Supplementary information
Minority control maximizes cumulative cooperation

Sha Song Ziyou Gao Jianjun Wu Qi Su

Contents

1 Model **2**

2 Results **8**

3 Extensions **19**

4 Supplementary Figures **20**

1 Model

1.1 Expressions for Λ_k and Γ_k under six control schemes

We consider a finite well-mixed population of size N, containing $n_{\rm C}$ cooperators (C) and $N-n_{\rm C}$ defectors (D). Each individual interacts with all other $N-1$ individuals. For the payoff matrix

$$\begin{array}{c} \\ \mathrm{C} \\ \mathrm{D} \end{array} \begin{array}{c} \begin{array}{cc} \mathrm{C} & \mathrm{D} \end{array} \\ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \end{array}, \tag{S1}$$

the average payoff of a cooperator is

$$\pi_{\rm C}^0 = \frac{(n_{\rm C}-1)a + (N-n_{\rm C})b}{N-1}, \tag{S2}$$

and the average payoff of a defector is

$$\pi_{\rm D}^0 = \frac{n_{\rm C}c + (N-n_{\rm C}-1)d}{N-1}. \tag{S3}$$

The payoff difference between cooperation and defection is defined as $\Phi = \pi_{\rm C}^0 - \pi_{\rm D}^0$.

We consider six control schemes: C-based, D-based, CC-based, CD-based, DC-based, and DD-based control. Under control, an individual i may receive an additional payoff, either positive or negative, and its payoff is therefore modified. For each scheme, we derive the modified average payoffs and specify the corresponding instantaneous cost function for implementing control.

(1) C-based control

In C-based control, individual i receives an additional payoff $u_{\rm r}$ if it cooperates and receives no additional payoff if it defects. The average payoff of individual i, conditional on i being a cooperator, is

$$\pi_{\rm C} = \frac{(a+u_{\rm r})\cdot(n_{\rm C}-1) + (b+u_{\rm r})\cdot(N-n_{\rm C})}{N-1} = \pi_{\rm C}^0 + u_{\rm r},$$

whereas its average payoff, conditional on i being a defector, is $\pi_{\rm D} = \pi_{\rm D}^0$.

The average payoff difference between the two strategies is

$$\Phi_{\rm r} = \pi_{\rm C}^0 - \pi_{\rm D}^0 + u_{\rm r} \triangleq \Phi + \Lambda_{\rm r} u_{\rm r}. \tag{S4}$$

A fraction $n_{\rm C}/N$ of the population consists of cooperators, each of whom receives an additional

payoff in every interaction. Thus, the instantaneous cost is

$$U_{\mathrm{r}} = \left(\frac{n_{\mathrm{C}}}{N} \cdot u_{\mathrm{r}}\right)^2 \triangleq \left(\Gamma_{\mathrm{r}} u_{\mathrm{r}}\right)^2 . \tag{S5}$$

(2) D-based control

In D-based control, individual i receives an additional payoff u_{p} if it defects and receives no additional payoff if it cooperates. The average payoff of individual i, conditional on i being a cooperator, is $\pi_{\mathrm{C}} = \pi_{\mathrm{C}}^0$. Its average payoff, conditional on i being a defector, is

$$\pi_{\mathrm{D}} = \frac{(c + u_{\mathrm{p}}) \cdot n_{\mathrm{C}} + (d + u_{\mathrm{p}}) \cdot (N - n_{\mathrm{C}} - 1)}{N - 1} = \pi_{\mathrm{D}}^0 + u_{\mathrm{p}}.$$

The average payoff difference between the two strategies is

$$\Phi_{\mathrm{p}} = \pi_{\mathrm{C}}^0 - \pi_{\mathrm{D}}^0 - u_{\mathrm{p}} \triangleq \Phi + \Lambda_p u_{\mathrm{p}}, \tag{S6}$$

A fraction $(N - n_{\mathrm{C}})/N$ of the population consists of defectors, each of whom receives an additional payoff in every interaction. Thus, the instantaneous cost is

$$U_{\mathrm{p}} = \left(\frac{N - n_{\mathrm{C}}}{N} \cdot u_{\mathrm{p}}\right)^2 \triangleq \left(\Gamma_{\mathrm{p}} u_{\mathrm{p}}\right)^2 . \tag{S7}$$

(3) CC-based control

In CC-based control, individual i receives u_{CC} only when both individual i and its opponent cooperate. The average payoff of individual i, conditional on i being a cooperator, is

$$\pi_{\mathrm{C}} = \frac{(a + u_{\mathrm{CC}}) \cdot (n_{\mathrm{C}} - 1) + b \cdot (N - n_{\mathrm{C}})}{N - 1} = \pi_{\mathrm{C}}^0 + \frac{n_{\mathrm{C}} - 1}{N - 1} u_{\mathrm{CC}},$$

whereas its average payoff, conditional on i being a defector, is $\pi_{\mathrm{D}} = \pi_{\mathrm{D}}^0$.

The average payoff difference between the two strategies is

$$\Phi_{\mathrm{CC}} = \pi_{\mathrm{C}}^0 - \pi_{\mathrm{D}}^0 + \frac{n_{\mathrm{C}} - 1}{N - 1} u_{\mathrm{CC}} \triangleq \Phi + \Lambda_{\mathrm{CC}} u_{\mathrm{CC}}. \tag{S8}$$

A fraction n_{C}/N of the population consists of cooperators, each of whom receives an additional payoff in a fraction $(n_{\mathrm{C}} - 1)/(N - 1)$ of its interactions. Thus, the instantaneous cost is

$$U_{\mathrm{CC}} = \left(\frac{n_{\mathrm{C}} \cdot (n_{\mathrm{C}} - 1)}{N(N - 1)} \cdot u_{\mathrm{CC}}\right)^2 \triangleq \left(\Gamma_{\mathrm{CC}} u_{\mathrm{CC}}\right)^2 . \tag{S9}$$

(4) CD-based control

In CD-based control, individual i receives u_{CD} when it cooperates and its opponent defects. The average payoff of individual i, conditional on i being a cooperator, is

$$\pi_{\mathrm{C}} = \frac{a \cdot (n_{\mathrm{C}} - 1) + (b + u_{\mathrm{CD}}) \cdot (N - n_{\mathrm{C}})}{N - 1} = \pi_{\mathrm{C}}^{0} + \frac{N - n_{\mathrm{C}}}{N - 1} u_{\mathrm{CD}},$$

whereas its average payoff, conditional on i being a defector, is $\pi_{\mathrm{D}} = \pi_{\mathrm{D}}^{0}$.

The average payoff difference between the two strategies is

$$\Phi_{\mathrm{CD}} = \pi_{\mathrm{C}}^{0} - \pi_{\mathrm{D}}^{0} + \frac{N - n_{\mathrm{C}}}{N - 1} u_{\mathrm{CD}} \triangleq \Phi + \Lambda_{\mathrm{CD}} u_{\mathrm{CD}}, \tag{S10}$$

A fraction n_{C}/N of the population consists of cooperators, each of whom receives an additional payoff in a fraction $(N - n_{\mathrm{C}})/(N - 1)$ of its interactions. Thus, the instantaneous cost is

$$U_{\mathrm{CD}} = \left(\frac{n_{\mathrm{C}} \cdot (N - n_{\mathrm{C}})}{N(N - 1)} \cdot u_{\mathrm{CD}} \right)^{2} \triangleq \left(\Gamma_{\mathrm{CD}} u_{\mathrm{CD}} \right)^{2}. \tag{S11}$$

(5) DC-based control

In DC-based control, individual i receives u_{DC} when it defects and its opponent cooperates. The average payoff of individual i, conditional on i being a cooperator, is $\pi_{\mathrm{C}} = \pi_{\mathrm{C}}^{0}$. Its average payoff, conditional on i being a defector, is

$$\pi_{\mathrm{D}} = \frac{(c + u_{\mathrm{DC}}) \cdot n_{\mathrm{C}} + d \cdot (N - n_{\mathrm{C}} - 1)}{N - 1} = \pi_{\mathrm{D}}^{0} + \frac{n_{\mathrm{C}}}{N - 1} u_{\mathrm{DC}}.$$

The average payoff difference between the two strategies is

$$\Phi_{\mathrm{DC}} = \pi_{\mathrm{C}}^{0} - \pi_{\mathrm{D}}^{0} - \frac{n_{\mathrm{C}}}{N - 1} u_{\mathrm{DC}} \triangleq \Phi + \Lambda_{\mathrm{DC}} u_{\mathrm{DC}}, \tag{S12}$$

