{sign}(\Phi(x)) = \text{sign}(x_f - x_0), \\ -\dfrac{2\Phi(x)}{\Lambda_k(x)}, & \text{sign}(\Phi(x)) \neq \text{sign}(x_f - x_0), \end{cases} \tag{8}$$

Here, $\text{sign}(\cdot)$ denotes the sign function, with $\text{sign}(z) = 1$ for $z > 0$ and $\text{sign}(z) = -1$ for $z < 0$. We assume throughout that $x_f \neq x_0$.

We next compare the state trajectories and instantaneous costs generated by the six basic control schemes under their respective optimal controls. This comparison identifies the most cost-effective scheme at each population state and thereby determines the optimal regulation strategy. We begin with the state trajectories. Substituting the optimal control in Eq. (8) into the controlled replicator dynamics in Eq. (6) shows that zero control, $u_k^*(x) = 0$, gives $\dot{x} = F(x)$, whereas the active control $u_k^*(x) = -2\Phi(x)/\Lambda_k(x)$ yields $\dot{x} = -F(x)$. Thus, active control exactly reverses

the direction of the uncontrolled dynamics while preserving their magnitude. The implicit time-dependent solution for the optimal trajectory and the explicit expression for the terminal time $t_f$ are derived in Supplementary Information Section 2.2. Neither the optimal trajectory nor the terminal time depends on the scheme-specific factors $\Lambda_k(x)$ and $\Gamma_k(x)$. Consequently, all six control schemes generate the same optimal state trajectory and reach the target state at the same terminal time, as shown in Supplementary Fig. 1. The schemes therefore differ not in the state trajectory they produce, but in the control cost required to generate that trajectory.

We then compare the instantaneous costs of the six control schemes under their respective optimal controls. When $u_k^* = 0$, the minimum instantaneous cost vanishes, $U_k^* = 0$. When $u_k^*(x) = -2\Phi(x)/\Lambda_k(x)$, the minimum instantaneous cost $U_k^*(x)$ becomes $(-2\Phi(x)x)^2$ for $k \in \{\mathrm{r}, \mathrm{CC}, \mathrm{CD}\}$, and $(2\Phi(x)(1-x))^2$ for $k \in \{\mathrm{p}, \mathrm{DC}, \mathrm{DD}\}$ (see Supplementary Information section 2.3). Thus, the three schemes that act on cooperators have the same optimal instantaneous cost, whereas the three schemes that act on defectors also have the same optimal instantaneous cost. It is therefore sufficient to compare one representative cooperator-based scheme with one representative defector-based scheme, namely the C-based control scheme r and the D-based control scheme p.

The condition $U_{\mathrm{r}}^*(x) = U_{\mathrm{p}}^*(x)$ gives $x = 0.5$. Moreover, $U_{\mathrm{r}}^*(x) < U_{\mathrm{p}}^*(x)$ for $x < 0.5$ and $U_{\mathrm{r}}^*(x) > U_{\mathrm{p}}^*(x)$ for $x > 0.5$. Because all six basic control schemes generate the same optimal state trajectory and terminal time, the scheme with the lowest instantaneous cost at each state also minimizes the cumulative control cost. The resulting optimal regulation is therefore

$$u^*(x) = \begin{cases} u_{\mathrm{r}}^*(x), & x < 0.5, \\ u_{\mathrm{p}}^*(x), & x \geq 0.5. \end{cases} \tag{9}$$

Here, $u_{\mathrm{r}}^*(x)$ and $u_{\mathrm{p}}^*(x)$ are given in Eq. (8), that is, $u_{\mathrm{r}}^* = u_{\mathrm{p}}^* = 0$ when $\mathrm{sign}(\Phi(x)) = \mathrm{sign}(x_f - x_0)$, and $u_{\mathrm{r}}^*(x) = -2\Phi(x)$ and $u_{\mathrm{p}}^*(x) = 2\Phi(x)$ when $\mathrm{sign}(\Phi(x)) \neq \mathrm{sign}(x_f - x_0)$. Equation (9) reveals a minority control principle: control should be applied to cooperators when cooperators constitute the minority, $x < 0.5$, and to defectors when defectors constitute the minority, $x > 0.5$. In other words, the minimum-cost strategy always targets the less abundant strategy. We illustrate this result using two examples in which the system is driven from a low initial cooperation level, $x_0 = 0.1$, to a high target cooperation level, $x_f = 0.9$. In HG, the uncontrolled dynamics naturally drive the system from $x_0$ toward $x_f$. No control is therefore required, and $u(t) = U(t) = J(t) = 0$ throughout the trajectory (Fig. 2a). In PD, by contrast, the natural dynamics oppose the desired increase in cooperation over the entire trajectory. The system must therefore be driven entirely by active control. As the cooperation level crosses $x = 0.5$, the optimal scheme switches from controlling cooperators to controlling defectors, in accordance with the minority control principle, as shown in Fig. 2b.

We proceed with the study of games with an interior equilibrium, namely SG and SH. Note that

for the games without an interior equilibrium considered above, the state dynamics under the optimal control can carry the system continuously from $x_0$ to $x_f$. By contrast, when the game admits an interior equilibrium $x^* = -K_0/K_1$ (determined by $\Phi(x) = K_0 + K_1 x = 0$) and $x^*$ lies between $x_0$ and $x_f$, the control problem requires additional treatment. Under the optimal controls derived above, the state dynamics take the form $\dot{x} = \pm F(x)$. Because $F(x^*) = 0$, the interior equilibrium remains non-crossable and therefore acts as a dynamical barrier between the initial and target states.

To overcome this barrier while preserving the payoff-based interpretation of the control input, we introduce a local deterministic pulse in a neighbourhood of $x^*$, inspired by pulse-control approaches developed for other dynamical systems [32–34]. We define the local equilibrium neighbourhood as $\mathcal{N}_\epsilon(x^*) = \{x : |x - x^*| < \epsilon\}$, where $\epsilon > 0$ is sufficiently small. For each control scheme $k \in \{\mathrm{r}, \mathrm{p}, \mathrm{CC}, \mathrm{CD}, \mathrm{DC}, \mathrm{DD}\}$, let $u_{k,\epsilon}$ denote the local pulse required to drive the state across $\mathcal{N}_\epsilon(x^*)$ within the local characteristic time scale $\Delta t = 2/(\omega x^*(1 - x^*)|K_1|)$ . This time scale is determined by the linearized dynamics near the interior equilibrium, as detailed in Supplementary Information Section 2.4. The required pulse is approximately

$$u_{k,\epsilon} \approx \frac{\mathrm{sign}(x_f - x_0)|K_1|\epsilon}{\Lambda_k(x^*)} \coth\left(\frac{1}{2}\right). \tag{10}$$

The local pulse therefore removes the non-crossability of the interior equilibrium while changing the dynamics only within $\mathcal{N}_\epsilon(x^*)$. The instantaneous cost induced by this pulse is $U^*_{k,\epsilon} = (|K_1|\epsilon \coth(1/2)x^*)^2$ for $k \in \{\mathrm{r}, \mathrm{CC}, \mathrm{CD}\}$, and $U^*_{k,\epsilon} = (|K_1|\epsilon \coth(1/2)(1 - x^*))^2$ for $k \in \{\mathrm{p}, \mathrm{DC}, \mathrm{DD}\}$. Thus, the three schemes acting on cooperators have identical pulse costs, as do the three schemes acting on defectors. It is therefore sufficient to compare $u_{\mathrm{r},\epsilon}$ and $u_{\mathrm{p},\epsilon}$. The optimal regulation in $\mathcal{N}_\epsilon(x^*)$ is

$$u^*_\epsilon = \begin{cases} u_{\mathrm{r},\epsilon}, & x^* < 0.5, \\ u_{\mathrm{p},\epsilon}, & x^* \geq 0.5. \end{cases} \tag{11}$$

This result extends the minority control principle to the neighbourhood of an interior equilibrium. When cooperators constitute the minority at $x^*$, the pulse is applied to cooperators. Conversely, when defectors constitute the minority, the pulse is applied to defectors. At $x^* = 0.5$, the two pulse schemes have the same cost, and either can be selected without affecting optimality.

We illustrate these results using two examples in which the interior equilibrium $x^* = 0.375$ lies between the initial state $x_0 = 0.1$ and the target state $x_f = 0.9$, thereby preventing the state from crossing continuously under the dynamics $\dot{x} = \pm F(x)$. Because $x^* < 0.5$, the optimal pulse within $\mathcal{N}_\epsilon(x^*)$ is $u^*_\epsilon = u_{\mathrm{r},\epsilon}$. In SG, the state first evolves naturally from $x_0$ toward $x^*$, and hence $u^* = 0$ for $x < x^* - \epsilon$. The pulse $u_{\mathrm{r},\epsilon}$ then drives the state across $\mathcal{N}_\epsilon(x^*)$. Once the equilibrium neighbourhood has been crossed, active control is applied to drive the system toward the target, as shown in Fig. 2c. In SH, active control is initially required for $x < x^* - \epsilon$ to drive the state out of the low-cooperation basin. The pulse $u_{\mathrm{r},\epsilon}$ then carries the state across $\mathcal{N}_\epsilon(x^*)$. Once $x > x^* + \epsilon$,

the natural dynamics drive the system toward the target, and the optimal control returns to zero, as shown in Fig. 2d.

Because regulation is costly, we further examine the minimum cumulative cost across a broader parameter space. We fix $a = 1$ and $d = 0$, and vary $b \in [-1, 1]$ and $c \in [0, 2]$. This parameter plane is partitioned into four classical game regions: HG for $b > 0$ and $c < 1$, SG for $b > 0$ and $c > 1$, SH for $b < 0$ and $c < 1$, and PD for $b < 0$ and $c > 1$. Supplementary Fig. S2 summarizes the optimal regulation strategies for these four game classes. Based on these strategies, Fig. 2e shows the minimum cumulative cost required to drive the system from $x_0 = 0.1$ to $x_f = 0.9$. As $b$ decreases and $c$ increases, cooperation becomes progressively less favorable, thereby intensifying the social dilemma and increasing the minimum cumulative cost. We also examine the time required to reach the target under optimal regulation. As shown in Fig. 2f, the terminal time is generally longer in SG and SH, where the presence of an interior equilibrium slows the controlled transition toward the target state.

