Electronic Supplementary Material for

Optimization of institutional incentives for cooperation in structured populations

Shengxian Wang, Xiaojie Chen, Zhilong Xiao, Attila Szolnoki, and Vítor V. Vasconcelos

In electronic supplementary material, we provide a detailed theoretical analysis to explore optimal incentive protocols for the promotion of cooperation in structured populations. Specifically, we consider four different strategy update rules, describing DB updating in section 1, BD updating in section 2, IM updating in section 3, and finally PC updating in section 4. In each section, we first use the pair approximation method to explore the dynamical equation with positive or negative incentive and theoretically obtain the conditions of the minimal amounts of incentives needed for the evolution of cooperation. After that, we formulate optimal control problems for both positive and negative incentive protocols and obtain the optimal positive and negative incentive protocols by means of the approach of HJB equation. As a result, the requested cumulative costs to reach the expected final state are determined for the optimal incentive protocols.

1. DB Updating

1.1. Positive Incentive

Population structure is represented by a regular network of N nodes with degree $k > 2$. The vertices of network correspond to individuals and the edges represent who interacts with whom. Each individual plays the Prisoner's Dilemma game with its neighbors, who can either cooperate (C) or defect (D). Here, we introduce some notations. Let p_i denote the proportion of individuals with strategy i, let p_{ij} denote the proportion of ij–pairs, and finally let $q_{i|j}$ denote the conditional probability of finding an i–individual given that the neighboring node is a j–individual, where $i, j \in \{C, D\}$. By using these notations, we have that $p_C + p_D = 1$, $q_{C|i} + q_{D|i} = 1$, $p_{ij} = q_{i|j}p_j$, and $p_{CD} = p_{DC}$.

We first consider the positive incentive into the networked Prisoner's Dilemma game with DB updating [1, 2]. According to DB updating, we randomly select a focal individual to die with probability p_i ($i \in \{C, D\}$), where i represents the strategy of the focal individual. Let k_C and k_D denote the numbers of cooperators and defectors among its k neighbors with $k_D + k_C = k$. If the focal individual adopts strategy D, then the fitness of a C–neighbor is

$$f_C = 1 - \omega + \omega\{(b - c + \mu_R)(k-1)q_{C|C} + (\mu_R - c)[(k-1)q_{D|C} + 1]\}, \tag{S1}$$

and the fitness of a D–neighbor is

$$f_D = 1 - \omega + \omega[b(k-1)q_{C|D}], \tag{S2}$$

where $0 \leq \omega \leq 1$ measures the strength of selection.

Since all the neighbors of the focal individual compete for the empty site with probability proportional to their fitness, the probability that a C–neighbor replaces this empty site is given by

$$\Gamma = \frac{k_C f_C}{k_C f_C + k_D f_D}. \tag{S3}$$

Based on the above equations, p_C increases by $1/N$ with probability

$$P(\Delta p_C = \frac{1}{N}) = p_D \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} \Gamma, \tag{S4}$$

and the number of CC-pairs increases by k_C and hence p_{CC} increases by $k_C/(kN/2)$ with probability

$$P(\Delta p_{CC} = \frac{2k_C}{kN}) = p_D \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} \Gamma. \tag{S5}$$

Furthermore, we consider another case, i.e., the randomly selected focal individual adopts strategy C. In this case, the fitness of a C–neighbor is

$$f_C = 1 - \omega + \omega\{(b - c + \mu_R)[(k-1)q_{C|C} + 1] + (\mu_R - c)(k-1)q_{D|C}\}, \tag{S6}$$

and the fitness of a D–neighbor is

$$f_D = 1 - \omega + \omega\{b[(k-1)q_{C|D} + 1] + 0 \cdot (k-1)q_{D|D}\}. \tag{S7}$$

The probability that a D–neighbor replaces the empty site is given by

$$M = \frac{k_D f_D}{k_C f_C + k_D f_D}. \tag{S8}$$

Therefore, p_C decreases by $1/N$ with probability

$$P(\Delta p_C = -\frac{1}{N}) = p_C \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} M, \tag{S9}$$

and the number of CC–pairs decreases by k_C and hence p_{CC} decreases by $k_C/(kN/2)$ with probability

$$P(\Delta p_{CC} = -\frac{2k_C}{kN}) = p_C \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} M. \tag{S10}$$

We suppose that one replacement event occurs in one unit of time, and the derivative of p_C can be written as

$$\begin{aligned} \frac{dp_C}{dt} &= \frac{E(\Delta p_C)}{\Delta t} = \frac{\frac{1}{N}P(\Delta p_C = \frac{1}{N}) - \frac{1}{N}P(\Delta p_C = -\frac{1}{N})}{\frac{1}{N}} \\ &= \omega \frac{k-1}{k} p_{CD}[(\mu_R - c) + (k-1)\eta_1](q_{C|C} + q_{D|D}) + o(\omega^2), \end{aligned} \tag{S11}$$

in which $\eta_1 = (b - c + \mu_R)q_{C|C} + (\mu_R - c)q_{D|C} - bq_{C|D}$. Accordingly, the derivative of p_{CC} is given by

$$\begin{aligned} \frac{dp_{CC}}{dt} &= \frac{E(\Delta p_{CC})}{\Delta t} = \frac{\sum_{k_C=0}^{k} \frac{2k_C}{kN} P(\Delta p_{CC} = \frac{2k_C}{kN}) - \sum_{k_C=0}^{k} \frac{2k_C}{kN} P(\Delta p_{CC} = -\frac{2k_C}{kN})}{\frac{1}{N}} \\ &= \frac{2p_{CD}}{k}[1 + (k-1)(q_{C|D} - q_{C|C})] + o(\omega). \end{aligned} \tag{S12}$$

Due to $q_{C|C} = \frac{p_{CC}}{p_C}$, we have

$$\frac{dq_{C|C}}{dt} = \frac{d}{dt}(\frac{p_{CC}}{p_C}) = \frac{2p_{CD}}{kp_C}[1 + (k-1)(q_{C|D} - q_{C|C})] + o(\omega). \tag{S13}$$

Other variables, such as p_D and $q_{D|C}$, can also be expressed by p_C and $q_{C|C}$ through appropriate calculation, and then the dynamical system can be described by p_C and $q_{C|C}$. Rewriting the right-hand expressions of Eqs. (S11) and (S13) as functions of p_C and $q_{C|C}$ yields the dynamical equation given by

$$\begin{cases} \frac{dp_C}{dt} = \omega \Psi^R_{\text{DB}}(p_C, q_{C|C}) + o(\omega^2), \\ \frac{dq_{C|C}}{dt} = \Phi^R_{\text{DB}}(p_C, q_{C|C}) + o(\omega), \end{cases} \tag{S14}$$

where

$$\begin{cases} \Psi^R_{\text{DB}}(p_C, q_{C|C}) = \frac{k-1}{k} p_{CD}[(-c + \mu_R) + (k-1)\eta_1](q_{C|C} + q_{D|D}), \\ \Phi^R_{\text{DB}}(p_C, q_{C|C}) = \frac{2p_{CD}}{kp_C}[1 + (k-1)(q_{C|D} - q_{C|C})]. \end{cases} \tag{S15}$$

Under weak selection ($w \ll 1$), the velocity of $q_{C|C}$ can be large, and it may rapidly converge to the root defined by $\Phi^R_{\text{DB}}(p_C, q_{C|C}) = 0$ as time $t \to +\infty$. Thus, we get

$$q_{C|C} = p_C + \frac{1}{k-1}(1 - p_C). \tag{S16}$$

Accordingly, the dynamical equation described by Eq. (S14) becomes

$$\frac{dp_C}{dt} = \frac{\omega(k-2)[b + k(\mu_R - c)]}{k-1} p_C(1 - p_C) + o(\omega^2), \tag{S17}$$

which has two fixed points $p_C = 0$ and $p_C = 1$. We define the function $F_{\text{DB}}(p_C, \mu_R, t)$ as

$$F_{\text{DB}}(p_C, \mu_R, t) = \frac{\omega(k-2)[b + k(\mu_R - c)]}{k-1} p_C(1 - p_C) + o(\omega^2). \quad \text{(S18)}$$

This function is a continuously differentiable function, and the derivative of $F_{\text{DB}}(p_C, \mu_R, t)$ with respect to p_C is

$$\frac{dF_{\text{DB}}}{dp_C} = \frac{\omega(k-2)[b + k(\mu_R - c)]}{k-1}(1 - 2p_C) + o(\omega^2). \quad \text{(S19)}$$

For $\mu_R > c - \frac{b}{k}$, we have $\frac{dF_{\text{DB}}}{dp_C}|_{p_C=1} = -\frac{\omega(k-2)[b+k(\mu_R-c)]}{k-1} < 0$ and $\frac{dF_{\text{DB}}}{dp_C}|_{p_C=0} = \frac{\omega(k-2)[b+k(\mu_R-c)]}{k-1} > 0$. This means that the fixed point $p_C = 1$ is stable and $p_C = 0$ unstable, i.e., cooperators prevail over defectors.

Lastly, we then study the special case of $\mu_R = 0$. In this case, we can see that for $b/c > k$, the fixed point $p_C = 1$ is stable and $p_C = 0$ unstable. Thus, we obtain the condition $b/c > k$ for the evolution of cooperation as previously obtained in Refs. [1, 2].

1.2. Negative Incentive

In this subsection, we then consider the negative incentive into the networked Prisoner's Dilemma game with DB updating, and the payoff matrix is given by Eq. (4.3) in the main text. According to DB updating, if the focal individual adopts strategy D, then the fitness of a C–neighbor is

$$f_C = 1 - \omega + \omega\{(b - c)(k-1)q_{C|C} - c[(k-1)q_{D|C} + 1]\}, \quad \text{(S20)}$$

and the fitness of a D–neighbor is

$$f_D = 1 - \omega + \omega\{(b - \mu_P)(k-1)q_{C|D} - \mu_P[(k-1)q_{D|D} + 1]\}. \quad \text{(S21)}$$

The probability that a C–neighbor replaces the empty site is given by the expression Γ in Eq. (S3). Therefore, p_C increases by $1/N$ with probability

$$P(\Delta p_C = \frac{1}{N}) = p_D \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} \Gamma. \quad \text{(S22)}$$

Accordingly, the number of CC–pairs increases by k_C and hence p_{CC} increases by $k_C/(kN/2)$ with probability

$$P(\Delta p_{CC} = \frac{2k_C}{kN}) = p_D \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} \Gamma. \quad \text{(S23)}$$

In addition, we consider another case where the focal individual adopts strategy D. In this case, the fitness of a C–neighbor is

$$f_C = 1 - \omega + \omega\{(b - c)[(k-1)q_{C|C} + 1] - c(k-1)q_{D|C}\}, \quad \text{(S24)}$$

and the fitness of a D–neighbor is

$$f_D = 1 - \omega + \omega\{(b - \mu_P)[(k-1)q_{C|D} + 1] - \mu_P(k-1)q_{D|D}\}. \tag{S25}$$

The probability that a D–neighbor replaces the empty site can be also given by the expression M in Eq. (S8). Thus, p_C decreases by $1/N$ with probability

$$P(\Delta p_C = -\frac{1}{N}) = p_C \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} M. \tag{S26}$$