A fraction $(N - n_{\mathrm{C}})/N$ of the population consists of defectors, each of whom receives an additional payoff in a fraction $n_{\mathrm{C}}/(N - 1)$ of its interactions. Thus, the instantaneous cost is

$$U_{\mathrm{DC}} = \left(\frac{(N - n_{\mathrm{C}}) \cdot n_{\mathrm{C}}}{N(N - 1)} \cdot u_{\mathrm{DC}} \right)^{2} \triangleq \left(\Gamma_{\mathrm{DC}} u_{\mathrm{DC}} \right)^{2}. \tag{S13}$$

(6) DD-based control

In DD-based control, individual i receives u_{DD} when both individual i and its opponent defect. The average payoff of individual i, conditional on i being a cooperator, is $\pi_{\mathrm{C}} = \pi_{\mathrm{C}}^{0}$. Its average payoff, conditional on i being a defector, is

$$\pi_{\mathrm{D}} = \frac{c \cdot n_{\mathrm{C}} + (d + u_{\mathrm{DD}}) \cdot (N - n_{\mathrm{C}} - 1)}{N - 1} = \pi_{\mathrm{D}}^{0} + \frac{N - n_{\mathrm{C}} - 1}{N - 1} u_{\mathrm{DD}}.$$

The average payoff difference between the two strategies is

$$\Phi_{\mathrm{DD}} = \pi_{\mathrm{C}}^{0} - \pi_{\mathrm{D}}^{0} - \frac{N - n_{\mathrm{C}} - 1}{N-1} u_{\mathrm{DD}} \triangleq \Phi + \Lambda_{\mathrm{DD}} u_{\mathrm{DD}}, \tag{S14}$$

A fraction $(N - n_{\mathrm{C}})/N$ of the population consists of defectors, each of whom receives an additional payoff in a fraction $(N - n_{\mathrm{C}} - 1)/(N - 1)$ of its interactions. Thus, the instantaneous cost is

$$U_{\mathrm{DD}} = \left(\frac{(N - n_{\mathrm{C}}) \cdot (N - n_{\mathrm{C}} - 1)}{N(N-1)} \cdot u_{\mathrm{DD}} \right)^{2} \triangleq \left(\Gamma_{\mathrm{DD}} u_{\mathrm{DD}} \right)^{2}. \tag{S15}$$

Let $x = n_{\mathrm{C}}/N$. The expressions for Λ_k and Γ_k under six control schemes are summarized in Supplementary Table 1.

Supplementary Table 1: Expressions for Λ_k and Γ_k under six control schemes.

Control scheme k	$\Lambda_k(x)$		$\Gamma_k(x)$	
	Finite N	$N \to \infty$	Finite N	$N \to \infty$
r	1	1	x	x
p	-1	-1	$1 - x$	$1 - x$
CC	$\frac{Nx - 1}{N - 1}$	x	$\frac{x(Nx - 1)}{N - 1}$	x^2
CD	$\frac{N(1 - x)}{N - 1}$	$1 - x$	$\frac{Nx(1 - x)}{N - 1}$	$x(1 - x)$
DC	$-\frac{Nx}{N - 1}$	$-x$	$\frac{Nx(1 - x)}{N - 1}$	$x(1 - x)$
DD	$-\frac{N - Nx - 1}{N - 1}$	$-(1 - x)$	$\frac{(1 - x)\big(N - Nx - 1\big)}{N - 1}$	$(1 - x)^2$

1.2 Replicator dynamics

Using the average payoffs of cooperators, π_{C}, and defectors, π_{D}, defined above, we derive the replicator dynamics with and without control.

In one elementary update, the number of cooperators can increase from n_{C} to $n_{\mathrm{C}} + 1$, decrease from n_{C} to $n_{\mathrm{C}} - 1$, or remain unchanged. We denote the corresponding transition probabilities by $T^{+}(n_{\mathrm{C}})$, $T^{-}(n_{\mathrm{C}})$ and $1 - T^{+}(n_{\mathrm{C}}) - T^{-}(n_{\mathrm{C}})$, respectively, where $T^{+}(n_{\mathrm{C}})$ corresponds to

a defector imitating a cooperator and thus an increase in the number of cooperators, whereas $T^-(n_{\mathrm{C}})$ corresponds to a cooperator imitating a defector and thus a decrease in the number of cooperators. These transitions are generated by 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 π_i, adopts the strategy of individual j, with payoff π_j, with probability

$$P^{i\to j} = \frac{1}{1+\exp(-\omega(\pi_j-\pi_i))}, \tag{S16}$$

where ω measures the intensity of selection. Thus, the transition probabilities are

$$T^+(n_{\mathrm{C}}) = \frac{n_{\mathrm{C}}}{N}\frac{N-n_{\mathrm{C}}}{N}P^{\mathrm{D}\to\mathrm{C}}, \tag{S17}$$

and

$$T^-(n_{\mathrm{C}}) = \frac{n_{\mathrm{C}}}{N}\frac{N-n_{\mathrm{C}}}{N}P^{\mathrm{C}\to\mathrm{D}}. \tag{S18}$$

If $q^\tau(n_{\mathrm{C}})$ is the probability that the system contains n_{C} cooperators after τ elementary updates, the stochastic process satisfies the master equation

$$\begin{aligned} q^{\tau+1}(n_{\mathrm{C}}) - q^\tau(n_{\mathrm{C}}) = {} & q^\tau(n_{\mathrm{C}}-1)T^+(n_{\mathrm{C}}-1) + q^\tau(n_{\mathrm{C}}+1)T^-(n_{\mathrm{C}}+1) \\ & - q^\tau(n_{\mathrm{C}})\left[T^+(n_{\mathrm{C}}) + T^-(n_{\mathrm{C}})\right]. \end{aligned} \tag{S19}$$

To obtain a continuous description, we express the dynamics in terms of the cooperation frequency $x = n_{\mathrm{C}}/N$, and rescale time as $t = \tau/N$. We also define the probability density $\rho(x,t) = Nq^\tau(n_{\mathrm{C}})$. The master equation can then be rewritten as

$$\begin{aligned} \rho\left(x, t+N^{-1}\right) - \rho(x,t) = {} & \rho\left(x-N^{-1}, t\right)T^+\left(x-N^{-1}\right) + \rho\left(x+N^{-1}, t\right)T^-\left(x+N^{-1}\right) \\ & - \rho(x,t)\left[T^+(x) + T^-(x)\right]. \end{aligned} \tag{S20}$$

For $N \gg 1$, expanding Eq. (S20) to second order in N^{-1} gives the Fokker–Planck equation

$$\frac{\partial\rho(x,t)}{\partial t} = -\frac{\partial}{\partial x}\left[A(x)\rho(x,t)\right] + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left[B(x)\rho(x,t)\right], \tag{S21}$$

where the drift term is $A(x) = T^+(x) - T^-(x)$, and the diffusion coefficient is $B(x) = \left[T^+(x) + T^-(x)\right]/N$. Equivalently, the stochastic dynamics can be written as the Langevin equation

$$dx = A(x)\,dt + \sqrt{B(x)}\,dW_t, \tag{S22}$$

where W_t is a standard Wiener process. Since $B(x) = O(N^{-1})$, the demographic noise vanishes in the deterministic limit $N \to \infty$, and the stochastic dynamics reduce to an ordinary differential equation.

In the deterministic limit, the state dynamics are governed by the drift term,

$$\dot{x} = A(x) = T^{+}(x) - T^{-}(x).$$

Using the Fermi imitation probabilities in Eqs. (S16)–(S18), we obtain

$$\dot{x} = x(1-x)\tanh\left[\frac{\omega}{2}\left(\pi_{\mathrm{C}} - \pi_{\mathrm{D}}\right)\right]. \tag{S23}$$

Equation (S23) gives the exact deterministic dynamics associated with the Fermi imitation process. Under weak selection, such that $|\omega(\pi_C - \pi_D)| \ll 1$, we use $\tanh(z) \simeq z$. In the absence of control,

$$\pi_C - \pi_D = \Phi(x) = K_0 + K_1 x, \tag{S24}$$

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

$$\dot{x} = \frac{\omega}{2} x(1-x)\Phi(x). \tag{S25}$$

Under control scheme k, the payoff difference becomes

$$\pi_C - \pi_D = \Phi(x) + \Lambda_k(x) u_k. \tag{S26}$$

Accordingly, the replicator equation with control is

$$\dot{x} = \frac{\omega}{2} x(1-x)\left[\Phi(x) + \Lambda_k(x) u_k\right]. \tag{S27}$$

2 Results

2.1 Solving the optimal control problem by dynamic programming

For the six control schemes $k \in \{\mathrm{r}, \mathrm{p}, \mathrm{CC}, \mathrm{CD}, \mathrm{DC}, \mathrm{DD}\}$, the minimum-cost control problem with free terminal time can be written in a unified form as

$$\min_{u_k} J_k = \int_{t_0}^{t_f} U_k(x(t))\, dt,$$
$$\text{s.t.} \begin{cases} \dot{x} = \dfrac{\omega}{2} x(1-x)\left[\Phi(x) + \Lambda_k(x) u_k\right] \triangleq F_k(x, u_k), \\ x(t_0) = x_0, \\ x(t_f) = x_f. \end{cases} \tag{S28}$$

The instantaneous cost is

$$U_k(x(t)) = \left[\Gamma_k(x) u_k\right]^2. \tag{S29}$$

For notational simplicity, we omit the explicit dependence of Λ_k and Γ_k on x in the following derivation. The optimal control is obtained using dynamic programming.