## Maximizing cumulative cooperation under a fixed budget

In this section, we consider how to maximize total cooperation over a finite time interval under a fixed cost budget $\mathcal{B}$. We focus on cooperation dilemmas where cooperative behavior is disfavored, corresponding to $\Phi(x) < 0$. In the prisoner's dilemma, $\Phi(x) < 0$ throughout the state space; in the snowdrift game, $\Phi(x) < 0$ holds for $x > x^*$; and in the stag–hunt game, $\Phi(x) < 0$ holds for $x < x^*$. In these cases, cooperation would decay without external control, and therefore promoting or maintaining cooperation necessarily incurs a cost. To quantify this objective, we define cumulative cooperation, denoted by $X$, as the integral of the cooperation frequency $x(t)$ over time.

We propose a form of regulation in which the system is first driven to a certain state (defined as the reaching phase), the remaining budget is then used to maintain this state (defined as the maintenance phase), and finally the system follows its natural dynamics until cooperation declines to a small extinction threshold (defined as the uncontrolled phase). We denote by $x_f^{\mathrm{d}} \in (0, 1)$ the desired state to be reached and maintained during regulation. Whether this desired state is achievable depends on the available budget (Fig. 3a). When the budget is sufficient, the system starts from the initial state $x_0$ at time $t_0$ and reaches the desired state $x_f^{\mathrm{d}}$ at time $t_r$. It is then maintained until time $t_m$. After that, control is withdrawn, and cooperation declines to the extinction threshold at extinction time $t_e$, defined by $x(t_e) = 10^{-10}$, as shown in Fig. 3b. When the budget is insufficient, the system cannot reach the desired state $x_f^{\mathrm{d}}$. Instead, there exists a maximum attainable state, denoted by $\bar{x}(\mathcal{B})$, with $\bar{x}(\mathcal{B}) < x_f^{\mathrm{d}}$. Since no budget remains to maintain $\bar{x}(\mathcal{B})$, the system is left uncontrolled after reaching $\bar{x}(\mathcal{B})$ at $t_r$. In this case, no maintenance phase occurs, so the maintenance ending time coincides with the reaching time, $t_m = t_r$, as shown in Fig. 3c.

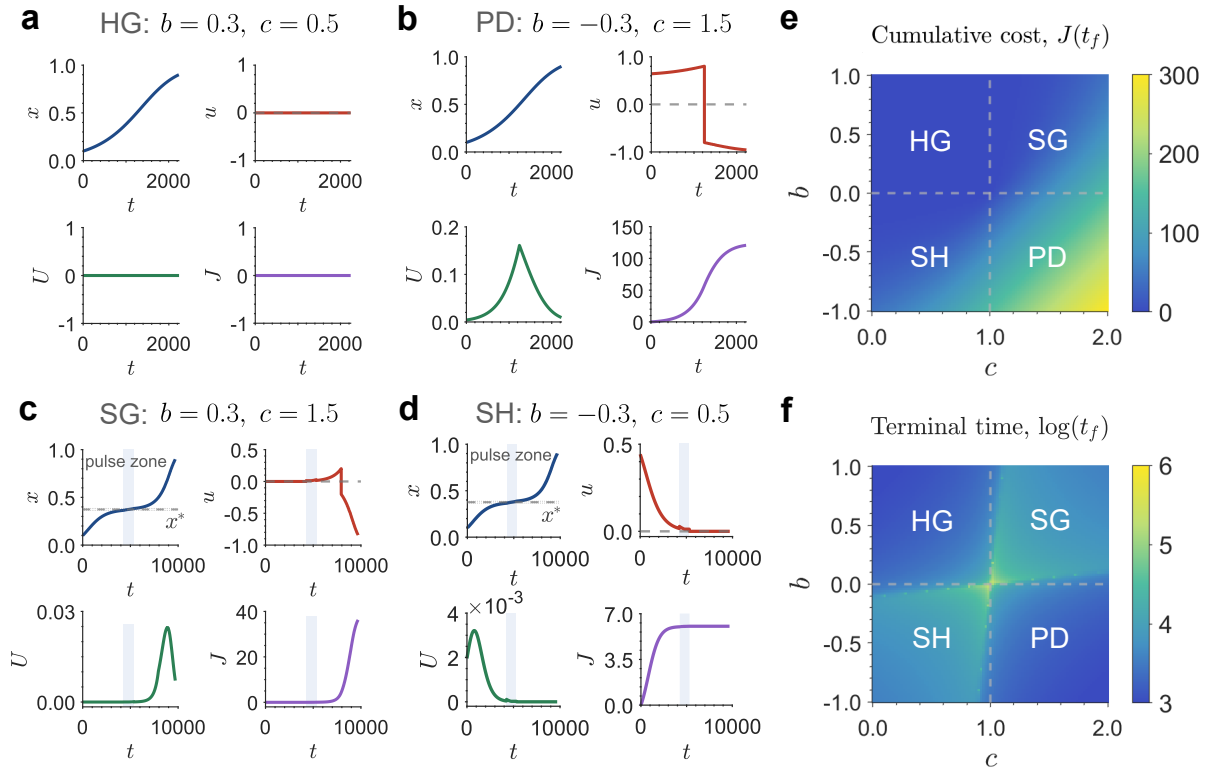



**Figure 2: Optimal regulation across general pairwise games.** **a–d,** State, control, and cost dynamics under optimal regulation for the four classes of pairwise games. In all panels, blue denotes the cooperation level $x(t)$, red the control input $u(t)$, green the instantaneous cost $U(t)$, and purple the cumulative cost $J(t)$. **a,** Harmony game, with $b = 0.3$ and $c = 0.5$. The natural dynamics drive the system toward the target state; therefore, no control is required, and $u(t) = U(t) = J(t) = 0$ throughout the trajectory. **b,** Prisoner's dilemma game, with $b = -0.3$ and $c = 1.5$. The natural dynamics oppose the desired direction of evolution over the entire trajectory, and active control is therefore required throughout. The instantaneous cost reaches its maximum at $x = 0.5$, where the optimal scheme switches from $u_{\mathrm{r}}^*$ for $x < 0.5$ to $u_{\mathrm{p}}^*$ for $x \geq 0.5$. **c,** Snowdrift game, with $b = 0.3$ and $c = 1.5$. The state initially evolves naturally toward the interior equilibrium $x^*$, with $u(t) = U(t) = 0$ until it reaches the equilibrium neighbourhood $\mathcal{N}\epsilon(x^*)$. The pale-blue band indicates the local pulse region around $x^*$, within which the optimal local pulse $u\epsilon^*$ enables the state to cross the otherwise impassable equilibrium. After crossing this region, the system continues toward the target under active control $u^*$. **d,** Stag-hunt game, with $b = -0.3$ and $c = 0.5$. Active control $u^*$ first drives the system out of the low-cooperation basin. The state then traverses $\mathcal{N}_\epsilon(x^*)$ under the optimal local pulse $u_\epsilon^*$, indicated by the pale-blue band. Once the state has crossed the equilibrium neighbourhood, such that $x > x^* + \epsilon$, the natural dynamics carry it toward the target, and both $u(t)$ and $U(t)$ return to zero. **e,** Minimum cumulative cost $J(t_f)$ required to drive the system from $x_0 = 0.1$ to $x_f = 0.9$ across the $(b, c)$ parameter plane. Grey dashed lines partition the plane into the four game regions: HG ($b > 0$, $c < 1$), SG ($b > 0$, $c > 1$), SH ($b < 0$, $c < 1$), and PD ($b < 0$, $c > 1$). Colours indicate the minimum cumulative cost, which increases as decreasing $b$ and increasing $c$ intensify the social dilemma against cooperation. **f,** Terminal time over the same parameter plane, plotted as $\log(t_f)$ to improve visual contrast. Terminal attainment is slowest in SG and SH. Parameters: $a = 1$, $d = 0$, $\omega = 0.01$, $\epsilon = 0.01$, $x_0 = 0.1$, $x_f = 0.9$.

During regulation, cooperation may need to be increased, maintained, or decreased, and we adopt different control methods for these cases. For increasing cooperation, we use the minority control principle (Eq. (9)). Here, $u_{\mathrm{r}}^*$ acts on cooperators and is interpreted as reward, whereas $u_{\mathrm{p}}^*$ acts on defectors and is interpreted as punishment. Specifically, cooperators are rewarded with $u_{\mathrm{r}}^* = -2\Phi(x)$ if $x < 0.5$, or defectors are punished with $u_{\mathrm{p}}^* = 2\Phi(x)$ if $x \geq 0.5$. For maintaining cooperation at desired state $x_f^{\mathrm{d}}$, the control needs to balance the natural drift. Thus, cooperators are rewarded with $u_{\mathrm{r}}^* = -\Phi(x_f^{\mathrm{d}})$ if $x_f^{\mathrm{d}} < 0.5$, or defectors are punished with $u_{\mathrm{p}}^* = \Phi(x_f^{\mathrm{d}})$ if $x_f^{\mathrm{d}} \geq 0.5$. In the decreasing case, no active control is applied, because the natural dynamics already reduce cooperation.

Taken together, we formulate an optimization problem: under a fixed budget $\mathcal{B}$, we solve for the desired state that maximizes cumulative cooperation $X$, thereby determining the optimal regulation. We use the same instantaneous cost as defined in Eq. (7), and the cumulative cost incurred over the reaching and maintenance phases is constrained by $\int_{t_0}^{t_m} U(x(t))\, dt = \mathcal{B}$. A detailed formulation is provided in the Methods.