Accordingly, the number of CC–pairs decreases by k_C and p_{CC} decreases by $k_C/(kN/2)$ with probability

$$P(\Delta p_{CC} = -\frac{2k_C}{kN}) = p_C \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} M. \tag{S27}$$

Based on these calculations, we obtain the time derivative of p_C given by

$$\begin{aligned} \frac{dp_C}{dt} &= \frac{E(\Delta p_C)}{\Delta t} = \frac{\frac{1}{N}P(\Delta p_C = \frac{1}{N}) - \frac{1}{N}P(\Delta p_C = -\frac{1}{N})}{\frac{1}{N}} \\ &= \frac{\omega(k-1)}{k} p_{CD}[\mu_P - c + (k-1)\eta_2](q_{C|C} + q_{D|D}) + o(\omega^2), \end{aligned} \tag{S28}$$

in which $\eta_2 = (b - c)q_{C|C} - cq_{D|C} + (\mu_P - b)q_{C|D} + \mu_P q_{D|D}$. And the time derivative of p_{CC} is given by

$$\begin{aligned} \frac{dp_{CC}}{dt} &= \frac{E(\Delta p_{CC})}{\Delta t} = \frac{\sum_{k_C=0}^{k} \frac{2k_C}{kN} P(\Delta p_{CC} = \frac{2k_C}{kN}) - \sum_{k_C=0}^{k} \frac{2k_C}{kN} P(\Delta p_{CC} = -\frac{2k_C}{kN})}{\frac{1}{N}} \\ &= \frac{2p_{CD}}{k}[1 + (k-1)(q_{C|D} - q_{C|C})] + o(\omega). \end{aligned} \tag{S29}$$

Furthermore, we have

$$\frac{dq_{C|C}}{dt} = \frac{2p_{CD}}{kp_C}[1 + (k-1)(q_{C|D} - q_{C|C})] + o(\omega). \tag{S30}$$

Hence, the dynamical equation can be described by

$$\begin{cases} \frac{dp_C}{dt} = \omega \Psi_{\text{DB}}^P(p_C, q_{C|C}) + o(\omega^2), \\ \frac{dq_{C|C}}{dt} = \Phi_{\text{DB}}^P(p_C, q_{C|C}) + o(\omega), \end{cases} \tag{S31}$$

where

$$\begin{cases} \Psi_{\text{DB}}^P(p_C, q_{C|C}) = \frac{k-1}{k} p_{CD}[\mu_P - c + (k-1)\eta_2](q_{C|C} + q_{D|D}), \\ \Phi_{\text{DB}}^P(p_C, q_{C|C}) = \frac{2p_{CD}}{kp_C}[1 + (k-1)(q_{C|D} - q_{C|C})]. \end{cases}$$

Under weak selection, the velocity of $q_{C|C}$ can be large, and it may rapidly converge to the root defined by $\Phi^{P}_{\text{DB}}(p_C, q_{C|C}) = 0$ as time $t \to +\infty$. Thus, we get

$$q_{C|C} = p_C + \frac{1}{k-1}(1-p_C). \tag{S32}$$

Correspondingly, the dynamical equation described by Eq. (S31) becomes

$$\frac{dp_C}{dt} = \frac{\omega(k-2)[b+k(\mu_P - c)]}{k-1} p_C(1-p_C) + o(\omega^2), \tag{S33}$$

which has two fixed points $p_C = 0$ and $p_C = 1$. We define the function $F_{\text{DB}}(p_C, \mu_P, t)$ as

$$F_{\text{DB}}(p_C, \mu_P, t) = \frac{\omega(k-2)[b+k(\mu_P - c)]}{k-1} p_C(1-p_C) + o(\omega^2), \tag{S34}$$

and the derivative of $F_{\text{DB}}(p_C, \mu_P, t)$ with respect to p_C is

$$\frac{dF_{\text{DB}}}{dp_C} = \frac{\omega(k-2)[b+k(\mu_P - c)]}{k-1}(1-2p_C) + o(\omega^2). \tag{S35}$$

Hence, for $\mu_P > c - \frac{b}{k}$ we have $\frac{dF_{\text{DB}}}{dp_C}|_{p_C=1} = -\frac{\omega(k-2)[b+k(\mu_P-c)]}{k-1} < 0$ and $\frac{dF_{\text{DB}}}{dp_C}|_{p_C=0} = \frac{\omega(k-2)[b+k(\mu_P-c)]}{k-1} > 0$, which means that the fixed point $p_C = 1$ is stable and $p_C = 0$ unstable, i.e., cooperators prevail over defectors. Particularly, when $\mu_P = 0$, we can see that for $b/c > k$, the fixed point $p_C = 1$ is stable and $p_C = 0$ unstable. Thus, we obtain the condition $b/c > k$ for the evolution of cooperation as obtained in Refs. [1, 2].

1.3. Optimal Incentive Protocols

In subsections 1.1 and 1.2, we have theoretically derived the dynamical system with positive or negative incentive by means of the pair approximation method in the limit of weak selection, which is given by

$$\frac{dp_C}{dt} = F_{\text{DB}}(p_C, \mu_v, t) = \frac{\omega(k-2)[b+k(\mu_v - c)]}{k-1} p_C(1-p_C) + o(\omega^2). \tag{S36}$$

As noted, this dynamical system has two equilibria which are $p_C = 0$ and $p_C = 1$. If $\mu_v > c - \frac{b}{k}$, the former is unstable and the latter is stable, which means that cooperation will be promoted in the long run. Since providing incentive is costly, our principal goal is to explore the optimal incentive protocol that is still able not only to promote cooperation, but also requires a minimal cost. To do that, we now solve the formulated optimal control problems for DB updating.

First, we solve the optimal control problem for rewarding. The Hamiltonian function $H_{\text{DB}}(p_C, \mu_R, t)$ is defined as

$$H_{\text{DB}}(p_C, \mu_R, t) = \frac{(kNp_C\mu_R)^2}{2} + \frac{\partial J_R^*}{\partial p_C} F_{\text{DB}}(p_C, \mu_R, t), \tag{S37}$$

where J_R^* is the optimal cost function of p_C and t for the optimal rewarding protocol, given as

$$J_R^* = \int_0^{t_f} \frac{(kNp_C\mu_R^*)^2}{2} dt. \tag{S38}$$

By solving $\frac{\partial H_{\mathrm{DB}}}{\partial \mu_R} = 0$, we know that the optimal rewarding protocol μ_R^* should satisfy

$$\mu_R^* = -\frac{\omega(k-2)(1-p_C)}{N^2 k(k-1)p_C} \frac{\partial J_R^*}{\partial p_C}. \tag{S39}$$

Generally, we should solve the canonical equations of Eq. (S37) to obtain the optimal rewarding protocol [3–5]. Yet, the obtained dynamical systems are nonlinear which greatly increases the complexity of obtaining the exact expression of the optimal protocols by a direct calculation. Instead, to solve the optimal control problem we use the dynamic programming method, HJB equation for continuous-time systems [3–5]. This equation can be written as

$$-\frac{\partial J_R^*}{\partial t} = H_{\mathrm{DB}}(p_C, \mu_R^*, t). \tag{S40}$$

By substituting Eq. (S39) into the above HJB equation, we have

$$-\frac{\partial J_R^*}{\partial t} = \frac{(kNp_C\mu_R^*)^2}{2} + \frac{\omega(k-2)[b+k(\mu_R^* - c)]}{k-1} p_C(1-p_C) \frac{\partial J_R^*}{\partial p_C}. \tag{S41}$$

Since we assume that the terminal time t_f is not fixed, the optimal cost function $J_R^*(p_C, t)$ is independent of t. Consequently, we have

$$\frac{\partial J_R^*}{\partial t} = 0. \tag{S42}$$

We then yield

$$\frac{\partial J_R^*}{\partial p_C} = 0 \ \text{ or } \ \frac{\partial J_R^*}{\partial p_C} = \frac{2N^2(k-1)(b-ck)p_C}{\omega(k-2)(1-p_C)}. \tag{S43}$$

As $\mu_R > 0$ and $p_C \in (0, 1)$, we have

$$\frac{\partial J_R^*}{\partial p_C} < 0. \tag{S44}$$

From Eq. (S44), we can see that this inequality is obviously satisfied for $b/c \geq k$. Instead, we consider the case, i.e., $b/c < k$, and hence only $\frac{\partial J_R^*}{\partial p_C} = \frac{2N^2(k-1)(b-ck)p_C}{\omega(k-2)(1-p_C)}$ holds. By substituting this equation into Eq. (S39), we obtain the optimal rewarding level as

$$\mu_R^* = \frac{2(ck-b)}{k}. \tag{S45}$$

With this μ_R^* the dynamical equation thus becomes

$$\frac{dp_C}{dt} = \frac{\omega(k-2)(ck-b)}{k-1} p_C(1-p_C), \tag{S46}$$

where the initial fraction of cooperators in the population is denoted by $p_0 = p_C(0)$. The solution of this equation is

$$p_C = \frac{1}{1 + \frac{1-p_0}{p_0} e^{-\beta_{\mathrm{DB}} t}}, \tag{S47}$$

where $\beta_{\mathrm{DB}} = \frac{\omega(k-2)(ck-b)}{k-1}$. It also means that the dynamical system needs infinitely long time to reach the full cooperation state from the initial $p_0 < 1$. To avoid it, we suppose that the terminal state $p_C(t_f)$ is $1-\delta$, where δ is the parameter determining the cooperation level at the terminal time. Due to $b/c < k$, p_C increases monotonically over time t, which leads to $p_C(t_f) > p_0$.

Furthermore, the cumulative cost required by the optimal rewarding level μ_R^* for the dynamical system to reach the expected terminal state $p_C(t_f)$ becomes

$$J_R^* = \frac{(kN\mu_R^*)^2}{2\beta_{\mathrm{DB}}}[p_0 - 1 + \delta + \ln(\frac{1-p_0}{\delta})]. \tag{S48}$$

The optimal control problem for punishment can be solved similarly and for the μ_P^* optimal level we have

$$\mu_P^* = \frac{2}{k}(ck-b) \tag{S49}$$

and

$$p_C = \frac{1}{1 + \frac{1-p_0}{p_0} e^{-\beta_{\mathrm{DB}} t}}. \tag{S50}$$

Hence, the cumulative cost produced by the optimal punishing protocol μ_P^* becomes

$$J_P^* = \frac{(kN\mu_P^*)^2}{2\beta_{\mathrm{DB}}}[p_0 - 1 + \delta + \ln(\frac{1-\delta}{p_0})]. \tag{S51}$$

From these results we can conclude that the optimal levels of negative and positive incentives are identical, i.e., $\mu_R^* = \mu_P^*$, but their cumulative costs could be different. For a proper comparison we can calculate their difference, which is

$$J_R^* - J_P^* = \frac{(kN\mu_v^*)^2}{2\beta_{\mathrm{DB}}} \ln[\frac{p_0(1-p_0)}{\delta(1-\delta)}], \tag{S52}$$

where $\mu_v^* = \mu_R^* = \mu_P^*$. In this work we assume that $p_0 > 0$, since we do not consider behavioral mutations or errors of strategy updating. In addition, $p_0 < p_C(t_f) = 1-\delta$ and we have $p_0 + \delta < 1$. Thus when $\delta < p_0$, we have $J_R^* > J_P^*$, which means that for DB updating the optimal punishment always requires lower cumulative cost than the usage of optimal reward. But when $\delta > p_0$, we have $J_R^* < J_P^*$, which means that for DB updating the optimal reward always requires lower cumulative cost than the usage of optimal punishment. This observations are supported by numerical calculations and Monte Carlo simulations as plotted in figure 4 and figure S5, respectively.