Let the value function be

$$J_k^*(x) = \min_{u_k} \int_t^{t_f} U_k(x(s))\, ds. \tag{S30}$$

Since the system is autonomous and the terminal time is free, the value function is time-independent. The stationary Hamilton–Jacobi–Bellman (HJB) equation is therefore

$$\min_{u_k} \left\{ U_k(x) + \frac{\partial J_k^*}{\partial x} \cdot F_k(x, u_k) \right\} = 0. \tag{S31}$$

Substituting Eqs. (S28) and (S29) into Eq. (S31) gives

$$\min_{u_k} \left\{ \left(\Gamma_k u_k\right)^2 + \frac{\partial J_k^*}{\partial x} \cdot \frac{\omega}{2} x(1-x)\left[\Phi(x) + \Lambda_k u_k\right] \right\} = 0. \tag{S32}$$

The first-order condition for the minimizing control is

$$\frac{\partial}{\partial u_k} \left\{ \left(\Gamma_k u_k\right)^2 + \frac{\partial J_k^*}{\partial x} \cdot \frac{\omega}{2} x(1-x)\left[\Phi(x) + \Lambda_k u_k\right] \right\} = 0. \tag{S33}$$

This yields

$$u_k^* = -\frac{\omega x(1-x)\Lambda_k}{4\Gamma_k^2} \cdot \frac{\partial J_k^*}{\partial x}. \tag{S34}$$

Substituting Eq. (S34) into the left-hand side of Eq. (S32), we obtain

$$\begin{aligned}
&(\Gamma_k u_k^*)^2 + \frac{\partial J_k^*}{\partial x} \cdot \frac{\omega}{2} x(1-x)\left[\Phi(x) + \Lambda_k u_k^*\right] \\
&= \Gamma_k^2 \left[-\frac{\omega x(1-x)\Lambda_k}{4\Gamma_k^2} \cdot \frac{\partial J_k^*}{\partial x}\right]^2 + \frac{\partial J_k^*}{\partial x} \cdot \frac{\omega}{2} x(1-x)\left[\Phi(x) - \frac{\omega x(1-x)\Lambda_k^2}{4\Gamma_k^2} \cdot \frac{\partial J_k^*}{\partial x}\right] \\
&= \frac{\omega}{2} x(1-x)\frac{\partial J_k^*}{\partial x}\left[\Phi(x) - \frac{\omega x(1-x)\Lambda_k^2}{8\Gamma_k^2}\frac{\partial J_k^*}{\partial x}\right].
\end{aligned} \tag{S35}$$

Equating this expression to zero gives

$$\frac{\partial J_k^*}{\partial x} = 0, \qquad \frac{\partial J_k^*}{\partial x} = \frac{8\Gamma_k^2 \Phi(x)}{\omega x(1-x)\Lambda_k^2}. \tag{S36}$$

Substituting these two possible value gradients into Eq. (S34) gives two solutions:

$$u_{k_1}^*(x) = 0, \qquad u_{k_2}^*(x) = -\frac{2\Phi(x)}{\Lambda_k(x)}. \tag{S37}$$

The first solution corresponds to following the natural evolutionary dynamics without active control. The second solution reverses the sign of the payoff-difference term in the controlled dynamics, because

$$\Phi(x) + \Lambda_k(x) u_k^*(x) = -\Phi(x). \tag{S38}$$

Therefore, the controlled vector field under the second solution is

$$\dot{x} = -\frac{\omega}{2} x(1-x)\Phi(x). \tag{S39}$$

The optimal control is selected from these two solutions according to the direction required to reach the prescribed target state. If the natural drift already moves the system toward the target, the zero-control branch $u_k^* = 0$ is optimal. If the natural drift moves the system away from the target, active control is required, and the optimal control is $u_k^* = -2\Phi(x)/\Lambda_k(x)$.

2.2 State dynamics and terminal time under optimal control

The HJB analysis gives two solutions, $u_{k_1}^*(x) = 0$ and $u_{k_2}^*(x) = -2\Phi(x)/\Lambda_k(x)$. The first follows the natural evolutionary dynamics, whereas the second reverses the sign of the payoff-difference term. These two cases can be written in a unified form as

$$\dot{x} = \sigma \frac{\omega}{2} x(1-x)\Phi(x), \tag{S40}$$

where

$$\sigma = \begin{cases} +1, & u_k^*(x) = 0, \\ -1, & u_k^*(x) = -2\Phi(x)/\Lambda_k(x). \end{cases} \tag{S41}$$

Using

$$\Phi(x) = K_0 + K_1 x, \qquad K_0 = b - d, \qquad K_1 = a - b - c + d,$$

Eq. (S40) gives

$$\frac{dx}{x(1-x)(K_1 x + K_0)} = \sigma \frac{\omega}{2}\, dt. \tag{S42}$$

For $K_0 \neq 0$ and $K_0 + K_1 \neq 0$, the partial-fraction decomposition is

$$\frac{1}{x(1-x)(K_1 x + K_0)} = \frac{P}{x} + \frac{Q}{1-x} + \frac{R}{K_1 x + K_0}, \tag{S43}$$

with

$$P = \frac{1}{K_0}, \qquad Q = \frac{1}{K_0 + K_1}, \qquad R = -\frac{K_1^2}{K_0(K_0 + K_1)}. \tag{S44}$$

Thus,

$$\begin{aligned} \int \frac{dx}{x(1-x)(K_1 x + K_0)} = {} & \frac{1}{K_0} \ln x - \frac{1}{K_0 + K_1} \ln(1-x) \\ & - \frac{K_1}{K_0(K_0 + K_1)} \ln |K_1 x + K_0| + C. \end{aligned} \tag{S45}$$

Define

$$\Psi(x) = \frac{1}{K_0} \ln x - \frac{1}{K_0 + K_1} \ln(1-x) - \frac{K_1}{K_0(K_0 + K_1)} \ln |K_1 x + K_0|. \tag{S46}$$

Then Eq. (S42) gives the implicit solution

$$\Psi(x) = \sigma \frac{\omega}{2} t + C. \tag{S47}$$

Applying the boundary conditions $x(t_0) = x_0$ and $x(t_f) = x_f$, we obtain

$$\Psi(x_f) - \Psi(x_0) = \sigma \frac{\omega}{2}(t_f - t_0). \tag{S48}$$

Therefore, the terminal time is

$$t_f - t_0 = \frac{2}{\sigma\omega} \left[\Psi(x_f) - \Psi(x_0) \right]. \tag{S49}$$

Equivalently, for the zero-control branch,

$$t_f - t_0 = \frac{2}{\omega} \left[\Psi(x_f) - \Psi(x_0) \right], \qquad u_k^*(x) = 0, \tag{S50}$$

whereas for the active-control branch,

$$t_f - t_0 = \frac{2}{\omega}\left[\Psi(x_0) - \Psi(x_f)\right], \qquad u_k^*(x) = -\frac{2\Phi(x)}{\Lambda_k(x)}. \tag{S51}$$

If $t_0 = 0$, Eq. (S49) gives

$$t_f = \frac{2}{\sigma\omega}\left[\frac{1}{K_0}\ln\left(\frac{x_f}{x_0}\right) - \frac{1}{K_0+K_1}\ln\left(\frac{1-x_f}{1-x_0}\right) - \frac{K_1}{K_0(K_0+K_1)}\ln\left|\frac{K_1x_f+K_0}{K_1x_0+K_0}\right|\right]. \tag{S52}$$