We first investigate the optimal desired state in the prisoner's dilemma. In the main analysis, we restrict the admissible desired state to $x_f^{\mathrm{d}} \in [0.01, 0.99]$ and verify the robustness of the results by extending this range closer to the boundaries $0$ and $1$ in Supplementary Fig. S3. As shown in Fig. 4a, the optimal desired state remains at the low-cooperation endpoint $x_f^{\mathrm{d}} = 0.01$ under small budgets. Rather than increasing continuously with the control budget, it undergoes an abrupt transition to the high-cooperation endpoint $x_f^{\mathrm{d}} = 0.99$, once the budget exceeds the critical value $\mathcal{B}_c = 125.6$. The optimal desired state is defined as the value of $x_f^{\mathrm{d}}$ that maximizes cumulative cooperation $X(x_f^{\mathrm{d}})$. Accordingly, the low-cooperation boundary is optimal when $X(0.01) > X(0.99)$, whereas the high endpoint $x_f^{\mathrm{d}} = 0.99$ is optimal when $X(0.99) > X(0.01)$. As the budget increases, the difference $X(0.01) - X(0.99)$ changes from positive to negative at $\mathcal{B}_c$, marking the transition from maintaining a low level of cooperation to locking the system at a high level of cooperation (Fig. 4b, upper panel). At the critical budget, $X$ depends non-monotonically on $x_f^{\mathrm{d}}$ and attains its largest values at the two boundaries of the admissible interval (Fig. 4b, lower panel). Notably, the low desired state can be maintained for a longer duration across the entire budget range. The transition therefore does not occur because the low-cooperation state loses its lifetime advantage. Instead, beyond the critical budget, the substantially higher cooperation level associated with $x_f^{\mathrm{d}} = 0.99$ compensates for its shorter maintenance duration and ultimately produces greater cumulative cooperation.

The trajectories at the critical budget further reveal how the available budget is allocated under the two optimal solutions. For $x_f^{\mathrm{d}} = 0.01$, the desired state is reached without control because the natural dynamics already drive cooperation downward. The entire budget can therefore be devoted to maintaining the low-cooperation state through a positive reward applied to cooperators. This produces a long controlled lifetime and cumulative cooperation of approximately $X = 1.378 \times 10^5$, as shown in Fig. 4c. For $x_f^{\mathrm{d}} = 0.99$, by contrast, active control must first drive

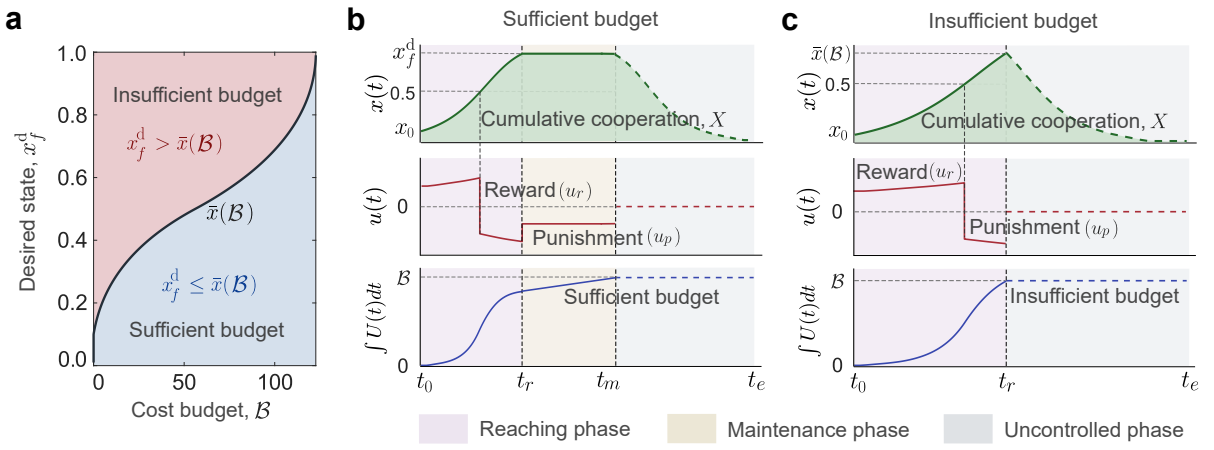



**Figure 3: The optimal regulation of cumulative cooperation under a fixed budget. a,** Attainable desired states under budget constraints. The horizontal axis denotes the cost budget $\mathcal{B}$, and the vertical axis denotes the desired state $x_f^{\mathrm{d}}$. The black curve shows the maximum attainable state $\bar{x}(\mathcal{B})$, derived from the minimum-cost control analysis. Desired states below the curve, $x_f^{\mathrm{d}} \leq \bar{x}(\mathcal{B})$, are attainable because the minimum cost required to reach them does not exceed the available budget (the blue region), whereas desired states above the curve, $x_f^{\mathrm{d}} > \bar{x}(\mathcal{B})$, are infeasible because the minimum cost required to reach them exceeds the budget (the pink region). **b,** Schematic illustration of the regulation process under a sufficient budget. The system starts from the initial state $x_0$ at time $t_0$ and reaches the desired target $x_f^{\mathrm{d}}$ at time $t_r$. The state is then maintained until time $t_m$. After that, the system evolves naturally until time $t_e$. The upper panel shows the state trajectory, with the green shaded area denoting cumulative cooperation $X$. The middle panel shows the optimal control that depends on the state, switching from rewarding cooperators $u_{\mathrm{r}}^*$ to punishing defectors $u_{\mathrm{p}}^*$ after $x = 0.5$. The lower panel shows that the cumulative cost reaches the budget $\mathcal{B}$ at the end of the maintenance phase. **c,** Schematic illustration of the regulation process under an insufficient budget. When the desired target exceeds the attainable limit under the available budget $\mathcal{B}$, the system can only reach the maximum attainable state $\bar{x}(\mathcal{B})$ at $t_r$, and then evolves naturally until $t_e$. The middle panel shows that the control follows the minority control principle during the reaching phase. The lower panel shows that the budget is exhausted at time $t_r$, so no maintenance phase occurs. In panels **b,c**, solid line segments denote evolution under control, dashed line segments denote evolution without control, and purple, yellow and grey backgrounds denote the reaching, maintenance, and uncontrolled phases, respectively. Parameters: $a = 1$, $b = -0.3$, $c = 1.5$, $d = 0$, $\omega = 0.01$, $x_0 = 0.1$.