2. BD Updating

2.1. Positive Incentive

According to BD update rule [1, 2], we randomly choose a focal individual for reproduction proportional to fitness who has k_C cooperators and k_D defectors among its k neighbors. If the focal individual adopts strategy C, then the fitness of the focal individual is given by

$$f_C = 1 - \omega + \omega[(b-c)k_C - ck_D]. \tag{S53}$$

Since the offspring of the selected individual replaces one of its neighbors randomly, the probability that p_C increases by $1/N$ is

$$P(\Delta p_C = \frac{1}{N}) = p_C \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|C})^{k_C} (q_{D|C})^{k_D} \frac{f_C}{\bar{f}} \frac{k_D}{k}, \tag{S54}$$

where $\bar{f}$ represents the average fitness of the whole population. In this case, the number of CC–pairs increases by $(k-1)q_{C|D}+1$ and p_{CC} increases by $[(k-1)q_{C|D}+1]/(kN/2)$ with probability

$$P(\Delta p_{CC} = \frac{(k-1)q_{C|D}+1}{kN/2}) = p_C \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|C})^{k_C} (q_{D|C})^{k_D} \frac{f_C}{\bar{f}} \frac{k_D}{k}. \tag{S55}$$

In the alternative case, the randomly selected focal individual adopts strategy D. Here the fitness of the focal individual is given by

$$f_D = 1 - \omega + \omega(bk_C + 0 \cdot k_D),$$

and therefore p_C decreases by $1/N$ with probability

$$P(\Delta p_C = -\frac{1}{N}) = p_D \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} \frac{f_D}{\bar{f}} \frac{k_C}{k}. \tag{S56}$$

Consequently, the number of CC–pairs decreases by $(k-1)q_{C|C}$ and p_{CC} decreases by $(k-1)q_{C|C}/(kN/2)$ with probability

$$P(\Delta p_{CC} = -\frac{(k-1)q_{C|C}}{kN/2}) = p_D \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} \frac{f_D}{\bar{f}} \frac{k_C}{k}. \tag{S57}$$

Here, the average fitness of the whole population is thus denoted by

$$\begin{aligned}\bar{f} &= p_C \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|C})^{k_C} (q_{D|C})^{k_D} f_C + p_D \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} f_D \\ &= 1 - \omega + k\omega[(b - c + \mu_R)p_{CC} + (\mu_R - c)p_{CD} + bp_{CD}].\end{aligned} \tag{S58}$$

Based on these calculations, we respectively obtain the time derivatives of p_C and p_{CC} as

$$\begin{aligned}\frac{dp_C}{dt} &= \frac{E(\Delta p_C)}{\Delta t} = \frac{\frac{1}{N}P(\Delta p_C = \frac{1}{N}) - \frac{1}{N}P(\Delta p_C = -\frac{1}{N})}{\frac{1}{N}} \\ &= \frac{\omega p_{CD}}{\bar{f}}\{(\mu_R - c - b) + (k-1)[(b - c + \mu_R)q_{C|C} + (\mu_R - c)q_{D|C} - bq_{C|D}]\} + o(\omega^2)\end{aligned} \tag{S59}$$

and

$$\begin{aligned}\frac{dp_{CC}}{dt} &= \frac{E(\Delta p_{CC})}{\Delta t} = \frac{\frac{(k-1)q_{C|D}+1}{kN/2}P(\Delta p_{CC} = \frac{(k-1)q_{C|D}+1}{kN/2}) - \frac{(k-1)q_{C|C}}{kN/2}P(\Delta p_{CC} = -\frac{(k-1)q_{C|C}}{kN/2})}{\frac{1}{N}} \\ &= \frac{2p_{CD}}{k}[(k-1)(q_{C|D} - q_{C|C}) + 1] + o(\omega).\end{aligned} \tag{S60}$$

Furthermore, we have

$$\frac{dq_{C|C}}{dt} = \frac{2p_{CD}}{kp_C}[(k-1)(q_{C|D} - q_{C|C}) + 1] + o(\omega). \tag{S61}$$

Hence, the dynamical equation is described by

$$\begin{cases} \frac{dp_C}{dt} = \omega\Psi^R_{\text{BD}}(p_C, q_{C|C}) + o(\omega^2), \\ \frac{dq_{C|C}}{dt} = \Phi^R_{\text{BD}}(p_C, q_{C|C}) + o(\omega), \end{cases} \tag{S62}$$

where

$$\begin{cases} \Psi^R_{\text{BD}}(p_C, q_{C|C}) = \frac{p_{CD}}{\bar{f}}\{(\mu_R - c - b) + (k-1)[(b - c + \mu_R)q_{C|C} + (\mu_R - c)q_{D|C} - bq_{C|D}]\}, \\ \Phi^R_{\text{BD}}(p_C, q_{C|C}) = \frac{2p_{CD}}{kp_C}[(k-1)(q_{C|D} - q_{C|C}) + 1]. \end{cases}$$

Under weak selection, the velocity of $q_{C|C}$ can be large, and it may rapidly converge to the root defined by $\Phi^R_{\text{BD}}(p_C, q_{C|C}) = 0$ as time $t \to +\infty$. Thus, we get

$$q_{C|C} = p_C + \frac{1}{k-1}(1 - p_C). \tag{S63}$$

Accordingly, the dynamical equation described by Eq. (S62) becomes

$$\frac{dp_C}{dt} = \frac{\omega k(k-2)(\mu_R - c)}{k-1}p_C(1 - p_C) + o(\omega^2), \tag{S64}$$

which has two fixed points $p_C = 0$ and $p_C = 1$. We define the function $F_{\text{BD}}(p_C, \mu_R, t)$ as

$$F_{\text{BD}}(p_C, \mu_R, t) = \frac{\omega k(k-2)(\mu_R - c)}{k-1}p_C(1 - p_C) + o(\omega^2), \tag{S65}$$

and the derivative of $F_{\text{BD}}(p_C, \mu_R, t)$ with respect to p_C is

$$\frac{dF_{\text{BD}}}{dp_C} = \frac{\omega k(k-2)(\mu_R - c)}{k-1}(1 - 2p_C) + o(\omega^2). \tag{S66}$$

For $\mu_R > c$, we have $\frac{dF_{\rm BD}}{dp_C}|_{p_C=1} = -\frac{\omega k(k-2)(\mu_R-c)}{k-1} < 0$ and $\frac{dF_{\rm BD}}{dp_C}|_{p_C=0} = \frac{\omega k(k-2)(\mu_R-c)}{k-1} > 0$. This implies that the fixed point $p_C = 1$ is stable and $p_C = 0$ unstable, i.e., cooperators prevail over defectors. Particularly, when $\mu_R = 0$, we can see that the fixed point $p_C = 0$ is always stable and $p_C = 1$ unstable, which means that cooperation cannot emerge under BD update rule as obtained in Refs. [1, 2].

2.2. Negative Incentive

In this subsection, we consider the negative incentive into the networked Prisoner's Dilemma game with BD updating. According to BD updating, a focal individual is randomly selected for reproduction who has k_C cooperators and k_D defectors among its k neighbors. Here, we first consider the focal individual adopts strategy C. Then, the fitness of the focal individual is given by

$$f_C = 1 - \omega + \omega[k_C(b-c) + k_D(-c)], \tag{S67}$$

and therefore p_C increases by $1/N$ with probability

$$P(\Delta p_C = \frac{1}{N}) = p_C \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|C})^{k_C} (q_{D|C})^{k_D} \frac{f_C}{\bar{f}} \frac{k_D}{k}, \tag{S68}$$

where $\bar{f}$ denotes the average fitness of the whole population. And the number of CC–pairs increases by $(k-1)q_{C|D}+1$ and therefore p_{CC} increases by $[(k-1)q_{C|D}+1]/(kN/2)$ with probability

$$P(\Delta p_{CC} = \frac{(k-1)q_{C|D}+1}{kN/2}) = p_C \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|C})^{k_C} (q_{D|C})^{k_D} \frac{f_C}{\bar{f}} \frac{k_D}{k}. \tag{S69}$$

In addition, we consider another case, that is, the randomly selected focal individual adopts strategy D. In this case, the fitness of the focal individual is given by

$$f_D = 1 - \omega + \omega[k_C(b-\mu_P) + k_D(-\mu_P)],$$

and therefore p_C decreases by $1/N$ with probability

$$P(\Delta p_C = -\frac{1}{N}) = p_D \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} \frac{f_D}{\bar{f}} \frac{k_C}{k}. \tag{S70}$$

And the number of CC–pairs decreases by $(k-1)q_{C|C}$ and therefore p_{CC} decreases by $(k-1)q_{C|C}/(kN/2)$ with probability

$$P(\Delta p_{CC} = -\frac{(k-1)q_{C|C}}{kN/2}) = p_D \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} \frac{f_D}{\bar{f}} \frac{k_C}{k}. \tag{S71}$$

Here, the average fitness of whole population can be calculated by

$$\begin{aligned}\bar{f} &= p_C \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|C})^{k_C} (q_{D|C})^{k_D} f_C + p_D \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} f_D \\ &= 1-\omega+\omega k[(b-c)p_{CC} - cp_{CD} + (b-\mu_P)p_{CD} - \mu_P p_{DD}].\end{aligned} \tag{S72}$$

From these calculations, we respectively obtain the time derivatives of p_C and p_{CC} as

$$\begin{aligned}\frac{dp_C}{dt} &= \frac{E(\Delta p_C)}{\Delta t} = \frac{\frac{1}{N}P(\Delta p_C = \frac{1}{N}) - \frac{1}{N}P(\Delta p_C = -\frac{1}{N})}{\frac{1}{N}} \\ &= \frac{\omega p_{CD}}{\bar{f}}\{\mu_P - c - b + (k-1)[(b-c)q_{C|C} - cq_{D|C} + (\mu_P - b)q_{C|D} + \mu_P q_{D|D}]\} + o(\omega^2),\end{aligned} \tag{S73}$$

and

$$\begin{aligned}\frac{dp_{CC}}{dt} &= \frac{E(\Delta p_{CC})}{\Delta t} = \frac{\frac{(k-1)q_{C|D}+1}{kN/2}P(\Delta p_{CC} = \frac{(k-1)q_{C|D}+1}{kN/2}) - \frac{(k-1)q_{C|C}}{kN/2}P(\Delta p_{CC} = -\frac{(k-1)q_{C|C}}{kN/2})}{\frac{1}{N}} \\ &= \frac{2}{k}p_{CD}[(k-1)(q_{C|D} - q_{C|C}) + 1] + o(\omega).\end{aligned} \tag{S74}$$

Furthermore, we have

$$\frac{dq_{C|C}}{dt} = \frac{2p_{CD}}{kp_C}[(k-1)(q_{C|D} - q_{C|C}) + 1] + o(\omega). \tag{S75}$$