2.3 Instantaneous cost under optimal control

The two optimal controls have been obtained above by dynamic programming (Eq. (S37)). According to Eq. (S29), the instantaneous cost under the optimal control, $U_k^*(x)$, can then be derived. When $u_{k_1}^* = 0$, we have $U_k^*(x) = 0$. When $u_{k_2}^*(x) = -2\Phi(x)/\Lambda_k(x)$, the instantaneous cost is

$$\begin{aligned} U_k^*(x) = [\Gamma_k(x)u_k]^2 &= \left[\Gamma_k(x)\cdot\left(-\frac{2\Phi(x)}{\Lambda_k(x)}\right)\right]^2 \\ &= \left[-2\Phi(x)\frac{\Gamma_k(x)}{\Lambda_k(x)}\right]^2. \end{aligned} \tag{S53}$$

Using the expressions in Supplementary Table 1, the three schemes targeting cooperators satisfy

$$\frac{\Gamma_k(x)}{\Lambda_k(x)} = x, \qquad k \in \{\mathrm{r}, \mathrm{CC}, \mathrm{CD}\}, \tag{S54}$$

whereas the three schemes targeting defectors satisfy

$$\frac{\Gamma_k(x)}{\Lambda_k(x)} = -(1-x), \qquad k \in \{\mathrm{p}, \mathrm{DC}, \mathrm{DD}\}. \tag{S55}$$

Thus, the six schemes can be grouped into two classes,

$$U_k^*(x) = \begin{cases} \left[-2\Phi(x)x\right]^2, \ k \in \{\mathrm{r}, \mathrm{CC}, \mathrm{CD}\}, \\ \left[2\Phi(x)(1-x)\right]^2, \ k \in \{\mathrm{p}, \mathrm{DC}, \mathrm{DD}\}. \end{cases} \tag{S56}$$

It is therefore sufficient to compare two strategy-based control schemes, namely C-based control and D-based control. The cost difference between the two control schemes is

$$U_{\mathrm{r}}^*(x) - U_{\mathrm{p}}^*(x) = 4\Phi^2(x)(2x-1). \tag{S57}$$

Because $\Phi^2(x) \geq 0$, the sign of this difference is determined only by whether x is below or above $1/2$. Therefore, C-based control is less costly when $x < 1/2$, whereas D-based control is

less costly when $x > 1/2$.

2.4 Local deterministic pulse control around an interior equilibrium

This appendix provides a detailed derivation of the local deterministic pulse control used to overcome the non-crossability of an interior equilibrium.

Since $\Phi(x) = K_0 + K_1 x$, an interior equilibrium x^* satisfies $x^* = -K_0/K_1$. Therefore, near x^*, it follows that $\Phi(x) = K_1(x - x^*)$. For sufficiently small ϵ, we use the local approximations

$$x(1-x) \approx x^*(1-x^*), \tag{S58}$$

and

$$\Lambda_k(x) \approx \Lambda_k(x^*), \qquad \Gamma_k(x) \approx \Gamma_k(x^*). \tag{S59}$$

Thus, in a small neighbourhood of x^*, the uncontrolled or reversed local dynamics take the linearized form

$$\dot{x} \approx \pm\frac{\omega}{2} x^*(1-x^*)K_1(x-x^*). \tag{S60}$$

Let $z = x - x^*$, then $\dot{z} = \lambda z$, where $|\lambda| = (\omega/2)x^*(1-x^*)|K_1|$. The natural local time scale induced by the linearized dynamics is therefore

$$\Delta t = \frac{1}{|\lambda|} = \frac{2}{\omega x^*(1-x^*)|K_1|}. \tag{S61}$$

This time scale represents the characteristic time required for a small deviation $z = x - x^*$ to change appreciably near the interior equilibrium.

To force the trajectory across the otherwise non-crossable neighbourhood $\mathcal{N}_\epsilon(x^*) = \{x : |x - x^*| < \epsilon\}$, we introduce a constant local deterministic pulse control $u_{k,\epsilon}$ within this neighbourhood. Let $s = \text{sign}(x_f - x_0)$, where $s = 1$ corresponds to crossing from left to right and $s = -1$ corresponds to crossing from right to left.

During the pulse interval, $u_{k,\epsilon}$ replaces the zero- or active-control branch. Under the local approximation, the controlled dynamics is therefore

$$\dot{x} \approx \frac{\omega}{2} x^*(1-x^*)\left[K_1(x-x^*) + \Lambda_k(x^*)u_{k,\epsilon}\right]. \tag{S62}$$

Accordingly, the crossing time across $\mathcal{N}_\epsilon(x^*)$ is

$$\frac{2}{\omega}\int_{x^*-s\epsilon}^{x^*+s\epsilon} \frac{dx}{x^*(1-x^*)\left[K_1(x-x^*) + \Lambda_k(x^*)u_{k,\epsilon}\right]}. \tag{S63}$$

We choose the local pulse so that this crossing time equals the natural local time scale in Eq. (S61).

This gives

$$\frac{2}{\omega}\int_{x^*-s\epsilon}^{x^*+s\epsilon}\frac{dx}{x^*(1-x^*)\left[K_1(x-x^*)+\Lambda_k(x^*)u_{k,\epsilon}\right]}=\frac{2}{\omega x^*(1-x^*)|K_1|}. \tag{S64}$$

Under $z=x-x^*$, the left-hand side of Eq. (S64) becomes

$$\begin{aligned}&\frac{2}{\omega x^*(1-x^*)}\int_{-s\epsilon}^{s\epsilon}\frac{dz}{K_1 z+\Lambda_k(x^*)u_{k,\epsilon}}\\&=\frac{2}{\omega x^*(1-x^*)K_1}\ln\left|\frac{\Lambda_k(x^*)u_{k,\epsilon}+sK_1\epsilon}{\Lambda_k(x^*)u_{k,\epsilon}-sK_1\epsilon}\right|.\end{aligned} \tag{S65}$$

To ensure that the local vector field points in the crossing direction throughout the interval, we require

$$s\Lambda_k(x^*)u_{k,\epsilon}>|K_1|\epsilon. \tag{S66}$$

Under this condition, Eq. (S65) gives the positive crossing time

$$\frac{2}{\omega x^*(1-x^*)|K_1|}\ln\left(\frac{s\Lambda_k(x^*)u_{k,\epsilon}+|K_1|\epsilon}{s\Lambda_k(x^*)u_{k,\epsilon}-|K_1|\epsilon}\right). \tag{S67}$$

Equating this expression to the natural local time scale in Eq. (S61) yields

$$\ln\left(\frac{s\Lambda_k(x^*)u_{k,\epsilon}+|K_1|\epsilon}{s\Lambda_k(x^*)u_{k,\epsilon}-|K_1|\epsilon}\right)=1. \tag{S68}$$

Solving for $u_{k,\epsilon}$, we obtain

$$u_{k,\epsilon}\approx\frac{s|K_1|\epsilon}{\Lambda_k(x^*)}\coth\left(\frac{1}{2}\right). \tag{S69}$$

Thus, the local deterministic pulse control removes the non-crossability of the interior equilibrium by producing a finite-time crossing within $\mathcal{N}_\epsilon(x^*)$.

2.5 Theoretical analysis of the model for maximizing cumulative cooperation under a fixed budget

In this section, we consider the case $\Phi(x)<0$ over the admissible target range, so that cooperation declines without control. The desired target state is denoted by $x_f^{\mathrm{d}}\in(0,1)$. Under a fixed cost budget $\mathcal{B}$, this target may not be feasible, and the highest target state attainable is defined as $\bar{x}(\mathcal{B})$. Thus, the actual attained target is

$$x_f^{\mathrm{a}}=\min\{x_f^{\mathrm{d}},\bar{x}(\mathcal{B})\}. \tag{S70}$$

When the budget is sufficient to reach the desired target, $x_f^{\mathrm{a}} = x_f^{\mathrm{d}} \leq \bar{x}(\mathcal{B})$; when the budget is insufficient, the system reaches only the highest feasible state, $x_f^{\mathrm{a}} = \bar{x}(\mathcal{B}) < x_f^{\mathrm{d}}$.

The replicator dynamics under control can be written as

$$\dot{x} = \frac{\omega}{2} x(1-x)\left[\Phi(x) \pm u(x)\right] \equiv G(x,u). \tag{S71}$$

The plus sign corresponds to reward control, with $u(x) = u_{\mathrm{r}}(x)$, whereas the minus sign corresponds to punishment control, with $u(x) = u_{\mathrm{p}}(x)$.