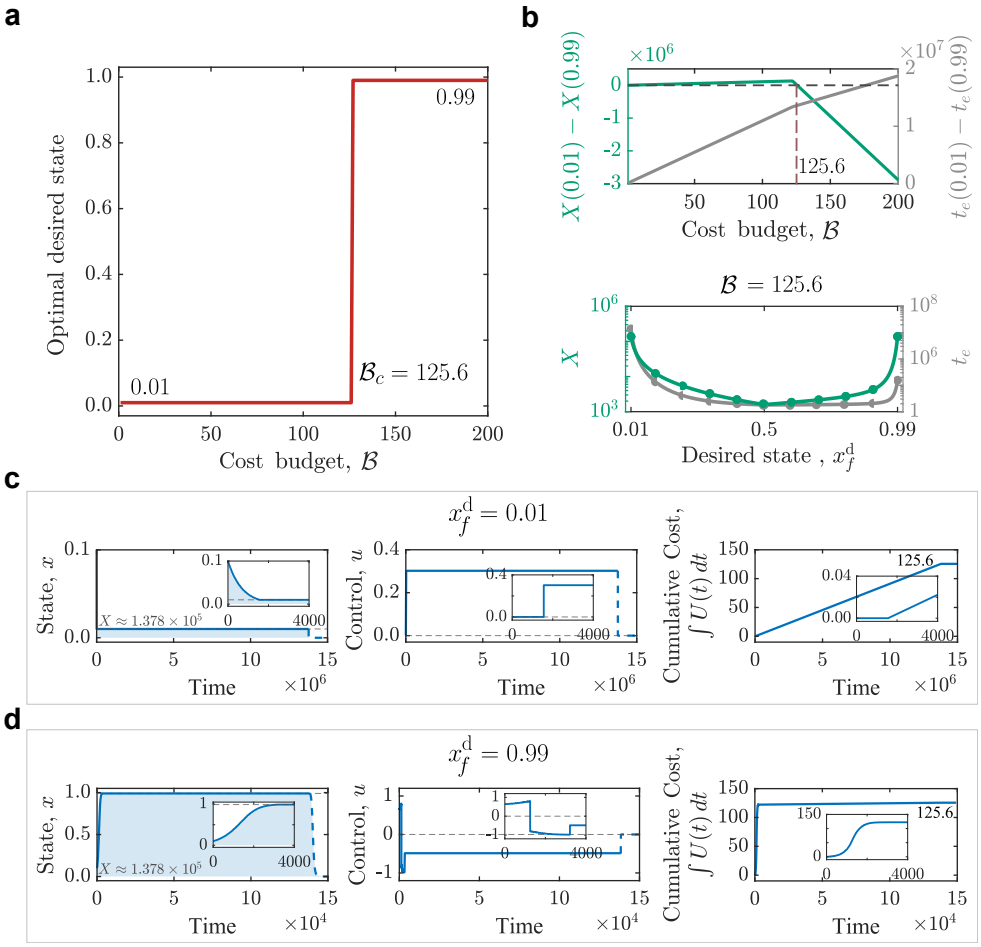



**Figure 4: Optimal desired state shifts from low cooperation to high cooperation in prisoner's dilemma.** **a,** Optimal desired state $x_f^{\mathrm{d}}$ as a function of the cost budget $\mathcal{B}$. Under small budgets, the optimum remains at the low-cooperation endpoint $x_f^{\mathrm{d}} = 0.01$, but it switches abruptly to the high-cooperation endpoint $x_f^{\mathrm{d}} = 0.99$ once the budget exceeds the critical threshold $\mathcal{B}_c = 125.6$. **b,** The optimal desired state is determined by which endpoint, $x_f^{\mathrm{d}} = 0.01$ or $x_f^{\mathrm{d}} = 0.99$, yields greater cumulative cooperation. The upper panel shows the difference in cumulative cooperation, $X(0.01) - X(0.99)$ (green, left axis), and the difference in extinction time, $t_e(0.01) - t_e(0.99)$ (grey, right axis). The brown dashed line marks $\mathcal{B} = 125.6$. The sign change of $X(0.01) - X(0.99)$ identifies the optimal desired state, with positive values favouring $x_f^{\mathrm{d}} = 0.01$ and negative values favouring $x_f^{\mathrm{d}} = 0.99$. By contrast, $t_e(0.01) - t_e(0.99) > 0$ throughout, indicating that the low desired state lasts longer. The lower panel shows, at the critical budget $\mathcal{B}_c$, cumulative cooperation $X$ and extinction time $t_e$ as functions of $x_f^{\mathrm{d}}$. Both vary non-monotonically, with cumulative cooperation maximized at the two endpoints. **c,d,** At the critical budget $\mathcal{B}_c$, trajectories of the state, control and cumulative cost are shown for $x_f^{\mathrm{d}} = 0.01$ and $x_f^{\mathrm{d}} = 0.99$, respectively. **c,** For $x_f^{\mathrm{d}} = 0.01$, the system first decays to the low desired state without control. The budget is then used to maintain this state by rewarding cooperators until it is exhausted. The blue shaded area denotes cumulative cooperation, $X \approx 1.378 \times 10^5$. **d,** For $x_f^{\mathrm{d}} = 0.99$, active control rapidly drives the system towards the high desired state. As cooperation increases, the optimal control switches from reward to punishment. Most of the budget is spent during the reaching phase, after which high desired state is maintained until the budget is exhausted. The blue shaded area denotes cumulative cooperation, $X \approx 1.378 \times 10^5$. Insets in **c,d** show early-time dynamics. Parameters: $a = 1$, $b = -0.3$, $c = 1.5$, $d = 0$, $\omega = 0.01$, $x_0 = 0.1$.

the system against its natural evolutionary drift. Following the minority control principle, the regulation switches from rewarding cooperators to punishing defectors as the cooperation level crosses $x = 0.5$. Most of the available budget is consumed during this reaching phase, leaving only a small fraction for maintaining the high-cooperation state (Fig. 4d). The early-time insets in Figs. 4c,d highlight these contrasting patterns of budget allocation. Despite their markedly different trajectories and control strategies, the two endpoint states yield nearly identical cumulative cooperation at the critical budget.

We then investigate the optimal desired state in the snowdrift game. In the case shown in Fig. 5a, $x_0 < x^* - \epsilon$, so reaching this region requires a local pulse across $\mathcal{N}_\epsilon(x^*)$. The pulse incurs only a small cumulative cost, denoted by $J_\epsilon$ ($J_\epsilon \approx 0.045$ in Fig. 5b). Once the pulse enables the system to cross $\mathcal{N}_\epsilon(x^*)$, admissible desired states are restricted to $x_f^{\mathrm{d}} \in [x^* + \epsilon, 0.99]$. The schematic in Fig. 5a shows the corresponding process. The state first approaches the stable interior equilibrium, crosses the pulse neighbourhood, and reaches the desired state. It is then maintained until the budget is exhausted, after which the system evolves without control. Since the system would remain near this equilibrium in the absence of control, performance is measured by the cooperation gain $\Delta X$. The results show that, in the snowdrift game, the optimal desired state switches discontinuously from $x^* + \epsilon$ to $0.99$ at the critical budget $\mathcal{B}_c = 39.44$ (Fig. 5b). At this critical budget, $\Delta X$ first decreases and then increases as $x_f^{\mathrm{d}}$ varies, attaining maxima at the two admissible endpoints $x^* + \epsilon$ and $0.99$ (Fig. 5c).

We now investigate the optimal desired state in the stag–hunt game. The key distinction is whether the budget is sufficient to cross the unstable interior equilibrium $x^*$. When the budget is too small for equilibrium crossing, the state is driven to a desired state $x_f^{\mathrm{d}} \in [0.01, x^* - \epsilon]$, maintained there until the budget is exhausted, and then follows the natural dynamics, as shown by the green trajectory in Fig. 5d. When the budget is sufficiently large, the system is first driven to $x^* - \epsilon$ and then crosses the equilibrium neighbourhood under pulse control. After reaching $x^* + \epsilon$, cooperation increases naturally without further control, as shown by the pale-green trajectory in Fig. 5d. Reaching $x^* + \epsilon$ requires a cumulative cost of $5.97$, of which $J_\epsilon \approx 0.045$ is spent on the pulse control (Fig. 2d). This amount is also the critical budget $\mathcal{B}_c$ at which the optimal desired state switches abruptly. When $\mathcal{B} < \mathcal{B}_c$, the optimal desired state is $x_f^{\mathrm{d}} = 0.01$. When $\mathcal{B} \geq \mathcal{B}_c$, the system can evolve toward higher cooperation without further cost after reaching $x^* + \epsilon$, and the optimal desired state is therefore $x_f^{\mathrm{d}} \in [x^* + \epsilon, 1)$ (Fig. 5e). At the critical budget $\mathcal{B}_c$, cumulative cooperation $X$ first decreases and then increases with $x_f^{\mathrm{d}}$ over $[0.01, x^* - \epsilon]$, yet remains largest at $x_f^{\mathrm{d}} = 0.01$ (Fig. 5f). Together, Figs. 4 and 5 show that, under a fixed budget, optimal regulation exhibits two distinct regimes across cooperation dilemmas. When the budget is small, optimal regulation maintains low cooperation for a long duration; when the budget is sufficiently large, it locks in high cooperation.

We next optimize over a broader class of regulation to determine whether the continuous-maintenance strategy discussed above remains optimal when intermittent regulation is permitted. In this class,