Hence, the dynamical equation is described by

$$\begin{cases} \frac{dp_C}{dt} = \omega\Psi^P_{\text{BD}}(p_C, q_{C|C}) + o(\omega^2), \\ fracdq_{C|C}dt = \Phi^P_{\text{BD}}(p_C, q_{C|C}) + o(\omega), \end{cases} \tag{S76}$$

where

$$\begin{cases} \Psi^P_{\text{BD}}(p_C, q_{C|C}) = \frac{p_{CD}}{\bar{f}}\{\mu_P - c - b + (k-1)[(b-c)q_{C|C} - cq_{D|C} + (\mu_P - b)q_{C|D} + \mu_P q_{D|D}]\}, \\ \Phi^P_{\text{BD}}(p_C, q_{C|C}) = \frac{2p_{CD}}{kp_C}[(k-1)(q_{C|D} - q_{C|C}) + 1]. \end{cases}$$

Under weak selection, the velocity of $q_{C|C}$ can be large, and it may rapidly converge to the root defined by $\Phi^P_{\text{BD}}(p_C, q_{C|C}) = 0$ as time $t \to +\infty$. Thus, we get

$$q_{C|C} = p_C + \frac{1}{k-1}(1-p_C). \tag{S77}$$

Accordingly, the system described by Eq. (S76) becomes

$$\frac{dp_C}{dt} = \frac{\omega k(k-2)(\mu_P - c)}{k-1}p_C(1-p_C) + o(\omega^2), \tag{S78}$$

which has two fixed points, $p_C = 0$ and $p_C = 1$. We define the function $F_{\text{BD}}(p_C, \mu_P, t)$ as

$$F_{\text{BD}}(p_C, \mu_P, t) = \frac{\omega k(k-2)(\mu_P - c)}{k-1} p_C(1-p_C) + o(\omega^2), \tag{S79}$$

and the derivative of $F_{\text{BD}}(p_C, \mu_P, t)$ with respect to p_C is

$$\frac{dF_{\text{BD}}}{dp_C} = \frac{\omega k(k-2)(\mu_P - c)}{k-1}(1-2p_C) + o(\omega^2). \tag{S80}$$

For $\mu_P > c$, we have $\frac{dF_{\text{BD}}}{dp_C}|_{p_C=1} = -\frac{\omega k(k-2)(\mu_P-c)}{k-1} < 0$ and $\frac{dF_{\text{BD}}}{dp_C}|_{p_C=0} = \frac{\omega k(k-2)(\mu_P-c)}{k-1} > 0$ which implies the fixed point $p_C = 1$ is stable and $p_C = 0$ is unstable, i.e., cooperators prevail over defectors. Particularly, when $\mu_P = 0$, we can see that the fixed point $p_C = 0$ is always stable and $p_C = 1$ unstable, which means that cooperation can never emerge under BD update rule as obtained in Refs. [1, 2].

2.3. Optimal Incentive Protocols

In subsections 2.1 and 2.2, we have theoretically obtained the dynamical equation with positive or negative incentive by means of the pair approximation approach in the limit of weak selection, which is given by

$$\frac{dp_C}{dt} = F_{\text{BD}}(p_C, \mu_v, t) = \frac{\omega k(k-2)(\mu_v - c)}{k-1} p_C(1-p_C) + o(\omega^2)\,, \tag{S81}$$

having $p_C = 0$ and $p_C = 1$ fixed points. If $\mu_v > c$, the former is unstable and the latter is stable, indicating that cooperation will be promoted in the long run. Furthermore, to explore the optimal rewarding and punishing protocols, we now employ the approach of HJB equation to solve the formulated optimal control problems for this BD updating.

First we solve the optimal control problem for rewarding. We define the Hamiltonian function $H_{\text{BD}}(p_C, \mu_R, t)$ as

$$H_{\text{BD}}(p_C, \mu_R, t) = \frac{(kNp_C\mu_R)^2}{2} + \frac{\partial J_R^*}{\partial p_C} F_{\text{BD}}(p_C, \mu_R, t), \tag{S82}$$

where J_R^* is the optimal cost function of p_C and t for the optimal rewarding protocol given as

$$J_R^* = \int_0^{t_f} \frac{(kNp_C\mu_R)^2}{2} dt. \tag{S83}$$

Solving $\frac{\partial H_{\text{BD}}}{\partial \mu_R} = 0$, we know that the optimal rewarding protocol μ_R^* should satisfy

$$\mu_R^* = -\frac{\omega(k-2)(1-p_C)}{k(k-1)N^2 p_C} \frac{\partial J_R^*}{\partial p_C}. \tag{S84}$$

The corresponding HJB equation [3–5] for dynamical system with positive incentive can be written as

$$-\frac{\partial J_R^*}{\partial t} = H_{\text{BD}}(p_C, \mu_R^*, t). \tag{S85}$$

Since we assume that the terminal time t_f is not fixed, the optimal cost function $J_R^*(p_C, t)$ is independent of t. Consequently, we have

$$\frac{\partial J_R^*}{\partial t} = 0. \tag{S86}$$

We then obtain

$$\frac{\partial J_R^*}{\partial p_C} = 0 \quad \text{or} \quad \frac{\partial J_R^*}{\partial p_C} = -\frac{2N^2 k(k-1)cp_C}{\omega(k-2)(1-p_C)}. \tag{S87}$$

As $\mu_R > 0$ and $p_C \in (0,1)$, we have

$$\frac{\partial J_R^*}{\partial p_C} < 0. \tag{S88}$$

Therefore only $\frac{\partial J_R^*}{\partial p_C} = -\frac{2N^2 k(k-1)cp_C}{\omega(k-2)(1-p_C)}$ holds. By substituting this equation into Eq. (S84), we obtain the optimal rewarding protocol as

$$\mu_R^* = 2c. \tag{S89}$$

With the optimal rewarding protocol μ_R^*, the dynamical equation thus becomes

$$\frac{dp_C}{dt} = \frac{\omega k(k-2)c}{k-1} p_C(1-p_C), \tag{S90}$$

where the initial fraction of cooperators in the population is denoted by $p_C(0) = p_0$. Solving the above equation, we have

$$p_C = \frac{1}{1 + \frac{1-p_0}{p_0} e^{-\beta_{\text{BD}} t}}, \tag{S91}$$

where $\beta_{\text{BD}} = \frac{\omega k(k-2)c}{k-1}$. Hence, the cumulative cost produced by the optimal rewarding protocol is given by

$$J_R^* = \frac{(kN\mu_R^*)^2}{2\beta_{\text{BD}}}[p_0 - 1 + \delta + \ln(\frac{1-p_0}{\delta})]. \tag{S92}$$

If we solve the optimal control problem for punishment, we obtain for the optimal protocol of negative incentive

$$\mu_P^* = 2c, \tag{S93}$$

and

$$p_C = \frac{1}{1 + \frac{1-p_0}{p_0} e^{-\beta_{\text{BD}} t}}. \tag{S94}$$

Accordingly, the cumulative cost required by the optimal punishing protocol is

$$J_P^* = \frac{(kN\mu_P^*)^2}{2\beta_{\text{BD}}}[p_0 - 1 + \delta + \ln(\frac{1-\delta}{p_0})]. \tag{S95}$$

Therefore the cumulative cost difference of optimal rewarding and punishing protocols is

$$J_R^* - J_P^* = \frac{(kN\mu_v^*)^2}{2\beta_{\text{BD}}}\ln[\frac{p_0(1-p_0)}{\delta(1-\delta)}], \tag{S96}$$

where $\mu_v^* = \mu_R^* = \mu_P^*$. Similarly to the analysis of Eq. (S52) in subsection 1.3, we also find that $J_R^* > J_P^*$ when $\delta < p_0$ and $J_R^* < J_P^*$ when $\delta > p_0$. This implies that for BD updating executing the optimal punishing protocol can induce a lower cumulative cost in comparison with the optimal rewarding one when $\delta < p_0$, and this conclusion is reversed when $\delta > p_0$, which has also been confirmed by numerical calculations and Monte Carlo simulations as presented in figure 4 and figure S5, respectively.

3. IM Updating

3.1. Positive Incentive

For IM updating [1, 2], a focal individual is randomly chosen to update its strategy who has k_C cooperators and k_D defectors among its k neighbors. If the focal individual adopts strategy D, then the fitness of a C–neighbor is

$$f_C = 1 - \omega + \omega\{(b - c + \mu_R)(k-1)q_{C|C} + (\mu_R - c)[(k-1)q_{D|C} + 1]\}, \tag{S97}$$

and the fitness of a D–neighbor is

$$f_D = 1 - \omega + \omega\{b(k-1)q_{C|D} + 0 \cdot [(k-1)q_{D|D} + 1]\}. \tag{S98}$$

Besides, the fitness of the focal individual is

$$f_0 = 1 - \omega + \omega b k_C. \tag{S99}$$

Since the focal individual can keep its own strategy or imitate a neighbor's strategy with probability proportional to the fitness, the probability that the focal individual adopts strategy C is given by

$$\Theta = \frac{k_C f_C}{k_C f_C + k_D f_D + f_0}. \tag{S100}$$

Therefore, p_C increases by $1/N$ with probability

$$P(\Delta p_C = \frac{1}{N}) = p_D \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} \Theta. \tag{S101}$$

Consequently, the number of CC–pairs increases by k_C and hence p_{CC} increases by $k_C/(kN/2)$ with probability

$$P(\Delta p_{CC} = \frac{2k_C}{kN}) = p_D \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} \Theta. \tag{S102}$$

In the alternative case, the randomly selected focal individual adopts strategy C. Here the fitness of a C–neighbor is

$$f_C = 1 - \omega + \omega\{(b - c + \mu_R)[(k-1)q_{C|C} + 1] + (\mu_R - c)(k-1)q_{D|C}\}, \tag{S103}$$

and the fitness of a D–neighbor is

$$f_D = 1 - \omega + \omega[(k-1)q_{C|D} + 1]b. \tag{S104}$$

Besides, the fitness of the focal individual is

$$f_0 = 1 - \omega + \omega[(b - c + \mu_R)k_C + (\mu_R - c)k_D]. \tag{S105}$$

The probability that the focal individual adopts the strategy D is

$$\Upsilon = \frac{k_D f_D}{k_C f_C + k_D f_D + f_0}. \tag{S106}$$

Thus, p_C decreases by $1/N$ with probability

$$P(\Delta p_C = -\frac{1}{N}) = p_C \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|C})^{k_C} (q_{D|C})^{k_D} \Upsilon. \tag{S107}$$

Therefore the number of CC–pairs decreases by k_C and hence p_{CC} decreases by $k_C/(kN/2)$ with probability

$$P(\Delta p_{CC} = -\frac{2k_C}{kN}) = p_C \binom{k}{k_C} (q_{C|C})^{k_C} (q_{D|C})^{k_D} \Upsilon. \tag{S108}$$