The optimization problem is

$$\begin{aligned}
&\max_{x_f^{\mathrm{d}}} X(x_f^{\mathrm{d}}) = \int_{t_0}^{t_r} x(t)\,dt + \int_{t_r}^{t_m} x_f^{\mathrm{a}}\,dt + \int_{t_m}^{t_e} x(t \mid u=0)\,dt, \\
&\text{s.t.} \quad \begin{cases}
\dot{x} = G(x,u^*), & t \in [t_0, t_r], \\
x(t_0) = x_0, \qquad x(t_r) = x_f^{\mathrm{a}}, & \\
x(t) = x_f^{\mathrm{a}}, & t \in [t_r, t_m], \\
\dot{x} = G(x,0), & t \in [t_m, t_e], \\
\displaystyle\int_{t_0}^{t_m} U(x(t))\,dt = \mathcal{B}, & \\
x_f^{\mathrm{a}} = \min\left\{x_f^{\mathrm{d}}, \bar{x}(\mathcal{B})\right\}, & \\
x(t_e) = 10^{-10}. &
\end{cases}
\end{aligned} \tag{S72}$$

The cumulative cost during the controlled phase (reaching phase and maintenance phase) is

$$\mathcal{B} = J_1 + J_2 = \int_{t_0}^{t_r} U(x(t))dt + \int_{t_r}^{t_m} U(x(t))dt. \tag{S73}$$

For a desired state above initial state x_0, if the budget is insufficient to reach it, then $x_f^{\mathrm{a}} = \bar{x}(\mathcal{B})$ is determined by $J_1(x_f^{\mathrm{a}}) = \mathcal{B}$, and the maintenance phase vanishes.

We first analyze the cumulative cost J_1 during the reaching phase. In the case of $x_0 < x_f^{\mathrm{a}}$, active control is applied during $t \in [t_0, t_r]$.

When $x_0 < x_f^{\mathrm{a}} \leq 0.5$,

$$\begin{aligned} J_1 &= \int_{t_0}^{t_r} U(x(t))dt = \int_{t_0}^{t_r} (-2\Phi(x)\cdot x)^2 dt \\ &= 4\int_{x_0}^{x_f^{\mathrm{a}}} (\Phi(x))^2 \cdot x^2 \cdot \frac{dx}{\dot{x}} \\ &= 4\int_{x_0}^{x_f^{\mathrm{a}}} (\Phi(x))^2 \cdot x^2 \cdot \frac{2}{-\omega x(1-x)\Phi(x)} dx \\ &= -8\int_{x_0}^{x_f^{\mathrm{a}}} \frac{(K_1 x + K_0)x}{\omega(1-x)} dx \\ &= \frac{8}{\omega}\left(\frac{K_1}{2}\left((x_f^{\mathrm{a}})^2 - x_0^2\right) + (K_1+K_0)\left((x_f^{\mathrm{a}} - x_0) + \ln\frac{1-x_f^{\mathrm{a}}}{1-x_0}\right)\right); \end{aligned} \tag{S74}$$

When $0.5 \leq x_0 < x_f^{\mathrm{a}}$,

$$\begin{aligned} J_1 &= \int_{t_0}^{t_r} U(x(t))dt = \int_{t_0}^{t_r} (2\Phi(x)\cdot(1-x))^2 dt \\ &= 4\int_{x_0}^{x_f^{\mathrm{a}}} (\Phi(x))^2 \cdot (1-x)^2 \cdot \frac{dx}{\dot{x}} \\ &= 4\int_{x_0}^{x_f^{\mathrm{a}}} (\Phi(x))^2 \cdot (1-x)^2 \cdot \frac{2}{-\omega x(1-x)\Phi(x)} dx \\ &= -4\int_{x_0}^{x_f^{\mathrm{a}}} \frac{2(K_1 x + K_0)(1-x)}{\omega x} dx \\ &= \frac{8}{\omega}\left(\frac{K_1}{2}\left((x_f^{\mathrm{a}})^2 - x_0^2\right) - (K_1-K_0)(x_f^{\mathrm{a}} - x_0) - K_0 \ln\frac{x_f^{\mathrm{a}}}{x_0}\right); \end{aligned} \tag{S75}$$

When $x_0 < 0.5 < x_f^{\mathrm{a}}$, let $x(t_1) = 0.5$,

$$\begin{aligned} J_1 &= \int_{t_0}^{t_1} U(x(t))\,dt + \int_{t_1}^{t_r} U(x(t))\,dt \\ &= \frac{8}{\omega}\Bigg(\frac{K_1}{2}\left((x_f^{\mathrm{a}})^2 - x_0^2\right) + (K_1+K_0)\left(\frac{1}{2} - x_0 + \ln\frac{1}{2(1-x_0)}\right) \\ &\qquad - (K_1-K_0)\left(x_f^{\mathrm{a}} - \frac{1}{2}\right) - K_0\ln(2x_f^{\mathrm{a}})\Bigg); \end{aligned} \tag{S76}$$

In the case of $x_0 \geq x_f^{\mathrm{a}}$, no control is applied during the reaching phase and $J_1 = 0$.

We then analyze the cumulative cost J_2 during the maintenance phase.

$$\begin{aligned} J_2 &= \int_{t_r}^{t_m} U(x(t))dt \\ &= \int_{t_r}^{t_m} (-\Phi(x) \cdot \min\{x, 1-x\})^2 dt \\ &= ((K_1 x_f^{\mathrm{a}} + K_0) \cdot \min\{x_f^{\mathrm{a}}, 1-x_f^{\mathrm{a}}\})^2 \cdot (t_m - t_r), \end{aligned} \tag{S77}$$

Based on the cumulative cost derived above, we analyze the time spent in each phase. According to Eq. (S52), when $x_0 < x_f^{\mathrm{a}}$, the time taken by the reaching phase is

$$t_r - t_0 = \frac{2}{\omega}\left(\frac{1}{K_0}\ln\left(\frac{x_0}{x_f^{\mathrm{a}}}\right) - \frac{1}{K_1+K_0}\ln\left(\frac{1-x_0}{1-x_f^{\mathrm{a}}}\right) - \frac{K_1}{K_0(K_1+K_0)}\ln\left(\frac{K_1x_0+K_0}{K_1x_f^{\mathrm{a}}+K_0}\right)\right), \tag{S78}$$

when $x_0 \geq x_f^{\mathrm{a}}$, the time taken by the reaching phase is

$$t_r - t_0 = \frac{2}{\omega}\left(\frac{1}{K_0}\ln\left(\frac{x_f^{\mathrm{a}}}{x_0}\right) - \frac{1}{K_1+K_0}\ln\left(\frac{1-x_f^{\mathrm{a}}}{1-x_0}\right) - \frac{K_1}{K_0(K_1+K_0)}\ln\left(\frac{K_1x_f^{\mathrm{a}}+K_0}{K_1x_0+K_0}\right)\right). \tag{S79}$$

According to Eqs. (S52) and (S77), the time taken by the maintenance phase is

$$t_m - t_r = \frac{\max\{\mathcal{B} - J_1, 0\}}{((K_1 x_f^{\mathrm{a}} + K_0) \cdot \min\{x_f^{\mathrm{a}}, 1-x_f^{\mathrm{a}}\})^2}. \tag{S80}$$

The evaluation time t_e is defined by the extinction threshold $x(t_e) = 10^{-10}$. According to Eq. (S52), the time taken by the uncontrolled phase is

$$t_e - t_m = \frac{2}{\omega}\left(\frac{1}{K_0}\ln\left(\frac{10^{-10}}{x_f^{\mathrm{a}}}\right) - \frac{1}{K_1+K_0}\ln\left(\frac{1-10^{-10}}{1-x_f^{\mathrm{a}}}\right) - \frac{K_1}{K_0(K_1+K_0)}\ln\left(\frac{K_1 10^{-10}+K_0}{K_1x_f^{\mathrm{a}}+K_0}\right)\right). \tag{S81}$$