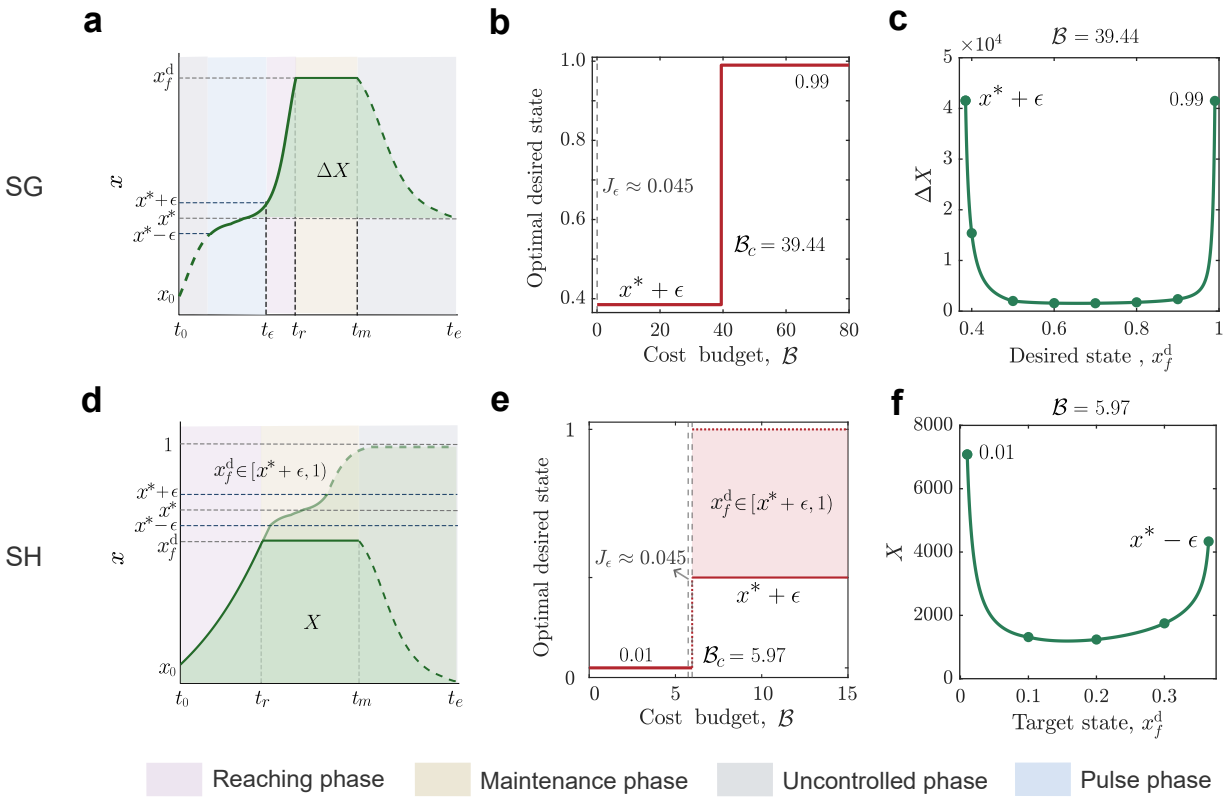



**Figure 5: The optimal desired state shifts abruptly from low to high cooperation in snowdrift and stag-hunt games.** Panels **a–c** show the snowdrift game (SG), and panels **d–f** show the stag-hunt game (SH). In panels **a,d**, solid and dashed lines indicate the state trajectory with and without control, respectively. Purple, yellow and grey backgrounds denote reaching, maintenance and uncontrolled phases, respectively, and blue marks the pulse phase in SG. **a,** Schematic for SG. Starting from $x_0 < x^* - \epsilon$, the state first approaches the neighbourhood of the interior equilibrium $x^*$ without control. It then crosses this neighbourhood through a local pulse, reaches the desired state $x_f^{\mathrm{d}} \in [x^* + \epsilon, 0.99]$, and is maintained there until $t_m$. After the budget is exhausted, the state evolves without control until $t_e$. The green shaded area denotes the cooperation gain $\Delta X$ relative to the trajectory without control. **b,** Optimal desired state $x_f^{\mathrm{d}}$ as a function of the cost budget $\mathcal{B}$, with $x_f^{\mathrm{d}} \in [x^* + \epsilon, 0.99]$. The optimum switches abruptly from $x^* + \epsilon$ to $0.99$ at critical budget $\mathcal{B}_c = 39.44$. The grey dashed line marks the cumulative cost of the local pulse, $J_\epsilon \approx 0.045$. **c,** At $\mathcal{B} = 39.44$, $\Delta X$ first decreases and then increases with $x_f^{\mathrm{d}}$, reaching its maxima at the two admissible endpoints, $x^* + \epsilon$ and $0.99$. **d,** Schematic for SH. The system may follow two possible evolutionary trajectories, depending on the available budget. When the budget is insufficient to cross the interior equilibrium, the state is driven to and maintained at $x_f^{\mathrm{d}} \in [0.01, x^* - \epsilon]$, and then evolves without control until $t_e$. The green shaded area denotes the cumulative cooperation $X$. When the budget is sufficient, the state is first driven to $x^* - \epsilon$, then reaches $x^* + \epsilon$ under pulse control, and subsequently evolves naturally toward higher cooperation without further control. The corresponding cumulative cooperation $X$ is indicated by the pale-green shaded area. **e,** Optimal desired state $x_f^{\mathrm{d}}$ as a function of the cost budget $\mathcal{B}$. The critical budget is $\mathcal{B}_c = 5.97$, with a cumulative pulse cost of $J_\epsilon \approx 0.045$. When $\mathcal{B} < \mathcal{B}_c$, the optimal desired state is $x_f^{\mathrm{d}} = 0.01$. When $\mathcal{B} \geq \mathcal{B}_c$, all states $x_f^{\mathrm{d}} \in [x^* + \epsilon, 1)$ can be regarded as optimal desired states. **f,** At $\mathcal{B} = 5.97$, $X$ first decreases and then increases with $x_f^{\mathrm{d}}$, reaching its maximum at $x_f^{\mathrm{d}} = 0.01$. Parameters: $a = 1$, $d = 0$, $\omega = 0.01$, $\epsilon = 0.01$, $x_0 = 0.1$, with $b = 0.3$, $c = 1.5$ for SG and $b = -0.3$, $c = 0.5$ for SH.

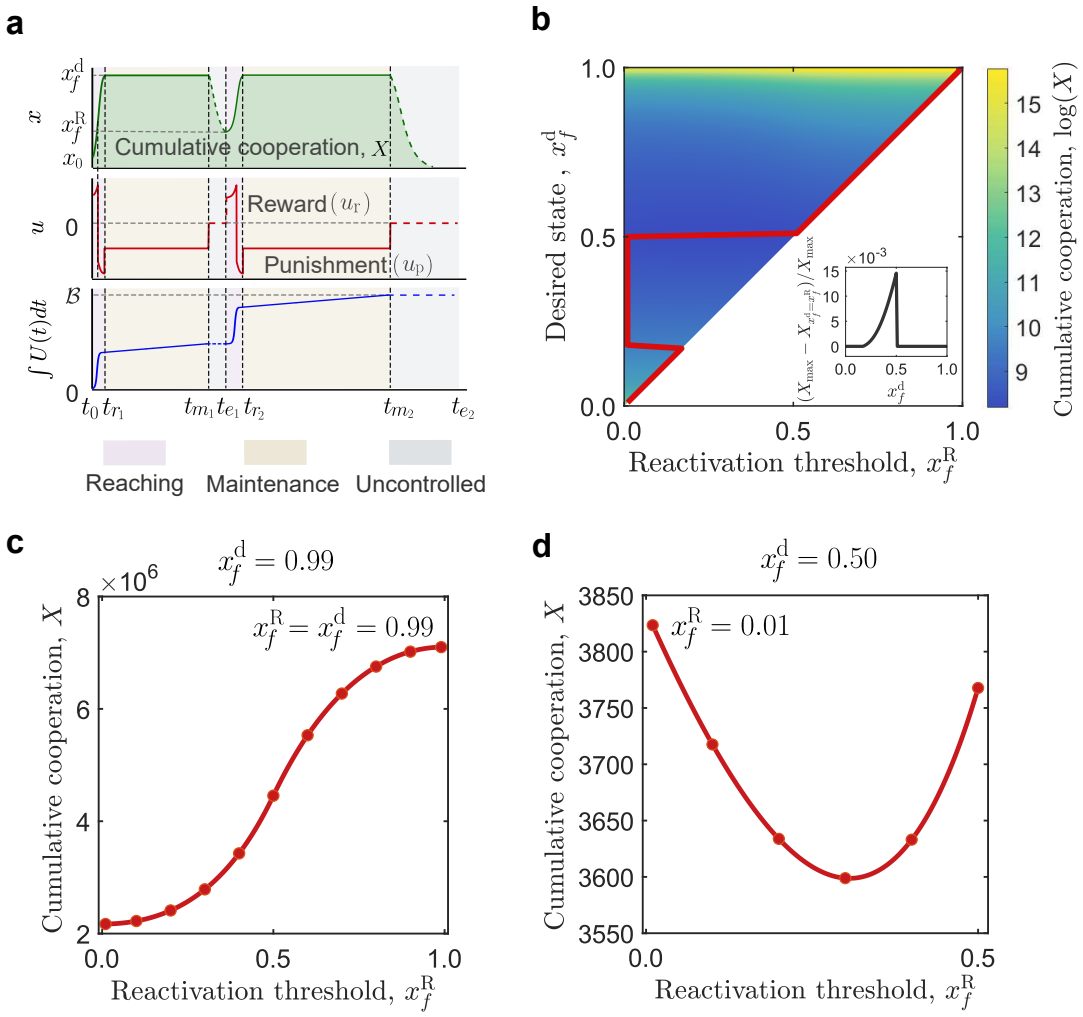



**Figure 6: The regulation with continuous maintenance remains optimal or nearly optimal.** **a,** Schematic of intermittent regulation under a sufficient budget. Starting from the initial state $x_0$ at $t_0$, the system reaches the desired state $x_f^{\mathrm{d}}$ at $t_{r_1}$ and maintains it until $t_{m_1}$. Control is then withdrawn, and the system decays to the reactivation threshold $x_f^{\mathrm{R}}$ at $t_{e_1}$, before control drives it back to $x_f^{\mathrm{d}}$ at $t_{r_2}$. The desired state is maintained again until the budget is exhausted at $t_{m_2}$, after which the system undergoes a final decay without control until $t_{e_2}$. Solid and dashed lines indicate the evolution with and without control, respectively. Purple, yellow and grey backgrounds denote reaching, maintenance and uncontrolled phases, respectively. The green shaded area denotes cumulative cooperation $X$, and the lower panels show the corresponding control $u(t)$ and cumulative cost $\int U(t)\,dt$. **b**, Heatmap of $\log(X)$ across the desired state $x_f^{\mathrm{d}}$ and the reactivation threshold $x_f^{\mathrm{R}}$. The red curve marks the value of $x_f^{\mathrm{R}}$ that maximizes $X$ for each $x_f^{\mathrm{d}}$. The optimum typically lies either at continuous maintenance, $x_f^{\mathrm{R}} = x_f^{\mathrm{d}}$, or at the low-cooperation endpoint, $x_f^{\mathrm{R}} = 0.01$. The inset shows that continuous maintenance remains nearly optimal: for each given $x_f^{\mathrm{d}}$, the relative gap $(X_{\max} - X_{x_f^{\mathrm{R}}=x_f^{\mathrm{d}}})/X_{\max}$ is negligible, where $X_{\max}$ denotes the maximum cumulative cooperation obtained by optimizing over $x_f^{\mathrm{R}}$, and $X_{x_f^{\mathrm{R}}=x_f^{\mathrm{d}}}$ denotes the cumulative cooperation on the continuous-maintenance line $x_f^{\mathrm{R}} = x_f^{\mathrm{d}}$. **c,** For $x_f^{\mathrm{d}} = 0.99$, cumulative cooperation increases with $x_f^{\mathrm{R}}$ and is maximized by continuous maintenance, $x_f^{\mathrm{R}} = x_f^{\mathrm{d}} = 0.99$. **d,** For $x_f^{\mathrm{d}} = 0.5$, cumulative cooperation is maximized at $x_f^{\mathrm{R}} = 0.01$, but the improvement over continuous maintenance, $x_f^{\mathrm{R}} = x_f^{\mathrm{d}} = 0.5$, is marginal. Parameters: $a = 1$, $b = -0.3$, $c = 1.5$, $d = 0$, $x_0 = 0.1$, $\mathcal{B} = 300$.

control is temporarily withdrawn during the maintenance phase and then restored to drive the state back to the desired state. Specifically, the system is first driven to $x_f^{\mathrm{d}}$ and maintained there. Control is then withdrawn, allowing the state to decay to the reactivation threshold $x_f^{\mathrm{R}}$. Once this threshold is reached, control is applied again to restore the state to $x_f^{\mathrm{d}}$. The remaining budget is then spent on maintaining $x_f^{\mathrm{d}}$ until it is exhausted, after which the system evolves without control (Fig. 6a). The continuous-maintenance strategy is recovered as the special case $x_f^{\mathrm{R}} = x_f^{\mathrm{d}}$. More broadly, regulation involving multiple cycles of decline, maintenance, and restoration can also be studied within this framework.