Based on these calculations, the time derivative of p_C is given by

$$\begin{aligned} \frac{dp_C}{dt} &= \frac{E(\Delta p_C)}{\Delta t} = \frac{\frac{1}{N}P(\Delta p_C = \frac{1}{N}) - \frac{1}{N}P(\Delta p_C = -\frac{1}{N})}{\frac{1}{N}} \\ &= \frac{\omega k p_{CD}}{(k+1)^2}[2(\mu_R - c - b) + 2(k-1)\xi_1 + (k-1)(\mu_R - c)(q_{C|C} \\ &\quad + q_{D|D}) + (k+1)^2 q_{C|C} + q_{D|D}\xi_1] + o(\omega^2), \end{aligned} \tag{S109}$$

where $\xi_1 = (b - c + \mu_R)q_{C|C} + (\mu_R - c)q_{D|C} - bq_{C|D}$. Accordingly, the time derivative of p_{CC} is given by

$$\begin{aligned} \frac{dp_{CC}}{dt} &= \frac{E(\Delta p_{CC})}{\Delta t} = \frac{\sum_{k_C=0}^{k} \frac{2k_C}{kN} P(\Delta p_{CC} = \frac{2k_C}{kN}) - \sum_{k_C=0}^{k} \frac{2k_C}{kN} P(\Delta p_{CC} = -\frac{2k_C}{kN})}{\frac{1}{N}} \\ &= \frac{2p_{CD}}{k+1}[1 + (k-1)(q_{C|D} - q_{C|C})] + o(\omega). \end{aligned} \tag{S110}$$

Furthermore, we have

$$\frac{dq_{C|C}}{dt} = \frac{2p_{CD}}{(k+1)p_C}[1+(k-1)(q_{C|D}-q_{C|C})]+o(\omega). \tag{S111}$$

Hence, the dynamical equation is described by

$$\begin{cases} \frac{dp_C}{dt} = \omega\Psi_{\text{IM}}^R(p_C, q_{C|C}) + o(\omega^2), \\ \frac{dq_{C|C}}{dt} = \Phi_{\text{IM}}^R(p_C, q_{C|C}) + o(\omega), \end{cases} \tag{S112}$$

where

$$\begin{cases} \Psi_{\text{IM}}^R(p_C, q_{C|C}) = \frac{kp_{CD}}{(k+1)^2}[2(\mu_R - c - b) + 2(k-1)\xi_1 + (k-1)(\mu_R - c)(q_{C|C} \\ \qquad\qquad +q_{D|D}) + (k+1)^2 q_{C|C} + q_{D|D}\xi_1], \\ \Phi_{\text{IM}}^R(p_C, q_{C|C}) = \frac{2p_{CD}}{(k+1)p_C}[1+(k-1)(q_{C|D}-q_{C|C})]. \end{cases}$$

Under weak selection, the velocity of $q_{C|C}$ can be large, and it may rapidly converge to the root defined by $\Phi_{\text{IM}}^R(p_C, q_{C|C}) = 0$ as time $t \to +\infty$. Thus, we get

$$q_{C|C} = p_C + \frac{1}{k-1}(1-p_C). \tag{S113}$$

Accordingly, the dynamical equation described by Eq. (S112) becomes

$$\frac{dp_C}{dt} = \frac{\omega k^2(k-2)[b+(\mu_R-c)(k+2)]}{(k+1)^2(k-1)} p_C(1-p_C) + o(\omega^2), \tag{S114}$$

which has two fixed points $p_C = 0$ and $p_C = 1$. We define the function $F_{\text{IM}}(p_C, \mu_R, t)$ as

$$F_{\text{IM}}(p_C, \mu_R, t) = \frac{\omega k^2(k-2)[b+(\mu_R-c)(k+2)]}{(k+1)^2(k-1)} p_C(1-p_C) + o(\omega^2), \tag{S115}$$

and the derivative of $F_{\text{IM}}(p_C, \mu_R, t)$ with respect to p_C is

$$\frac{dF_{\text{IM}}}{dp_C} = \frac{\omega k^2(k-2)[b+(\mu_R-c)(k+2)]}{(k+1)^2(k-1)}(1-2p_C) + o(\omega^2). \tag{S116}$$

For $\mu_R > c - \frac{b}{k+2}$, we have $\frac{dF_{\text{IM}}}{dp_C}|_{p_C=1} = -\frac{\omega k^2(k-2)[b+(\mu_R-c)(k+2)]}{(k+1)^2(k-1)} < 0$ and $\frac{dF_{\text{IM}}}{dp_C}|_{p_C=0} = \frac{\omega k^2(k-2)[b+(\mu_R-c)(k+2)]}{(k+1)^2(k-1)} > 0$ which implies that the fixed point $p_C = 1$ is stable and $p_C = 0$ is unstable, i.e., cooperators prevail over defectors. Particularly, when $\mu_R = 0$, we can see that for $b/c > k+2$, the fixed point $p_C = 1$ is stable and $p_C = 0$ unstable. Thus, we obtain the condition $b/c > k+2$ for the evolution of cooperation under IM update rule as previously obtained in Refs. [1, 2].

3.2. Negative Incentive

In this subsection, we consider the negative incentive into the networked Prisoner's Dilemma game with IM updating. According to this rule, we randomly choose a focal individual to update its strategy who has k_C cooperators and k_D defectors among its k neighbors. If the focal individual adopts strategy D, then the fitness of a C–neighbor is

$$f_C = 1 - \omega + \omega\{(b-c)(k-1)q_{C|C} - c[(k-1)q_{D|C} + 1]\}, \tag{S117}$$

and the fitness of a D–neighbor is

$$f_D = 1 - \omega + \omega\{(b-\mu_P)(k-1)q_{C|D} - \mu_P[(k-1)q_{D|D} + 1]\}. \tag{S118}$$

Besides, the fitness of the focal individual is

$$f_0 = 1 - \omega + \omega[(b-\mu_P)k_C - \mu_P k_D]. \tag{S119}$$

The probability that the focal individual adopts strategy C is given by the expression Θ in E-q. (S100). Therefore, p_C increases by $1/N$ with probability

$$P(\Delta p_C = \frac{1}{N}) = p_D \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} \Theta. \tag{S120}$$

Furthermore, the number of CC–pairs increases by k_C and hence p_{CC} increases by $k_C/(kN/2)$ with probability

$$P(\Delta p_{CC} = \frac{2k_C}{kN}) = p_D \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} \Theta. \tag{S121}$$

In addition, we consider another case, that is, the randomly selected focal individual adopts strategy C. In this case, the fitness of a C–neighbor is

$$f_C = 1 - \omega + \omega\{(b-c)[(k-1)q_{C|C} + 1] - c(k-1)q_{D|C}\}, \tag{S122}$$

and the fitness of a D–neighbor is

$$f_D = 1 - \omega + \omega\{(b-\mu_P)[(k-1)q_{C|D} + 1] - \mu_P(k-1)q_{D|D}\}. \tag{S123}$$

Besides, the fitness of the focal individual is

$$f_0 = 1 - \omega + \omega[(b-c)k_C - ck_D]. \tag{S124}$$

The probability that the focal individual adopts strategy D is given by the expression Υ in E-q. (S106). Therefore, p_C decreases by $1/N$ with probability

$$P(\Delta p_C = -\frac{1}{N}) = p_C \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|C})^{k_C} (q_{D|C})^{k_D} \Upsilon. \tag{S125}$$

And the number of CC–pairs decreases by k_C and hence p_{CC} decreases by $k_C/(kN/2)$ with probability

$$P(\Delta p_{CC} = -\frac{2k_C}{kN}) = p_C \binom{k}{k_C} (q_{C|C})^{k_C} (q_{D|C})^{k_D} \Upsilon. \tag{S126}$$

Based on these calculations, we obtain the time derivative of p_C given by

$$\begin{aligned} \frac{dp_C}{dt} &= \frac{E(\Delta p_C)}{\Delta t} = \frac{\frac{1}{N}P(\Delta p_C = \frac{1}{N}) - \frac{1}{N}P(\Delta p_C = -\frac{1}{N})}{\frac{1}{N}} \\ &= \frac{\omega k p_{CD}}{(k+1)^2}\{2[(\mu_P - c - b) + (k-1)\xi_2] + (k-1)(\mu_P - c)(q_{C|C} + q_{D|D}) \\ &+ (k+1)^2(q_{C|C} + q_{D|D})\xi_2\} + o(\omega^2), \end{aligned} \tag{S127}$$

where $\xi_2 = (b-c)q_{C|C} - cq_{D|C} + (\mu_P - b)q_{C|D} + \mu_P q_{D|D}$. Accordingly, the time derivative of p_{CC} is given by

$$\begin{aligned} \frac{dp_{CC}}{dt} &= \frac{E(\Delta p_{CC})}{\Delta t} = \frac{\sum_{k_C=0}^{k} \frac{2k_C}{kN} P(\Delta p_{CC} = \frac{2k_C}{kN}) - \sum_{k_C=0}^{k} \frac{2k_C}{kN} P(\Delta p_{CC} = -\frac{2k_C}{kN})}{\frac{1}{N}} \\ &= \frac{2p_{CD}}{k+1}[1 + (k-1)(q_{C|D} - q_{C|C})] + o(\omega). \end{aligned} \tag{S128}$$

Furthermore, we have

$$\frac{dq_{C|C}}{dt} = \frac{2p_{CD}}{(k+1)p_C}[1 + (k-1)(q_{C|D} - q_{C|C})] + o(\omega). \tag{S129}$$

Hence, the dynamical equation is described by

$$\begin{cases} \frac{dp_C}{dt} = \omega \Psi_{\text{IM}}^P(p_C, q_{C|C}) + o(\omega^2), \\ \frac{dq_{C|C}}{dt} = \Phi_{\text{IM}}^P(p_C, q_{C|C}) + o(\omega), \end{cases} \tag{S130}$$

where

$$\begin{cases} \Psi_{\text{IM}}^P(p_C, q_{C|C}) = \frac{kp_{CD}}{(k+1)^2}\{2[(\mu_P - c - b) + (k-1)\xi_2] + (k-1)(\mu_P - c)(q_{C|C} + q_{D|D}) \\ \qquad\qquad\qquad + (k+1)^2(q_{C|C} + q_{D|D})\xi_2\}, \\ \Phi_{\text{IM}}^P(p_C, q_{C|C}) = \frac{2p_{CD}}{(k+1)p_C}[1 + (k-1)(q_{C|D} - q_{C|C})]. \end{cases}$$

Under weak selection, the velocity of $q_{C|C}$ can be large, and it may rapidly converge to the root defined by $\Phi_{\text{IM}}^{P}(p_C, q_{C|C}) = 0$ as time $t \to +\infty$. Thus, we get

$$q_{C|C} = p_C + \frac{1}{k-1}(1-p_C). \tag{S131}$$

Accordingly, the dynamical equation described by Eq. (S130) thus becomes

$$\frac{dp_C}{dt} = \frac{\omega k^2(k-2)[b+(\mu_P-c)(k+2)]}{(k+1)^2(k-1)} p_C(1-p_C) + o(\omega^2), \tag{S132}$$

which has two fixed points $p_C = 0$ and $p_C = 1$. We define the function $F_{\text{IM}}(p_C, \mu_P, t)$ as

$$F_{\text{IM}}(p_C, \mu_P, t) = \frac{\omega k^2(k-2)[b+(\mu_P-c)(k+2)]}{(k+1)^2(k-1)} p_C(1-p_C) + o(\omega^2), \tag{S133}$$

and the derivative of $F_{\text{IM}}(p_C, \mu_P, t)$ with respect to p_C is

$$\frac{dF_{\text{IM}}}{dp_C} = \frac{\omega k^2(k-2)[b+(\mu_P-c)(k+2)]}{(k+1)^2(k-1)} (1-2p_C) + o(\omega^2). \tag{S134}$$

For $\mu_P > c - \frac{b}{k+2}$, we have $\frac{dF_{\text{IM}}}{dp_C}|_{p_C=1} = -\frac{\omega k^2(k-2)[b+(\mu_P-c)(k+2)]}{(k+1)^2(k-1)} < 0$ and $\frac{dF_{\text{IM}}}{dp_C}|_{p_C=0} = \frac{\omega k^2(k-2)[b+(\mu_P-c)(k+2)]}{(k+1)^2(k-1)} > 0$, which implies that the fixed point $p_C = 1$ is stable and $p_C = 0$ unstable, i.e., cooperators prevail over defectors. Particularly, when $\mu_P = 0$, we can see that for $b/c > k+2$, the fixed point $p_C = 1$ is stable and $p_C = 0$ unstable. Thus, we obtain the condition $b/c > k+2$ for the evolution of cooperation under IM update rule as obtained in Refs. [1, 2].