For convenience, we set $t_0 = 0$. By summing the time spent in the three stages, we obtain that when $x_0 < x_f^{\mathrm{a}}$,

$$\begin{aligned} t_e =& \frac{2}{\omega}\left(\frac{1}{K_0}\ln\left(\frac{x_0 \cdot 10^{-10}}{(x_f^{\mathrm{a}})^2}\right) - \frac{1}{K_1+K_0}\ln\left(\frac{(1-x_0)(1-10^{-10})}{(1-x_f^{\mathrm{a}})^2}\right)\right. \\ &\left. - \frac{K_1}{K_0(K_1+K_0)}\ln\left(\frac{(K_1x_0+K_0)(K_1 10^{-10}+K_0)}{(K_1x_f^{\mathrm{a}}+K_0)^2}\right)\right) + \frac{\max\{\mathcal{B} - J_1, 0\}}{((K_1 x_f^{\mathrm{a}} + K_0) \cdot \min\{x_f^{\mathrm{a}}, 1-x_f^{\mathrm{a}}\})^2}. \end{aligned} \tag{S82}$$

when $x_0 \geq x_f^{\mathrm{a}}$,

$$\begin{aligned} t_e =& \frac{2}{\omega}\left(\frac{1}{K_0}\ln\left(\frac{10^{-10}}{x_0}\right) - \frac{1}{K_1+K_0}\ln\left(\frac{1-10^{-10}}{1-x_0}\right)\right. \\ & \left. - \frac{K_1}{K_0(K_1+K_0)}\ln\left(\frac{K_1 10^{-10}+K_0}{K_1 x_0+K_0}\right)\right) + \frac{\max\{\mathcal{B}-J_1,0\}}{((K_1 x_f^{\mathrm{a}}+K_0)\cdot\min\{x_f^{\mathrm{a}},1-x_f^{\mathrm{a}}\})^2}. \end{aligned} \tag{S83}$$

We next analyze cumulative cooperation:

$$\begin{aligned} X &= X_1 + X_2 + X_3 \\ &= \int_{t_0}^{t_r} x(t)\,dt + \int_{t_r}^{t_m} x_f^{\mathrm{a}}\,dt + \int_{t_m}^{t_e} x(t \mid u=0)\,dt. \end{aligned} \tag{S84}$$

When $x_0 < x_f^{\mathrm{a}}$, according to Eq. (S40), replicator evolution under optimal control is $\dot{x} = -(\omega/2)x(1-x)\Phi(x)$, thus, the cumulative cooperation X_1 is

$$\begin{aligned} X_1 &= \int_{t_0}^{t_r} x(t)dt = \int_{x_0}^{x_f^{\mathrm{a}}} x\cdot\frac{dx}{\dot{x}} \\ &= \int_{x_0}^{x_f^{\mathrm{a}}} x\cdot\frac{2dx}{-\omega x(1-x)\Phi(x)} \\ &= -\frac{2}{\omega}\int_{x_0}^{x_f^{\mathrm{a}}}\frac{1}{(K_1x+K_0)(1-x)}dx \\ &= \frac{2}{\omega(K_1+K_0)}\ln\frac{(1-x_f^{\mathrm{a}})(K_1x_0+K_0)}{(1-x_0)(K_1x_f^{\mathrm{a}}+K_0)}; \end{aligned} \tag{S85}$$

When $x_0 \geq x_f^{\mathrm{a}}$, according to Eq. (S40), the state evolution dynamics is $\dot{x} = (\omega/2)x(1-x)\Phi(x)$, thus, the cumulative cooperation X_1 is

$$\begin{aligned} X_1 &= \int_{t_0}^{t_r} x(t)dt = \int_{x_0}^{x_f^{\mathrm{a}}} x\cdot\frac{dx}{\dot{x}} \\ &= \int_{x_0}^{x_f^{\mathrm{a}}} x\cdot\frac{2dx}{\omega x(1-x)\Phi(x)} \\ &= \frac{2}{\omega}\int_{x_0}^{x_f^{\mathrm{a}}}\frac{1}{(K_1x+K_0)(1-x)}dx \\ &= \frac{2}{\omega(K_1+K_0)}\ln\frac{(1-x_0)(K_1x_f^{\mathrm{a}}+K_0)}{(1-x_f^{\mathrm{a}})(K_1x_0+K_0)}. \end{aligned} \tag{S86}$$

The cumulative cooperation X_2 is

$$X_2 = \int_{t_r}^{t_m} x(t)dt = x_f^{\mathrm{a}}\cdot(t_m - t_r). \tag{S87}$$

The cumulative cooperation X_3 is

$$\begin{aligned} X_3 &= \int_{t_m}^{t_e} x(t|u=0)dt = \int_{x_f^{\mathrm{a}}}^{10^{-10}} x \cdot \frac{dx}{\dot{x}} \\ &= \int_{x_f^{\mathrm{a}}}^{10^{-10}} x \cdot \frac{2dx}{\omega x(1-x)\Phi(x)} \\ &= \frac{2}{\omega(K_1+K_0)} \ln \frac{(1-x_f^{\mathrm{a}})(K_0+K_1\cdot 10^{-10})}{(1-10^{-10})(K_0+K_1 x_f^{\mathrm{a}})}. \end{aligned} \tag{S88}$$

3 Extensions

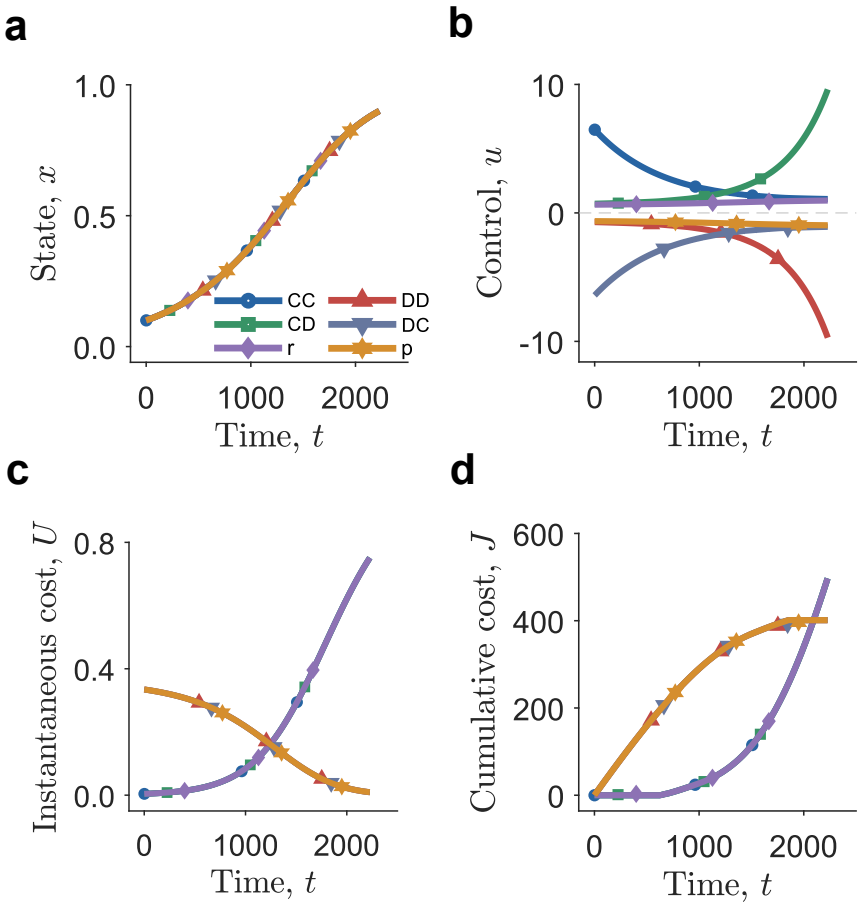
We examine whether the minority control principle remains valid when demographic stochasticity is explicitly included. We focus on the minimum-cost problem of driving cooperation from x_0 to x_f, and directly simulate the finite-population Fermi process described in Section 1.2.

The six control schemes are simulated independently. At each elementary update, the control is recalculated from the current number of cooperators using the finite-N expressions in Supplementary Table 1. The corresponding controlled payoffs are then used in the Fermi imitation rule. Thus, the stochastic state trajectories are generated separately for each scheme rather than being obtained from a common deterministic trajectory.