We investigate which form of regulation within this broader class maximizes cumulative cooperation $X$ and find that the continuous-maintenance strategy discussed above remains optimal or nearly optimal. Because the dynamics of cooperation decline and restoration are fixed, $X$ does not depend on when control is withdrawn, but only on the reactivation threshold. As shown in Fig. 6b, across the $(x_f^{\mathrm{R}}, x_f^{\mathrm{d}})$ plane, the maximum of cumulative cooperation occurs either on the continuous-maintenance line $x_f^{\mathrm{R}} = x_f^{\mathrm{d}}$ or at the low-cooperation endpoint $x_f^{\mathrm{R}} = 0.01$. The inset further shows that the relative gain over continuous maintenance remains negligible, indicating that allowing the system to decay and then restoring it almost never yields a meaningful gain in cumulative cooperation. To examine how cumulative cooperation varies with the reactivation threshold at fixed desired states, we consider two representative cases in which continuous maintenance is either exactly optimal or nearly optimal. For the high desired state $x_f^{\mathrm{d}} = 0.99$, cumulative cooperation increases with $x_f^{\mathrm{R}}$ and is maximized at $x_f^{\mathrm{R}} = x_f^{\mathrm{d}} = 0.99$, so any decay from the high state reduces cumulative cooperation (Fig. 6c). For the intermediate desired state $x_f^{\mathrm{d}} = 0.5$, cumulative cooperation is maximized at $x_f^{\mathrm{R}} = 0.01$, but the improvement over continuous maintenance at $x_f^{\mathrm{R}} = x_f^{\mathrm{d}} = 0.5$ is marginal (Fig. 6d). These results show that the regulation with continuous maintenance considered above is robust and nearly optimal. The same conclusion also holds for the snowdrift and stag-hunt games (Supplementary Fig. S4).

# Discussion

Cooperation contributes to collective welfare, yet it often conflicts with individual interests and is therefore disfavoured in social dilemmas. This makes cooperation difficult to sustain without regulatory support. Although many mechanisms for promoting cooperation have been proposed, their evaluation has often focused on the eventual state of evolving systems, while treating regulation as cost-free. In practice, cumulative cooperation throughout the evolutionary process is crucial, maintaining cooperation is costly, and regulatory budgets are finite. To this end, we formulate two optimization problems. The first identifies the optimal regulation for driving the system to a target state at minimum cumulative cost. By comparing six control schemes, including two strategy-based schemes and four interaction-based schemes, we show that the optimal regulation follows a minority control principle: rewarding minority cooperators or punishing

minority defectors. The minority control principle also holds in finite populations (see Supplementary Information section 3 and Supplementary Fig. S5). The second optimization problem identifies the regulation that maximizes cumulative cooperation over time under a fixed cost budget. We find that the optimal regulation is to apply the minority control principle to drive the system to a desired state (low when the budget is small and high when it is large), maintain it there until the budget is exhausted, and then allow natural evolution.

When regulatory budgets are finite, maximizing cumulative cooperation requires balancing the cooperation level achieved against how long it can be maintained. For example, in common-pool resource governance, a short burst of intensive monitoring may temporarily increase compliance, but sustained moderate enforcement can preserve resource use over a longer horizon [35, 36]. Similarly, in climate governance, large short-term subsidies may rapidly increase participation in mitigation efforts, whereas limited public budgets often need to be allocated over time to sustain emission reduction, monitoring and compliance [37–39]. Existing studies have mainly focused on increasing cooperation, without considering how finite regulatory budgets should be allocated to maintain it over time [9, 13, 40–42]. Our results show that pursuing the highest possible cooperation level is not always optimal. Under small budgets, maintaining low cooperation for longer can yield greater cumulative cooperation than inducing a brief increase to a high level that cannot be sustained. Moreover, the model predicts that cumulative cooperation increases as the desired state approaches $0$ or $1$. In reality, however, strategy mutation, random exploration, and environmental fluctuations make such homogeneous endpoint states difficult to achieve. Even so, the model provides useful practical guidance: with insufficient budgets, cooperation should remain low; with sufficient budgets, high cooperation should be maintained.

We examine the optimal regulation for reaching a target state at minimum cumulative cost across general pairwise games, and show that it differs between games without an interior equilibrium (the harmony game and the prisoner's dilemma) and those with an interior equilibrium (the snowdrift and stag-hunt games). In the harmony game and the prisoner's dilemma, no interior equilibrium blocks the dynamics, and control can continuously drive the system toward target state. In the snowdrift and stag-hunt games, however, the optimal control approaches zero at the interior equilibrium, making crossing impossible. We therefore introduce a local pulse within an arbitrarily small neighbourhood of the equilibrium, which enables the system to cross it while preserving the payoff interpretation of the original control term. From a practical perspective, transient intensification of intervention is also common in real-world regulation, as illustrated by threshold-triggered and large-scale non-pharmaceutical interventions during epidemic control [43–46].

Studies on the control of complex systems determine how external inputs should be applied to steer a system between states with limited resources [30, 47–51]. However, the controlled units are typically system components rather than individuals capable of autonomous decision-making [51–54]. Evolutionary game theory captures such strategic behaviour, but it has mainly

been used to identify mechanisms that promote cooperation rather than to optimize regulation under limited resources. Our framework fills this gap by incorporating costly regulation into evolutionary games. To quantify regulatory expenditure, we use a quadratic cost, a standard choice in control theory [29–31]. Other classical forms, such as absolute-value costs, can also be accommodated within our framework. Within this framework, we address two closely related optimization problems: reaching a prescribed cooperation target at minimum cumulative cost and maximizing cumulative cooperation under a fixed budget.

This work develops an optimal control framework for regulating the evolution of cooperation through rewards and punishments, jointly accounting for the effectiveness of promoting cooperation and the cost of implementing regulatory mechanisms. Nevertheless, the present framework is a minimal model designed to capture these key factors, and several limitations remain to be addressed in future work. First, allowing regulation to target only a fraction of cooperators or defectors could help identify the optimal control proportion. Second, extending the analysis beyond well-mixed populations to populations with asymmetric interactions or multilayer structures [55, 56] could reveal how social structure affects control placement. Third, incorporating higher-order interactions [57, 58], personalized update rates [59], and diverse behavioral motivations [60] could further clarify how group interactions and individual heterogeneity affect the optimal allocation of finite regulatory budgets.

# Methods

**Average payoffs under six control schemes.** In a well-mixed population of $N$ individuals, suppose that the population contains $n_{\mathrm{C}}$ cooperators and $N - n_{\mathrm{C}}$ defectors. Each individual interacts with all other $N-1$ individuals. We consider six basic control schemes, which modify the average payoffs of cooperators or defectors.

Under C-based control, a cooperator receives the additional payoff $u_{\mathrm{r}}$ in all interactions. Thus,

$$\pi_{\mathrm{C}} = \frac{(a+u_{\mathrm{r}})(n_{\mathrm{C}}-1)+(b+u_{\mathrm{r}})(N-n_{\mathrm{C}})}{N-1}, \tag{12}$$

whereas a defector obtains

$$\pi_{\mathrm{D}} = \frac{cn_{\mathrm{C}}+d(N-n_{\mathrm{C}}-1)}{N-1}. \tag{13}$$

Under D-based control, a cooperator obtains

$$\pi_{\mathrm{C}} = \frac{a(n_{\mathrm{C}}-1)+b(N-n_{\mathrm{C}})}{N-1}, \tag{14}$$

whereas every defector receives the additional payoff $u_{\mathrm{p}}$ in all interactions. Thus,

$$\pi_{\mathrm{D}} = \frac{(c+u_{\mathrm{p}})n_{\mathrm{C}} + (d+u_{\mathrm{p}})(N-n_{\mathrm{C}}-1)}{N-1}. \tag{15}$$

Under CC-based control, a cooperator receives the additional payoff $u_{\mathrm{CC}}$ only when interacting with another cooperator. Thus,

$$\pi_{\mathrm{C}} = \frac{(a+u_{\mathrm{CC}})(n_{\mathrm{C}}-1) + b(N-n_{\mathrm{C}})}{N-1}, \tag{16}$$

whereas a defector obtains the payoff given in Eq. (13).

Under CD-based control, a cooperator receives the additional payoff $u_{\mathrm{CD}}$ only when interacting with a defector. Thus,

$$\pi_{\mathrm{C}} = \frac{a(n_{\mathrm{C}}-1) + (b+u_{\mathrm{CD}})(N-n_{\mathrm{C}})}{N-1}, \tag{17}$$

whereas a defector obtains the payoff given in Eq. (13).

Under DC-based control, a cooperator obtains the payoff given in Eq. (14). A defector receives the additional payoff $u_{\mathrm{DC}}$ only when interacting with a cooperator. Thus,

$$\pi_{\mathrm{D}} = \frac{(c+u_{\mathrm{DC}})n_{\mathrm{C}} + d(N-n_{\mathrm{C}}-1)}{N-1}. \tag{18}$$

Under DD-based control, a cooperator obtains the payoff given in Eq. (14). A defector receives the additional payoff $u_{\mathrm{DD}}$ only when interacting with another defector. Thus,

$$\pi_{\mathrm{D}} = \frac{cn_{\mathrm{C}} + (d+u_{\mathrm{DD}})(N-n_{\mathrm{C}}-1)}{N-1}. \tag{19}$$

The details are provided in Supplementary Information section 1.1.

**Replicator dynamics.** The deterministic population dynamics follow from the pairwise imitation process defined in Eq. (4). Let $x = n_{\mathrm{C}}/N$ denote the fraction of cooperators. After rescaling time as $t = \tau/N$ and taking the deterministic limit $N \to \infty$, the expected rate of change in $x$ is

$$\dot{x} = x(1-x)\tanh\left[\frac{\omega}{2}\left(\pi_{\mathrm{C}} - \pi_{\mathrm{D}}\right)\right]. \tag{20}$$

Under weak selection, $\tanh(z) \approx z$, and the dynamics reduce to the replicator equation.