3.3. Optimal Incentive Protocols

In subsections 3.1 and 3.2, we have theoretically obtained the dynamical equation with positive or negative incentive by means of the pair approximation approach in the limit of weak selection, which is given by

$$\frac{dp_C}{dt} = F_{\text{IM}}(p_C, \mu_v, t) = \frac{\omega k^2(k-2)[b+(\mu_v-c)(k+2)]}{(k+1)^2(k-1)} p_C(1-p_C) + o(\omega^2). \tag{S135}$$

This dynamical equation has two equilibria which are $p_C = 0$ and $p_C = 1$. If $\mu_v > c - \frac{b}{k+2}$, the former is unstable and the latter is stable, which means that cooperation will be promoted in the long run. Furthermore, to identify the optimal rewarding and punishing protocols, we now use the approach of HJB equation.

The Hamiltonian function for the control problem is

$$H_{\text{IM}}(p_C, \mu_R, t) = \frac{(kNp_C\mu_R)^2}{2} + \frac{\partial J_R^*}{\partial p_C} F_{\text{IM}}(p_C, \mu_R, t), \tag{S136}$$

where J_R^* is the optimal cost function of p_C and t for the optimal rewarding protocol given as

$$J_R^* = \int_0^{t_f} \frac{(kNp_C\mu_R^*)^2}{2}dt. \tag{S137}$$

Solving $\frac{\partial H_{\text{IM}}}{\partial \mu_R} = 0$, we know that the optimal rewarding protocol μ_R^* should satisfy

$$\mu_R^* = -\frac{\omega(k-2)(k+2)(1-p_C)}{N^2(k-1)(k+1)^2 p_C}\frac{\partial J_R^*}{\partial p_C}. \tag{S138}$$

The HJB equation can be written as

$$-\frac{\partial J_R^*}{\partial t} = H_{\text{IM}}(p_C, \mu_R^*, t). \tag{S139}$$

As the terminal time t_f is not fixed, the optimal cost function $J_R^*(p_C, t)$ is independent of t. Consequently, we have

$$\frac{\partial J_R^*}{\partial t} = 0. \tag{S140}$$

We then obtain

$$\frac{\partial J_R^*}{\partial p_C} = 0 \;\text{ or }\; \frac{\partial J_R^*}{\partial p_C} = \frac{2N^2 p_C[b-c(k+2)](k+1)^2(k-1)}{\omega(k-2)(k+2)^2(1-p_C)}. \tag{S141}$$

As $\mu_R > 0$ and $p_C \in (0,1)$, we have

$$\frac{\partial J_R^*}{\partial p_C} < 0. \tag{S142}$$

From Eq. (S142), we can see that this inequality is obviously satisfied for $b/c \geq k+2$. Instead, we consider the case, i.e., $b/c < k+2$, and hence only $\frac{\partial J_R^*}{\partial p_C} = \frac{2N^2 p_C[b-c(k+2)](k+1)^2(k-1)}{\omega(k-2)(k+2)^2(1-p_C)}$ holds. By substituting this equation into Eq. (S138), we obtain the optimal rewarding protocol as

$$\mu_R^* = \frac{2[c(k+2)-b]}{k+2}. \tag{S143}$$

With the optimal rewarding protocol μ_R^*, the dynamical equation thus becomes

$$\frac{dp_C}{dt} = \frac{\omega k^2(k-2)[c(k+2)-b]}{(k+1)^2(k-1)}p_C(1-p_C), \tag{S144}$$

where the initial fraction of cooperators in the population is denoted by $p_0 = p_C(0)$. To solve the above equation, we have

$$p_C = \frac{1}{1+\frac{1-p_0}{p_0}e^{-\beta_{\text{IM}}t}}, \tag{S145}$$

where $\beta_{\text{IM}} = \frac{\omega k^2(k-2)[c(k+2)-b)]}{(k+1)^2(k-1)}$. Hence, the cumulative cost produced by the optimal rewarding protocol is given by

$$J_R^* = \frac{(kN\mu_R^*)^2}{2\beta_{\text{IM}}}[p_0 - 1 + \delta + \ln(\frac{1-p_0}{\delta})]. \tag{S146}$$

Then, we solve the optimal control problem for punishing described by Eq. (4.5) in the main text. After calculations, we respectively obtain the optimal punishing protocol μ_P^* and the corresponding solution of p_C given by

$$\mu_P^* = \frac{2[c(k+2)-b]}{k+2}, \tag{S147}$$

and

$$p_C = \frac{1}{1+\frac{1-p_0}{p_0}e^{-\beta_{\mathrm{IM}}t}}. \tag{S148}$$

Accordingly, the cumulative cost produced by the optimal punishing protocol is given by

$$J_P^* = \frac{(kN\mu_P^*)^2}{2\beta_{\mathrm{IM}}}[p_0 - 1 + \delta + \ln(\frac{1-\delta}{p_0})]. \tag{S149}$$

Consequently, the cumulative cost difference between optimal rewarding and punishing protocols is given by

$$J_R^* - J_P^* = \frac{(kN\mu_v^*)^2}{2\beta_{\mathrm{IM}}}\ln[\frac{p_0(1-p_0)}{\delta(1-\delta)}], \tag{S150}$$

where $\mu_v^* = \mu_R^* = \mu_P^*$. Similarly to Eq. (S52), we also find that $J_R^* > J_P^*$ when $\delta < p_)$, but when $\delta > p_0$ we have $J_R^* < J_P^*$. This implies that for IM updating the execution of the optimal punishing protocol requires lower cumulative cost in comparison with the optimal rewarding one for $\delta < p_0$ and this conclusion is reversed for $\delta > p_0$. These theoretical results can be confirmed by numerical calculations and Monte Carlo simulations as presented in figure 4 and figure S5, respectively.

4. PC Updating

4.1. Positive Incentive

For PC updating [1, 2] we randomly choose a focal individual to update its strategy who has k_C cooperators and k_D defectors among its k neighbors. If the focal individual adopts strategy D, then the fitness of the focal individual is

$$f_D = 1-\omega+\omega\pi_0^D = 1-\omega+\omega[bk_C + 0\cdot k_D], \tag{S151}$$

and the fitness of a C–neighbor is

$$f_C = 1-\omega+\omega\pi_C^D = 1-\omega+\omega\{(b-c+\mu_R)(k-1)q_{C|C} + (\mu_R - c)[(k-1)q_{D|C}+1]\}. \tag{S152}$$

where π_0^D represents the payoff of the focal individual, and π_C^D denotes the payoff of a C–neighbor.

Since the focal individual either keeps its current strategy or adopts the strategy of a neighbor with a probability that depends on the payoff difference, i.e., $\pi_C^D - \pi_0^D$, the probability that the focal individual adopts the strategy of a C–neighbor for $\omega \to 0$ is

$$\Lambda = \frac{1}{1 + e^{-\omega(\pi_C^D - \pi_0^D)}} = \frac{1}{2} + \omega \frac{\pi_C^D - \pi_0^D}{4}. \tag{S153}$$

Since $f_C - f_D = \omega(\pi_C^D - \pi_0^D)$ for weak selection, we further have

$$\Lambda = \frac{1}{1 + e^{-\omega(\pi_C^D - \pi_0^D)}} = \frac{1}{1 + e^{-(f_C - f_D)}} = \frac{1}{2} + \frac{f_C - f_D}{4}. \tag{S154}$$

Therefore, p_C increases by $1/N$ with probability

$$P(\Delta p_C = \frac{1}{N}) = p_D \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} \frac{k_C}{k} \Lambda. \tag{S155}$$

Hence the number of CC–pairs increases by $(k-1)q_{C|D} + 1$ and p_{CC} increases by $[(k-1)q_{C|D} + 1]/(kN/2)$ with probability

$$P(\Delta p_{CC} = \frac{(k-1)q_{C|D} + 1}{kN/2}) = p_D \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} \frac{k_C}{k} \Lambda. \tag{S156}$$

In addition, we consider another case where the randomly selected focal individual adopts strategy C. In this case, the fitness of the focal individual is

$$f_C = 1 - \omega + \omega \pi_0^C = 1 - \omega + \omega[(b - c + \mu_R)k_C + (\mu_R - c)k_D], \tag{S157}$$

and the fitness of a D–neighbor is

$$f_D = 1 - \omega + \omega \pi_D^C = 1 - \omega + \omega[(k-1)q_{C|D} + 1]b, \tag{S158}$$

where π_0^C represents the payoff of the focal individual, and π_D^C denotes the payoff of a neighbor with strategy D. The probability that the focal individual adopts the strategy of a D–neighbor for $\omega \to 0$ is

$$\Omega = \frac{1}{1 + e^{-\omega(\pi_D^C - \pi_0^C)}} = \frac{1}{2} + \omega \frac{\pi_D^C - \pi_0^C}{4} = \frac{1}{2} + \frac{f_D - f_C}{4}. \tag{S159}$$

Therefore, p_C decreases by $1/N$ with probability

$$P(\Delta p_C = -\frac{1}{N}) = p_C \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|C})^{k_C} (q_{D|C})^{k_D} \frac{k_D}{k} \Omega. \tag{S160}$$

Therefore the number of CC–pairs decreases by $(k-1)q_{C|C}$ and hence p_{CC} increases by $(k-1)q_{C|C}/(kN/2)$ with probability

$$P(\Delta p_{CC}=-\frac{(k-1)q_{C|C}}{kN/2})=p_C\sum_{k_C=0}^{k}\binom{k}{k_C}(q_{C|C})^{k_C}(q_{D|C})^{k_D}\frac{k_D}{k}\Omega. \tag{S161}$$