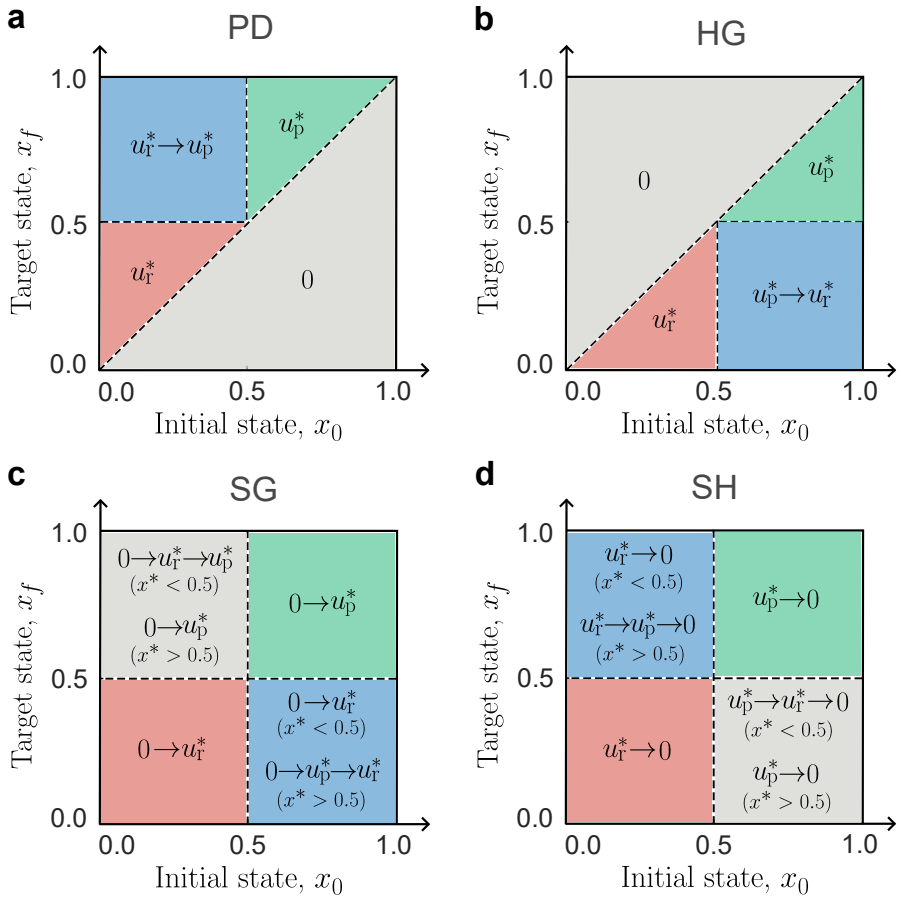
For each scheme, we collect 1000 realizations that reach x_f. The curves in Supplementary Fig. 5 show the pointwise means calculated from these realizations. To display the variability without excessive visual overlap, shaded regions are shown only for the r and p schemes and represent the pointwise mean plus or minus one standard deviation. After a trajectory first reaches x_f, its state and instantaneous cost are held at their first-passage values only for alignment on the common plotting time axis.

From Supplementary Fig. 5a, the mean cooperation level increases from x_0 towards x_f under all six schemes. The independently simulated mean trajectories are similar but not identical. The mean instantaneous costs separate into two classes whose ordering reverses as cooperation increases. At low mean cooperation levels, the schemes acting on cooperators, r, CC, and CD, tend to be less costly. At high mean cooperation levels, the schemes acting on defectors, p, DC, and DD, tend to be less costly (Supplementary Fig. 5b). For any given state, the costs of the two classes are equal at $x = 0.5$, as predicted by the deterministic analysis. However, Supplementary Fig. 5 shows costs averaged across stochastic trajectories at each time. Because the cost is a nonlinear function of the state, the crossover of the mean cost curves does not necessarily occur when the mean cooperation frequency is exactly 0.5. In the present simulations, it occurs while the mean cooperation frequency is slightly below 0.5. Independent simulation and finite sampling also produce small differences among the three curves within each class. Nevertheless, the stochastic results preserve the qualitative cost ordering predicted by the minority control principle.

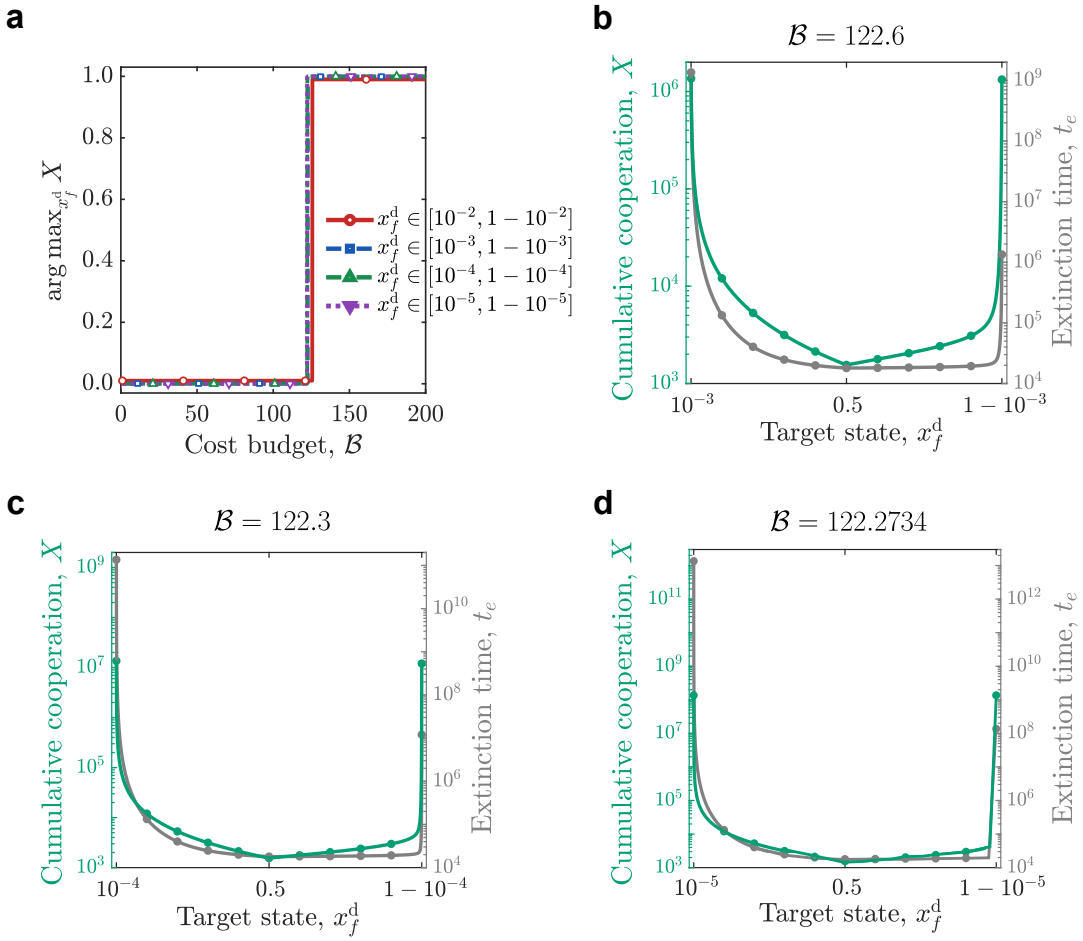
4 Supplementary Figures



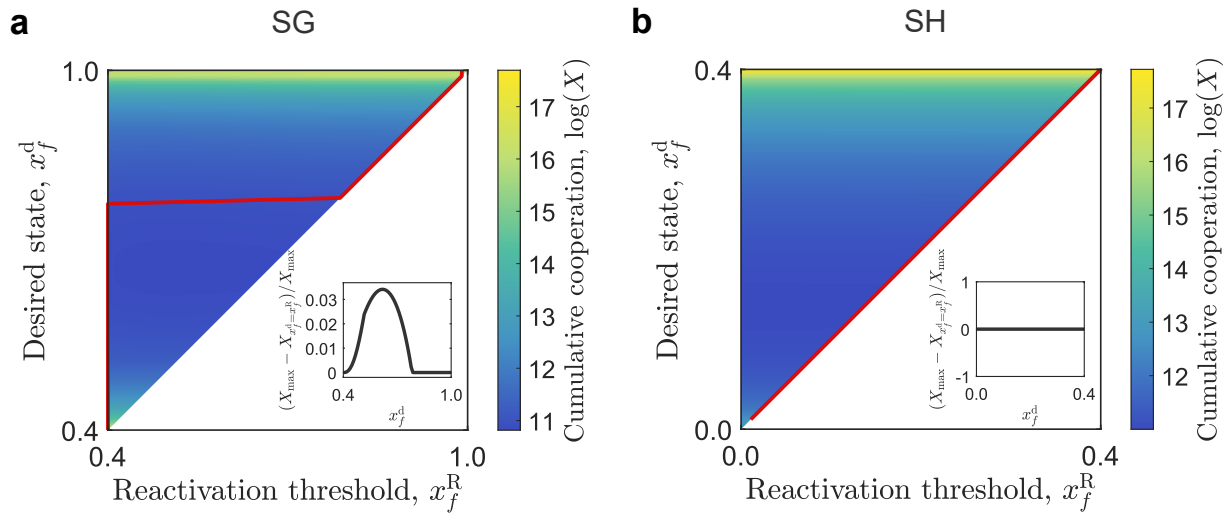
Supplementary Figure 1: Six control schemes produce an identical controlled trajectory, but their minimum costs depend only on whether control targets cooperators or defectors. Blue circles indicate control on cooperators in CC interactions, green squares indicate control on cooperators in CD interactions, purple diamonds indicate control on cooperators in all interactions (denoted by r), red upward triangles indicate control on defectors in DD interactions, dark-blue downward triangles indicate control on defectors in DC interactions, and orange stars indicate control on defectors in all interactions (denoted by p). **a**, The cooperation level x follows the same path from the initial low-cooperation state to the prescribed high-cooperation state under all six schemes, showing that distinct implementations can be dynamically equivalent. **b**, The control u required to generate this path differs across schemes. Schemes acting on cooperators require positive control, whereas schemes acting on defectors require negative control. **c**, The instantaneous cost U collapses the six schemes into two overlapping groups. The CC, CD, and r schemes share one cost curve, whereas the DD, DC, and p schemes share another cost curve. 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. **d**, The cumulative cost J preserves the same two-class structure. Costs accumulate slowly at early times for schemes acting on cooperators and rapidly at early times for schemes acting on defectors. Parameters: $a = 1$, $b = -0.3$, $c = 1.5$, $d = 0$, $\omega = 0.01$, $x_0 = 0.1$, $x_f = 0.9$.