In the absence of control, the payoff difference in the large-population limit is

$$\pi_{\mathrm{C}} - \pi_{\mathrm{D}} = \Phi(x) = K_0 + K_1 x, \tag{21}$$

where $K_0 = b - d$ and $K_1 = a - b - c + d$. The replicator equation without control is therefore

$$\dot{x} = \frac{\omega}{2} x(1-x)\Phi(x). \tag{22}$$

Under control scheme $k$, the additional payoff changes the payoff difference by $\Lambda_k(x)u_k$, giving

$$\pi_{\mathrm{C}} - \pi_{\mathrm{D}} = \Phi(x) + \Lambda_k(x)u_k. \tag{23}$$

Here, $\Lambda_k(x)$ measures the change in the payoff difference per unit of control under scheme $k$. The expressions for all six schemes follow directly from the corresponding average payoffs given above. Substituting Eq. (23) into the replicator equation gives

$$\dot{x} = \frac{\omega}{2} x(1-x)\left[\Phi(x) + \Lambda_k(x)u_k\right]. \tag{24}$$

The full technical details are provided in Supplementary Information section 1.2.

**Driving the system to a target state at minimum cumulative cost.** Under control scheme $k$, the replicator equation changes from Eq. (22) to Eq. (24). For notational simplicity, we denote the replicator equation under control by $\dot{x} = F(x, u_k)$. Implementing control incurs a cost. The instantaneous cost at time $t$ is denoted by $U_k(x(t))$, as defined in Eq. (7). The cumulative cost over a time interval is therefore obtained by integrating this instantaneous cost $J_k = \int U_k(x(t))dt$.

We then formulate the optimal control problem as driving the system from an initial state $x_0$ at time $t_0$ to a prescribed target state $x_f$ at a free terminal time $t_f$, while minimizing the cumulative cost. For control scheme $k$, this problem is written as

$$\begin{aligned} &\min_{u_k} J_k = \int_{t_0}^{t_f} U_k(x(t))\, dt, \\ &\text{s.t.} \begin{cases} \dot{x} = F(x, u_k), \\ x(t_0) = x_0, \\ x(t_f) = x_f. \end{cases} \end{aligned} \tag{25}$$

We solve this problem using dynamic programming. Let $J_k^*(x)$ denote the minimum cumulative cost of reaching the target state $x_f$ from state $x$. Because the system is autonomous and the terminal time is free, $J_k^*(x)$ is time-independent and satisfies the stationary Hamilton–Jacobi–Bellman (HJB) equation

$$\min_{u_k} \left\{ U_k(x) + \frac{\partial J_k^*(x)}{\partial x} F(x, u_k) \right\} = 0, \qquad J_k^*(x_f) = 0. \tag{26}$$

The minimizing control in the HJB equation determines the optimal feedback control $u_k^*(x)$, and

the minimum cumulative cost from the prescribed initial state is $J_k^*(x_0)$. A detailed derivation is provided in Supplementary Information section 2.1.

**Maximizing cumulative cooperation under a fixed budget.** We adopt the minority control principle, under which regulation rewards minority cooperators or punishes minority defectors. The control is therefore

$$u^*(x) = \begin{cases} u_{\mathrm{r}}^*(x), & x < 0.5, \\ u_{\mathrm{p}}^*(x), & x \geq 0.5. \end{cases} \tag{27}$$

Accordingly, the replicator dynamics under control can be written as

$$\dot{x} = \frac{\omega}{2} x(1-x)\left[\Phi(x) \pm u(x)\right] \equiv G(x,u). \tag{28}$$

The plus sign corresponds to reward control, with $u(x) = u_{\mathrm{r}}(x)$, whereas the minus sign corresponds to punishment control, with $u(x) = u_{\mathrm{p}}(x)$. This formulation is equivalent to the previous $\Lambda_k(x)$ notation with $\Lambda_{\mathrm{r}}(x) = 1$ and $\Lambda_{\mathrm{p}}(x) = -1$.

Under a fixed budget $\mathcal{B}$, the system may not be able to reach the desired state $x_f^{\mathrm{d}}$. In this case, the maximum attainable state is denoted by $\bar{x}(\mathcal{B})$. Therefore, the actual attained target state is $x_f^{\mathrm{a}} = \min\{x_f^{\mathrm{d}}, \bar{x}(\mathcal{B})\}$.

The process is divided into three phases. The first is the reaching phase, $t \in [t_0, t_r]$, during which the system is driven from the initial state $x_0$ to the attainable target state $x_f^{\mathrm{a}}$. The second is the maintenance phase, $t \in [t_r, t_m]$, during which the system is maintained at $x_f^{\mathrm{a}}$ by control. The third is the uncontrolled phase, $t \in [t_m, t_e]$, after the budget has been exhausted. Control is applied only in the reaching and maintenance phases. The cumulative cooperation is therefore the sum of cooperation over these three phases. The instantaneous cost remains $U(x(t))$, as defined in Eq. (7), and the cumulative cost is constrained by $\int_{t_0}^{t_m} U(x(t))\, dt = \mathcal{B}$.

Taken together, the optimization problem is to determine the desired state $x_f^{\mathrm{d}} \in (0,1)$ that maximizes cumulative cooperation $X$ from the initial state $x_0$ at time $t_0$ to the extinction time $t_e$,

subject to the fixed budget $\mathcal{B}$. This problem can be written as

$$\max_{x_f^{\mathrm{d}}} X(x_f^{\mathrm{d}}) = \int_{t_0}^{t_r} x(t)\,dt + \int_{t_r}^{t_m} x_f^{\mathrm{a}}\,dt + \int_{t_m}^{t_e} x(t \mid u = 0)\,dt,$$

$$\text{s.t.} \quad \begin{cases} \dot{x} = G(x, u^*), & t \in [t_0, t_r], \\ x(t_0) = x_0, \qquad x(t_r) = x_f^{\mathrm{a}}, & \\ x(t) = x_f^{\mathrm{a}}, & t \in [t_r, t_m], \\ \dot{x} = G(x, 0), & t \in [t_m, t_e], \\ \displaystyle\int_{t_0}^{t_m} U(x(t))\,dt = \mathcal{B}, & \\ x_f^{\mathrm{a}} = \min\left\{x_f^{\mathrm{d}}, \bar{x}(\mathcal{B})\right\}, & \\ x(t_e) = 10^{-10}. & \end{cases} \tag{29}$$

Detailed derivations of the costs, phase durations, and the three contributions to cumulative cooperation are provided in Supplementary Information Section 2.5.

# Data availability

All data generated or analysed during this study are included within the paper and its supplementary information files.

# References


[1] M. A. Nowak, Five rules for the evolution of cooperation. *Science* **314**, 1560–1563 (2006).

[2] M. A. Nowak, K. Sigmund, Tit for tat in heterogeneous populations. *Nature* **355**, 250–253 (1992).

[3] M. A. Nowak, K. Sigmund, A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature* **364**, 56–58 (1993).

[4] M. A. Nowak, K. Sigmund, Evolution of indirect reciprocity by image scoring. *Nature* **393**, 573–577 (1998).

[5] C. Wedekind, M. Milinski, Cooperation through image scoring in humans. *Science* **288**, 850–852 (2000).

[6] M. A. Nowak, K. Sigmund, Evolution of indirect reciprocity. *Nature* **437**, 1291–1298 (2005).

[7] K. Sigmund, C. Hauert, M. A. Nowak, Reward and punishment. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10757–10762 (2001).

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

[9] M. Alsobay, D. G. Rand, D. J. Watts, A. Almaatouq, Integrative experiments identify how punishment affects welfare in public goods games. *Science* **392**, eaeb5280 (2026).

[10] E. Fehr, S. Gächter, Altruistic punishment in humans. *Nature* **415**, 137–140 (2002).

[11] D. G. Rand, A. Dreber, T. Ellingsen, D. Fudenberg, M. A. Nowak, Positive interactions promote public cooperation. *Science* **325**, 1272–1275 (2009).

[12] X. Chen, T. Sasaki, Å. Brännström, U. Dieckmann, First carrot, then stick: how the adaptive hybridization of incentives promotes cooperation. *J. R. Soc. Interface* **12**, 20140935 (2015).

[13] S. Wang, M. Cao, X. Chen, Optimally combined incentive for cooperation among interacting agents in population games. *IEEE Trans. Autom. Control* **70**, 4562–4577 (2025).

[14] M. E. Renwick, R. D. Green, Do residential water demand side management policies measure up? An analysis of eight California water agencies. *J. Environ. Econ. Manag.* **40**, 37–55 (2000).

[15] S. M. Olmstead, R. N. Stavins, Comparing price and nonprice approaches to urban water conservation. *Water Resour. Res.* **45**, W04301 (2009).

[16] J. R. Beddington, D. J. Agnew, C. W. Clark, Current problems in the management of marine fisheries. *Science* **316**, 1713–1716 (2007).

[17] C. Costello, S. D. Gaines, J. Lynham, Can catch shares prevent fisheries collapse? *Science* **321**, 1678–1681 (2008).

[18] C. Mora, R. A. Myers, M. Coll, S. Libralato, T. J. Pitcher, R. U. Sumaila, D. Zeller, R. Watson, K. J. Gaston, B. Worm, Management effectiveness of the world's marine fisheries. *PLoS Biol.* **7**, e1000131 (2009).

[19] D. Nepstad, D. McGrath, C. Stickler, A. Alencar, A. Azevedo, B. Swette, T. Bezerra, M. DiGiano, J. Shimada, R. Seroa da Motta, E. Armijo, L. Castello, P. Brando, M. C. Hansen, M. McGrath-Horn, O. Carvalho, L. Hess, Slowing Amazon deforestation through public policy and interventions in beef and soy supply chains. *Science* **344**, 1118–1123 (2014).

[20] J. Assunção, C. Gandour, R. Rocha, DETER-ing deforestation in the Amazon: environmental monitoring and law enforcement. *Am. Econ. J. Appl. Econ.* **15**, 125–156 (2023).