Based on these calculations, we respectively obtain the time derivatives of p_C and p_{CC} as

$$\begin{aligned}\frac{dp_C}{dt}&=\frac{E(\Delta p_C)}{\Delta t}=\frac{\frac{1}{N}P(\Delta p_C=\frac{1}{N})-\frac{1}{N}P(\Delta p_C=-\frac{1}{N})}{\frac{1}{N}}\\&=\frac{\omega p_{CD}}{2}\{(\mu_R-c-b)+(k-1)[(b-c+\mu_R)q_{C|C}+(\mu_R-c)q_{D|C}-bq_{C|D}]\}+o(\omega^2).\end{aligned} \tag{S162}$$

and

$$\begin{aligned}\frac{dp_{CC}}{dt}&=\frac{E(\Delta p_{CC})}{\Delta t}=\frac{\frac{(k-1)q_{C|D}+1}{kN/2}P(\Delta p_{CC}=\frac{(k-1)q_{C|D}+1}{kN/2})-\frac{(k-1)q_{C|C}}{kN/2}P(\Delta p_{CC}=-\frac{(k-1)q_{C|C}}{kN/2})}{\frac{1}{N}}\\&=\frac{1}{k}p_{CD}[1+(k-1)(q_{C|D}-q_{C|C})]+o(\omega).\end{aligned} \tag{S163}$$

Furthermore, we have

$$\frac{dq_{C|C}}{dt}=\frac{p_{CD}}{kp_C}[1+(k-1)(q_{C|D}-q_{C|C})]+o(\omega). \tag{S164}$$

Hence, the dynamical equation is described by

$$\begin{cases}\frac{dp_C}{dt}=\omega\Psi^R_{\text{PC}}(p_C,q_{C|C})+o(\omega^2),\\ \frac{dq_{C|C}}{dt}=\Phi^R_{\text{PC}}(p_C,q_{C|C})+o(\omega),\end{cases} \tag{S165}$$

where

$$\begin{cases}\Psi^R_{\text{PC}}(p_C,q_{C|C})=\frac{p_{CD}}{2}\{(\mu_R-c-b)+(k-1)[(b-c+\mu_R)q_{C|C}+(\mu_R-c)q_{D|C}-bq_{C|D}]\}\\ \Phi^R_{\text{PC}}(p_C,q_{C|C})=\frac{p_{CD}}{kp_C}[1+(k-1)(q_{C|D}-q_{C|C})].\end{cases}$$

Under weak selection, the velocity of $q_{C|C}$ can be large, and it may rapidly converge to the root defined by $\Phi^R_{\text{PC}}(p_C,q_{C|C})=0$ as time $t\to+\infty$. Thus, we get

$$q_{C|C}=p_C+\frac{1}{k-1}(1-p_C). \tag{S166}$$

Accordingly, the dynamical equation described by Eq. (S165) becomes

$$\frac{dp_C}{dt}=\frac{\omega k(k-2)(\mu_R-c)}{2(k-1)}p_C(1-p_C)+o(\omega^2), \tag{S167}$$

which has two fixed points $p_C = 0$ and $p_C = 1$. We define the function $F_{\text{PC}}(p_C, \mu_R, t)$ as

$$F_{\text{PC}}(p_C, \mu_R, t) = \frac{\omega k(k-2)(\mu_R - c)}{2(k-1)} p_C(1-p_C) + o(\omega^2), \tag{S168}$$

and the derivative of $F_{\text{PC}}(p_C, \mu_R, t)$ with respect to p_C is

$$\frac{dF_{\text{PC}}}{dp_C} = \frac{\omega k(k-2)(\mu_R - c)}{2(k-1)}(1-2p_C) + o(\omega^2). \tag{S169}$$

For $\mu_R > c$, we have $\frac{dF_{\text{PC}}}{dp_C}|_{p_C=1} = -\frac{\omega k(k-2)(\mu_R-c)}{2(k-1)} < 0$ and $\frac{dF_{\text{PC}}}{dp_C}|_{p_C=0} = \frac{\omega k(k-2)(\mu_R-c)}{2(k-1)} > 0$ which implies that the fixed point $p_C = 1$ is stable and $p_C = 0$ is unstable, i.e., cooperators prevail over defectors. Particularly, when $\mu_R = 0$, we can see that the fixed point $p_C = 0$ is always stable and $p_C = 1$ unstable, which means that cooperation can never emerge as observed in previous work [2].

4.2. Negative Incentive

In this subsection, we consider how punishment works for PC updating. According to this rule, we randomly select a focal individual to update its strategy who has k_C cooperators and k_D defectors among its k neighbors. If the focal individual adopts strategy D, then the fitness of the focal individual is

$$f_D = 1 - \omega + \omega\pi_0^D = 1 - \omega + \omega[(b - \mu_P)k_C - \mu_P k_D], \tag{S170}$$

and the fitness of a C–neighbor is

$$f_C = 1 - \omega + \omega\pi_C^D = 1 - \omega + \omega\{(b-c)(k-1)q_{C|C} - c[(k-1)q_{D|C} + 1]\}, \tag{S171}$$

where π_0^D represents the payoff of the focal individual, and π_C^D denotes the payoff of a C–neighbor.

The probability that the focal individual adopts the strategy of a C-neighbor is given by the expression Λ in Eq. (S154). Therefore, p_C increases by $1/N$ with probability

$$P(\Delta p_C = \frac{1}{N}) = p_D \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} \frac{k_C}{k} \Lambda, \tag{S172}$$

and the number of CC–pairs increases by $(k-1)q_{C|D} + 1$ and hence p_{CC} increases by $[(k-1)q_{C|D} + 1]/(kN/2)$ with probability

$$P(\Delta p_{CC} = \frac{(k-1)q_{C|D} + 1}{kN/2}) = p_D \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} \frac{k_C}{k} \Lambda. \tag{S173}$$

In the alternative case, the randomly selected focal individual adopts strategy C. Here the fitness of the focal individual is

$$f_C = 1 - \omega + \omega\pi_0^C = 1 - \omega + \omega[(b-c)k_C - ck_D], \tag{S174}$$

and the fitness of a D–neighbor is

$$f_D = 1 - \omega + \omega\pi_D^C = 1 - \omega + \omega\{(b-\mu_P)[(k-1)q_{C|D} + 1] - \mu_P(k-1)q_{D|D}\}, \tag{S175}$$

where π_0^C represents the payoff of the focal individual, and π_D^C denotes the payoff of a neighbor with strategy D. The probability that the focal individual adopts the strategy of a D–neighbor is given by the expression Ω in Eq. (S159). Therefore, p_C decreases by $1/N$ with probability

$$P(\Delta p_C = -\frac{1}{N}) = p_C \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|C})^{k_C} (q_{D|C})^{k_D} \frac{k_D}{k} \Omega. \tag{S176}$$

And the number of CC–pairs decreases by $(k-1)q_{C|C}$ and hence p_{CC} increases by $(k-1)q_{C|C}/(kN/2)$ with probability

$$P(\Delta p_{CC} = -\frac{(k-1)q_{C|C}}{kN/2}) = p_C \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|C})^{k_C} (q_{D|C})^{k_D} \frac{k_D}{k} \Omega. \tag{S177}$$

Based on these calculations, we respectively obtain the time derivatives of p_C and p_{CC} as

$$\begin{aligned} \frac{dp_C}{dt} &= \frac{E(\Delta p_C)}{\Delta t} = \frac{\frac{1}{N}P(\Delta p_C = \frac{1}{N}) - \frac{1}{N}P(\Delta p_C = -\frac{1}{N})}{\frac{1}{N}} \\ &= \frac{\omega p_{CD}}{2}\{\mu_P - c - b + (k-1)[(b-c)q_{C|C} - cq_{D|C} + (\mu_P - b)q_{C|D} + \mu_P q_{D|D}]\} + o(\omega^2). \end{aligned} \tag{S178}$$

and

$$\begin{aligned} \frac{dp_{CC}}{dt} &= \frac{E(\Delta p_{CC})}{\Delta t} = \frac{\frac{(k-1)q_{C|D}+1}{kN/2}P(\Delta p_{CC} = \frac{(k-1)q_{C|D}+1}{kN/2}) - \frac{(k-1)q_{C|C}}{kN/2}P(\Delta p_{CC} = -\frac{(k-1)q_{C|C}}{kN/2})}{\frac{1}{N}} \\ &= \frac{p_{CD}}{k}[1 + (k-1)(q_{C|D} - q_{C|C})] + o(\omega). \end{aligned} \tag{S179}$$

Furthermore, we have

$$\frac{dq_{C|C}}{dt} = \frac{p_{CD}}{kp_C}[1 + (k-1)(q_{C|D} - q_{C|C})] + o(\omega). \tag{S180}$$

Hence, the dynamical equation is described by

$$\begin{cases} \frac{dp_C}{dt} = \omega\Psi_{\text{PC}}^P(p_C, q_{C|C}) + o(\omega^2), \\ \frac{dq_{C|C}}{dt} = \Phi_{\text{PC}}^P(p_C, q_{C|C}) + o(\omega), \end{cases} \tag{S181}$$

where

$$\begin{cases} \Psi_{\text{PC}}^{P}(p_C, q_{C|C}) = \frac{p_{CD}}{2}\{\mu_P - c - b + (k-1)[(b-c)q_{C|C} - cq_{D|C} + (\mu_P - b)q_{C|D} + \mu_P q_{D|D}]\}, \\ \Phi_{\text{PC}}^{P}(p_C, q_{C|C}) = \frac{p_{CD}}{kp_C}[1 + (k-1)(q_{C|D} - q_{C|C})]. \end{cases}$$

Under weak selection, the velocity of $q_{C|C}$ can be large, and it may rapidly converge to the root defined by $\Phi_{\text{PC}}^{P}(p_C, q_{C|C}) = 0$ as time $t \to +\infty$. Thus, we get

$$q_{C|C} = p_C + \frac{1}{k-1}(1 - p_C). \tag{S182}$$

Accordingly, the dynamical equation described by Eq. (S181) thus becomes

$$\frac{dp_C}{dt} = \frac{\omega k(k-2)(\mu_P - c)}{2(k-1)} p_C(1 - p_C) + o(\omega^2), \tag{S183}$$

which has two fixed points $p_C = 0$ and $p_C = 1$. We define the function $F_{\text{PC}}(p_C, \mu_P, t)$ as

$$F_{\text{PC}}(p_C, \mu_P, t) = \frac{\omega k(k-2)(\mu_P - c)}{2(k-1)} p_C(1 - p_C) + o(\omega^2), \tag{S184}$$

and the derivative of $F_{\text{PC}}(p_C, \mu_P, t)$ with respect to p_C is

$$\frac{dF_{\text{PC}}}{dp_C} = \frac{\omega k(k-2)(\mu_P - c)}{2(k-1)}(1 - 2p_C) + o(\omega^2). \tag{S185}$$

For $\mu_P > c$, we have $\frac{dF_{\text{PC}}}{dp_C}|_{p_C=1} = -\frac{\omega k(k-2)(\mu_P - c)}{2(k-1)} < 0$ and $\frac{dF_{\text{PC}}}{dp_C}|_{p_C=0} = \frac{\omega k(k-2)(\mu_P - c)}{2(k-1)} > 0$ which implies that the fixed point $p_C = 1$ is stable and $p_C = 0$ unstable, i.e., cooperators prevail over defectors. Particularly, when $\mu_P = 0$, we can see that the fixed point $p_C = 0$ is always stable and $p_C = 1$ unstable, which means that cooperation can never emerge as obtained in Ref. [2].