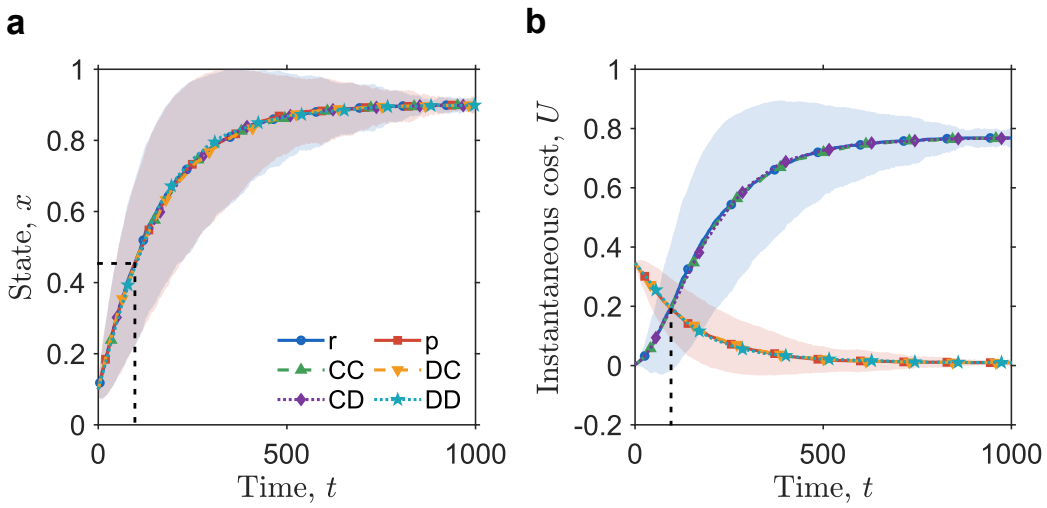
Supplementary Figure 2: The optimal regulation applies across general pairwise games. The diagrams show the optimal control required to drive the population from an initial state x_0 to a target state x_f. The framework applies to general pairwise games, including the prisoner's dilemma, harmony game, snowdrift game, and stag-hunt game. Here, u_r^* and u_p^* denote the optimal controls acting on cooperators and defectors, respectively, and 0 denotes no control. According to the minority control principle, the optimal control uses u_r^* when the evolving state satisfies $x < 0.5$ and uses u_p^* when $x > 0.5$. **a**, Optimal regulation for the prisoner's dilemma. When $x_0 < x_f$, control is required to increase cooperation. When $x_0 > x_f$, the target is reachable by natural evolution and no control is required. **b**, Optimal regulation for the harmony game. When $x_0 > x_f$, control is required to reduce cooperation. When $x_0 < x_f$, the target is reachable by natural evolution and no control is required. **c**, Optimal regulation for the snowdrift game when the stable interior equilibrium x^* lies between x_0 and x_f. Because x^* is an attractor, the system first moves naturally from x_0 toward x^*, and control is then required to move from x^* to x_f. The applied control is determined by the evolving state relative to 0.5, with u_r^* used below 0.5 and u_p^* used above 0.5. When $x^* < \min\{x_0, x_f\}$, the optimal control structure follows panel **a**, whereas when $x^* > \max\{x_0, x_f\}$, it follows panel **b**. **d**, Optimal regulation for the stag-hunt game when the unstable interior equilibrium x^* lies between x_0 and x_f. Because x^* is a repeller, control is required to move the system from x_0 to x^*, after which the system can move naturally from x^* to x_f. The applied control is determined by the evolving state relative to 0.5, with u_r^* used below 0.5 and u_p^* used above 0.5. When $x^* > \max\{x_0, x_f\}$, the optimal control structure follows panel **a**, whereas when $x^* < \min\{x_0, x_f\}$, it follows panel **b**. In **c** and **d**, crossing the interior equilibrium additionally requires a local pulse control u_ϵ^*. The pulse acts on cooperators, $u_{r,\epsilon}^*$, when $x^* < 0.5$, and on defectors, $u_{p,\epsilon}^*$, when $x^* > 0.5$.



Supplementary Figure 3: The abrupt endpoint switch persists as the admissible desired states approach the boundaries. a, Optimal desired state as a function of the cost budget $\mathcal{B}$ for four admissible ranges. In every case, increasing the budget induces a discontinuous switch from the lower endpoint to the upper endpoint, rather than a gradual movement through intermediate desired states. As the endpoints are moved closer to 0 and 1, the switching budget decreases and approaches a limiting value. **b–d,** Cumulative cooperation X (green) and extinction time t_e (grey) as functions of the desired state for $x_f^{\mathrm{d}} \in [10^{-3}, 1-10^{-3}]$, $x_f^{\mathrm{d}} \in [10^{-4}, 1-10^{-4}]$, and $x_f^{\mathrm{d}} \in [10^{-5}, 1-10^{-5}]$, respectively, with each panel shown near its corresponding critical budget. Both cumulative cooperation and extinction time first decrease and then increase as the desired state increases, and the two endpoints yield approximately equal cumulative cooperation. These results show that bringing the admissible endpoints closer to the boundaries does not alter the qualitative optimal control pattern: the optimal desired state still switches abruptly between the two endpoints as the available budget increases. Parameters: $a = 1$, $b = -0.3$, $c = 1.5$, $d = 0$, $\omega = 0.01$, $x_0 = 0.1$.



Supplementary Figure 4: The regulation with continuous maintenance remains optimal or nearly optimal in the snowdrift and stag-hunt games. The colour scale shows cumulative cooperation X over the admissible $(x_f^{\mathrm{R}}, x_f^{\mathrm{d}})$ plane, and the red curve indicates the reactivation threshold that maximizes X for each desired state. **a**, For the snowdrift game, the admissible desired states satisfy $x_f^{\mathrm{d}} \in [0.385, 0.99]$. The optimum lies either at the lowest admissible reactivation threshold or on the continuous-maintenance line $x_f^{\mathrm{R}} = x_f^{\mathrm{d}}$. The inset shows that the maximum improvement over continuous maintenance is only a few percent. **b**, For the stag-hunt game, the admissible desired states satisfy $x_f^{\mathrm{d}} \in [0.01, 0.375]$. The optimum coincides with $x_f^{\mathrm{R}} = x_f^{\mathrm{d}}$ throughout this range, and the inset confirms that intermittent regulation provides no gain over continuous maintenance. 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.



Supplementary Figure 5: Finite populations still follow the minority control principle. **a**, Evolution of the cooperation frequency x under the six control schemes. The mean trajectories under all schemes increase from x_0 to x_f and remain close to one another. **b**, Corresponding instantaneous cost U. At low mean cooperation levels, the schemes acting on cooperators, r, CC, and CD, are less costly than those acting on defectors. As mean cooperation increases, the two cost classes cross, after which p, DC, and DD become less costly. Because the costs are averaged across stochastic trajectories at each time, the crossover occurs while the mean cooperation frequency in **a** is slightly below 0.5. Nevertheless, the cost ordering on either side of the crossover is consistent with the minority control principle, showing that this principle remains qualitatively valid under finite-population stochasticity. For each scheme, we collect 1000 realizations that reach x_f. In both panels, the curves show pointwise means calculated from these realizations, and the shaded regions show the mean plus or minus one standard deviation for r and p. Parameters: $a = 1$, $b = -0.3$, $c = 1.5$, $d = 0$, $\omega = 0.01$, $N = 200$, $x_0 = 0.1$, $x_f = 0.9$.