[21] C. Nolte, A. Agrawal, K. M. Silvius, B. S. Soares-Filho, Governance regime and location influence avoided deforestation success of protected areas in the Brazilian Amazon. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 4956–4961 (2013).

[22] M. A. Janssen, R. Holahan, A. Lee, E. Ostrom, Lab experiments for the study of social-ecological systems. *Science* **328**, 613–617 (2010).

[23] R. Boyd, H. Gintis, S. Bowles, Coordinated punishment of defectors sustains cooperation and can proliferate when rare. *Science* **328**, 617–620 (2010).

[24] D. Baldassarri, G. Grossman, Centralized sanctioning and legitimate authority promote cooperation in humans. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 11023–11027 (2011).

[25] D. P. McCarthy, P. F. Donald, J. P. W. Scharlemann, G. M. Buchanan, A. Balmford, J. M. H. Green, L. A. Bennun, N. D. Burgess, L. D. C. Fishpool, S. T. Garnett, D. L. Leonard, R. F. Maloney, P. Morling, H. M. Schaefer, A. Symes, D. A. Wiedenfeld, S. H. M. Butchart, Financial costs of meeting global biodiversity conservation targets: current spending and unmet needs. *Science* **338**, 946–949 (2012).

[26] A. Waldron, A. O. Mooers, D. C. Miller, N. Nibbelink, D. Redding, T. S. Kuhn, J. T. Roberts, J. L. Gittleman, Targeting global conservation funding to limit immediate biodiversity declines. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 12144–12148 (2013).

[27] D. A. Gill, M. B. Mascia, G. N. Ahmadia, L. Glew, S. E. Lester, M. Barnes, I. Craigie, E. S. Darling, C. M. Free, J. Geldmann, S. Holst, O. P. Jensen, A. T. White, X. Basurto, L. Coad, R. D. Gates, G. Guannel, P. J. Mumby, H. Thomas, S. Whitmee, S. Woodley, H. E. Fox, Capacity shortfalls hinder the performance of marine protected areas globally. *Nature* **543**, 665–669 (2017).

[28] N. Masuda, Evolution via imitation among like-minded individuals. *J. Theor. Biol.* **349**, 100–108 (2014).

[29] G. Yan, J. Ren, Y.-C. Lai, C.-H. Lai, B. Li, Controlling complex networks: how much energy is needed? *Phys. Rev. Lett.* **108**, 218703 (2012).

[30] G. Yan, G. Tsekenis, B. Barzel, J.-J. Slotine, Y.-Y. Liu, A.-L. Barabási, Spectrum of controlling and observing complex networks. *Nat. Phys.* **11**, 779–786 (2015).

[31] I. Klickstein, A. Shirin, F. Sorrentino, Energy scaling of targeted optimal control of complex networks. *Nat. Commun.* **8**, 15145 (2017).

[32] Z. Wang, M. S. Sarandy, L. Wu, Almost exact state transfer in a spin chain via pulse control. *Phys. Rev. A* **102**, 022601 (2020).

[33] J. Jing, L.-A. Wu, M. Byrd, J. Q. You, T. Yu, Z.-M. Wang, Nonperturbative leakage elimination operators and control of a three-level system. *Phys. Rev. Lett.* **114**, 190502 (2015).

[34] K. Khodjasteh, J. Sastrawan, D. Hayes, T. J. Green, M. J. Biercuk, L. Viola, Designing a practical high-fidelity long-time quantum memory. *Nat. Commun.* **4**, 2045 (2013).

[35] D. Rustagi, S. Engel, M. Kosfeld, Conditional cooperation and costly monitoring explain success in forest commons management. *Science* **330**, 961–965 (2010).

[36] A. Chhatre, A. Agrawal, Forest commons and local enforcement. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 13286–13291 (2008).

[37] H. Luo, W. Peng, A. Fawcett, J. F. Green, G. Iyer, J. Meckling, J. Nahm, D. G. Victor, Modelling the impacts of policy sequencing on energy decarbonization. *Nat. Clim. Chang.* **16**, 43–51 (2026).

[38] B. Noll, D. Graff, T. S. Schmidt, A. Patt, C. Bauer, C. Agutu, P. K. Asiimwe, O. E. Diemuodeke, S. Krygsman, I. Nienkerke, T. Tröndle, C. Moretti, Battery-electric passenger vehicles will be cost-effective across Africa well before 2040. *Nat. Energy* **11**, 284–298 (2026).

[39] K. G. Austin, A. Favero, N. Forsell, B. L. Sohngen, C. M. Wade, S. B. Ohrel, S. Ragnauth, Targeting climate finance for global forests. *Nat. Commun.* **16**, 6443 (2025).

[40] S. Wang, X. Chen, A. Szolnoki, Exploring optimal institutional incentives for public cooperation. *Commun. Nonlinear Sci. Numer. Simul.* **79**, 104914 (2019).

[41] S. Wang, X. Chen, Z. Xiao, A. Szolnoki, V. V. Vasconcelos, Optimization of institutional incentives for cooperation in structured populations. *J. R. Soc. Interface* **20**, 20220653 (2023).

[42] Ö. Gürerk, B. Irlenbusch, B. Rockenbach, The competitive advantage of sanctioning institutions. *Science* **312**, 108–111 (2006).

[43] S. M. Kissler, C. Tedijanto, E. Goldstein, Y. H. Grad, M. Lipsitch, Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science* **368**, 860–868 (2020).

[44] S. Flaxman, S. Mishra, A. Gandy, H. J. T. Unwin, T. A. Mellan, H. Coupland, C. Whittaker, H. Zhu, T. Berah, J. W. Eaton, M. Monod, Imperial College COVID-19 Response Team, A. C. Ghani, C. A. Donnelly, S. Riley, M. A. C. Vollmer, N. M. Ferguson, L. C. Okell, S. Bhatt, Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* **584**, 257–261 (2020).

[45] S. Hsiang, D. Allen, S. Annan-Phan, K. Bell, I. Bolliger, T. Chong, H. Druckenmiller, L. Y. Huang, A. Hultgren, E. Krasovich, P. Lau, J. Lee, E. Rolf, J. Tseng, T. Wu, The effect of large-scale anti-contagion policies on the COVID-19 pandemic. *Nature* **584**, 262–267 (2020).

[46] J. M. Brauner, S. Mindermann, M. Sharma, D. Johnston, J. Salvatier, T. Gavenčiak, A. B. Stephenson, G. Leech, G. Altman, V. Mikulik, A. J. Norman, J. T. Monrad, T. Besiroglu, H. Ge, M. A. Hartwick, Y. W. Teh, L. Chindelevitch, Y. Gal, J. Kulveit, Inferring the effectiveness of government interventions against COVID-19. *Science* **371**, eabd9338 (2021).

[47] A. Rahmani, M. Ji, M. Mesbahi, M. Egerstedt, Controllability of multi-agent systems from a graph-theoretic perspective. *SIAM J. Control Optim.* **48**, 162–186 (2009).

[48] Y. Liu, J.-J. Slotine, A.-L. Barabási, Controllability of complex networks. *Nature* **473**, 167–173 (2011).

[49] J. Ruths, D. Ruths, Control profiles of complex networks. *Science* **343**, 1373–1376 (2014).

[50] J. Gao, Y. Liu, R. M. D'Souza, A.-L. Barabasi, Target control of complex networks. *Nat. Commun.* **5**, 5415 (2014).

[51] G. Baggio, D. S. Bassett, F. Pasqualetti, Data-driven control of complex networks. *Nat. Commun.* **12**, 1429 (2021).

[52] I. Rajapakse, M. Groudine, M. Mesbahi, Dynamics and control of state-dependent networks for probing genomic organization. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 17257–17262 (2011).

[53] S. P. Cornelius, W. L. Kath, A. E. Motter, Realistic control of network dynamics. *Nat. Commun.* **4**, 1942 (2013).

[54] G. Yan, P. E. Vertes, E. K. Towlson, Y. L. Chew, D. S. Walker, W. R. Schafer, A.-L. Barabasi, Network control principles predict neuron function in the Caenorhabditis elegans connectome. *Nature* **550**, 519–523 (2017).

[55] Q. Su, B. Allen, J. B. Plotkin, Evolution of cooperation with asymmetric social interactions. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2113468118 (2022).

[56] Q. Su, A. McAvoy, Y. Mori, J. B. Plotkin, Evolution of prosocial behaviours in multilayer populations. *Nat. Hum. Behav.* **6**, 338–348 (2022).

[57] U. Alvarez-Rodriguez, F. Battiston, G. F. de Arruda, Y. Moreno, M. Perc, V. Latora, Evolutionary dynamics of higher-order interactions in social networks. *Nat. Hum. Behav.* **5**, 586–595 (2021).

[58] A. Civilini, O. Sadekar, F. Battiston, J. Gómez-Gardeñes, V. Latora, Explosive cooperation in social dilemmas on higher-order networks. *Phys. Rev. Lett.* **132**, 167401 (2024).

[59] Y. Meng, S. P. Cornelius, Y. Liu, A. Li, Dynamics of collective cooperation under personalised strategy updates. *Nat. Commun.* **15**, 3125 (2024).

[60] Q. Su, A. J. Stewart, Evolutionary dynamics of behavioral motivations for cooperation. *Nat. Commun.* **16**, 4023 (2025).

# Acknowledgements

Q.S. acknowledges support from the National Natural Science Foundation of China (No. 62473252, 72642012).

# Author contributions

S.S., Z.G. and Q.S. conceived and designed the research. S.S., J.W. and Q.S. contributed to the methodology development. S.S.derived analytical results and performed numerical calculations. S.S. and Q.S. analyzed the data. S.S., Z.G., J.W. and Q.S. contributed to the interpretation of the results. S.S. wrote the main text and the Supplementary Information. Q.S. reviewed and edited the main text.

# Competing interests

The authors declare no competing interests.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementaryinformation.pdf