4.3. Optimal Incentive Protocols

By means of the pair approximation approach, in the weak selection limit we have the dynamical equation under PC update rule as

$$\frac{dp_C}{dt} = F_{\text{PC}}(p_C, \mu_v, t) = \frac{\omega k(k-2)(\mu_v - c)}{2(k-1)} p_C(1 - p_C) + o(\omega^2). \tag{S186}$$

This dynamical equation has two equilibria which are $p_C = 0$ and $p_C = 1$. If $\mu_v > c$, the former is unstable and the latter is stable, and hence cooperation will be promoted in the long run. Furthermore, to explore the optimal rewarding and punishing protocols, we now use the approach of HJB equation.

In case of reward, we define the Hamiltonian function $H_{\text{PC}}(p_C, \mu_R, t)$ as

$$H_{\text{PC}}(p_C, \mu_R, t) = \frac{(kNp_C\mu_R)^2}{2} + \frac{\partial J_R^*}{\partial p_C} F_{\text{PC}}(p_C, \mu_R, t), \tag{S187}$$

where J_R^* is the optimal cost function of p_C and t for the optimal rewarding protocol given as

$$J_R^* = \int_0^{t_f} \frac{(kNp_C\mu_R^*)^2}{2} dt. \tag{S188}$$

Solving $\frac{\partial H_{\text{PC}}}{\partial \mu_R} = 0$, we know that the optimal rewarding protocol μ_R^* should satisfy

$$\mu_R^* = -\frac{\omega(k-2)(1-p_C)}{2N^2k(k-1)p_C} \frac{\partial J_R^*}{\partial p_C}. \tag{S189}$$

The HJB equation can be written as

$$-\frac{\partial J_R^*}{\partial t} = H_{\text{PC}}(p_C, \mu_R^*, t). \tag{S190}$$

As the terminal time t_f is not fixed, the optimal cost function $J_R^*(p_C, t)$ is independent of t. Consequently, we have

$$\frac{\partial J_R^*}{\partial t} = 0. \tag{S191}$$

We then yield

$$\frac{\partial J_R^*}{\partial p_C} = 0 \ \text{ or } \ \frac{\partial J_R^*}{\partial p_C} = -\frac{4N^2k(k-1)cp_C}{\omega(k-2)(1-p_C)}. \tag{S192}$$

As $\mu_R > 0$ and $p_C \in (0, 1)$, we have

$$\frac{\partial J_R^*}{\partial p_C} < 0. \tag{S193}$$

Therefore only $\frac{\partial J_R^*}{\partial p_C} = -\frac{4N^2k(k-1)cp_C}{\omega(k-2)(1-p_C)}$ holds. By substituting this equation into Eq. (S189), we obtain the optimal rewarding protocol μ_R^* as

$$\mu_R^* = 2c. \tag{S194}$$

With the optimal rewarding protocol μ_R^*, the dynamical equation thus becomes

$$\frac{dp_C}{dt} = \frac{\omega k(k-2)c}{2(k-1)} p_C(1-p_C), \tag{S195}$$

where the initial fraction of cooperators in the population is denoted by $p_0 = p_C(0)$. By solving this equation, we have

$$p_C = \frac{1}{1 + \frac{1-p_0}{p_0} e^{-\beta_{\text{PC}} t}}, \tag{S196}$$

where $\beta_{\text{PC}} = \frac{\omega k(k-2)c}{2(k-1)}$. Hence, the cumulative cost produced by the optimal rewarding protocol is given by

$$J_R^* = \frac{(kN\mu_R^*)^2}{2\beta_{\text{PC}}}[p_0 - 1 + \delta + \ln(\frac{1-p_0}{\delta})]. \tag{S197}$$

For punishment, we respectively obtain the optimal protocol μ_P^* and the corresponding solution of p_C as

$$\mu_P^* = 2c \tag{S198}$$

and

$$p_C = \frac{1}{1 + \frac{1-p_0}{p_0} e^{-\beta_{\text{PC}} t}}. \tag{S199}$$

Accordingly, the cumulative cost produced by the optimal punishing protocol is given by

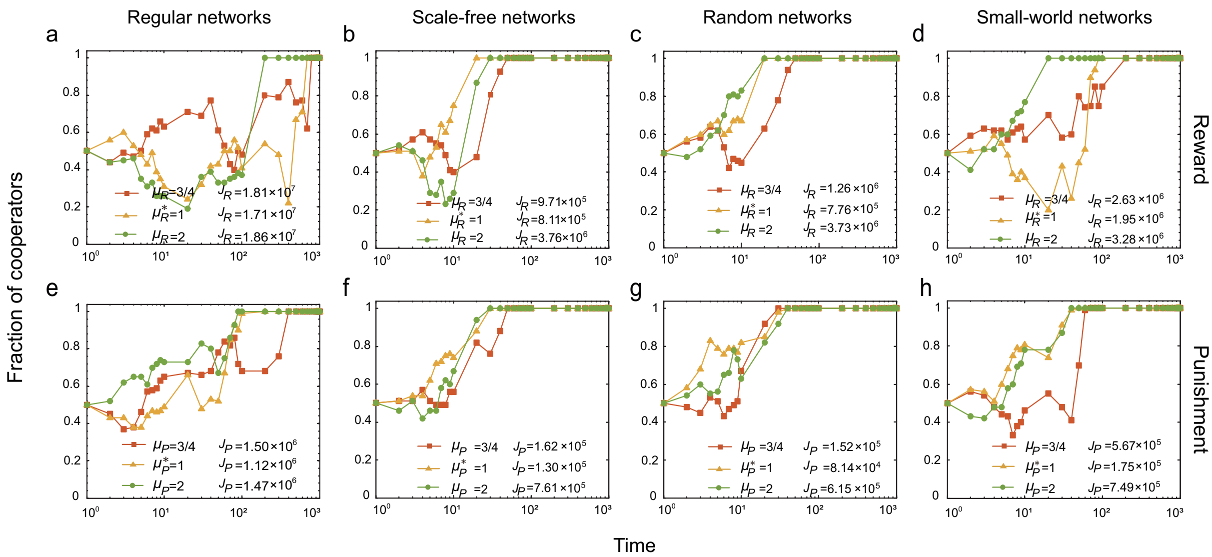
$$J_P^* = \frac{(kN\mu_P^*)^2}{2\beta_{\text{PC}}}[p_0 - 1 + \delta + \ln(\frac{1-\delta}{p_0})]. \tag{S200}$$

Consequently, the difference between the cumulative cost values is

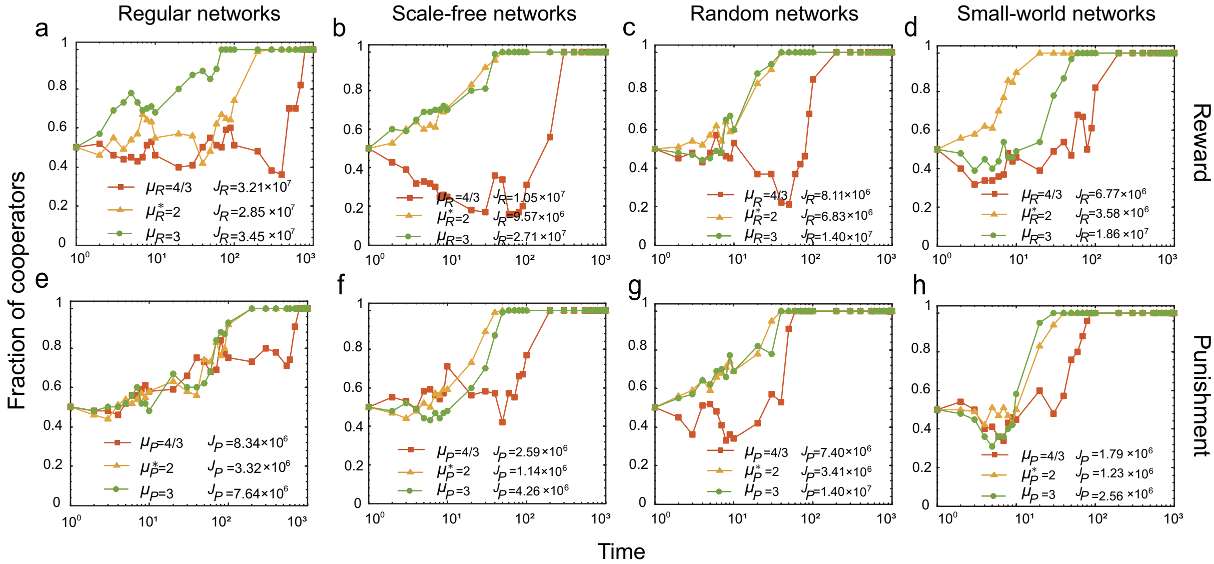
$$J_R^* - J_P^* = \frac{(kN\mu_v^*)^2}{2\beta_{\text{PC}}} \ln[\frac{p_0(1-p_0)}{\delta(1-\delta)}], \tag{S201}$$

where $\mu_v^* = \mu_R^* = \mu_P^*$. Similarly to Eq. (S52), we also find that $J_R^* > J_P^*$ when $\delta < p_0$, but when $\delta > p_0$ we have $J_R^* < J_P^*$. This implies that for PC updating the usage of optimal punishment requires less cost than the optimal rewarding protocol for $\delta < p_0$ and we have the opposite conclusion for $\delta > p_0$. These theoretical results can be confirmed by numerical calculations and Monte Carlo simulations as presented in figure 4 and figure S5, respectively.

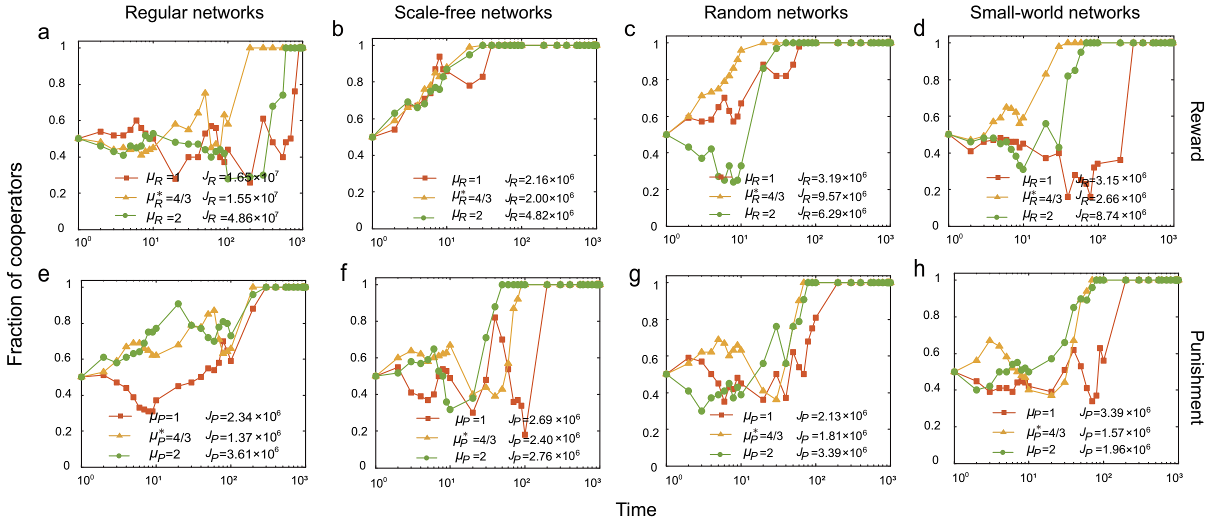
Supplementary Figures



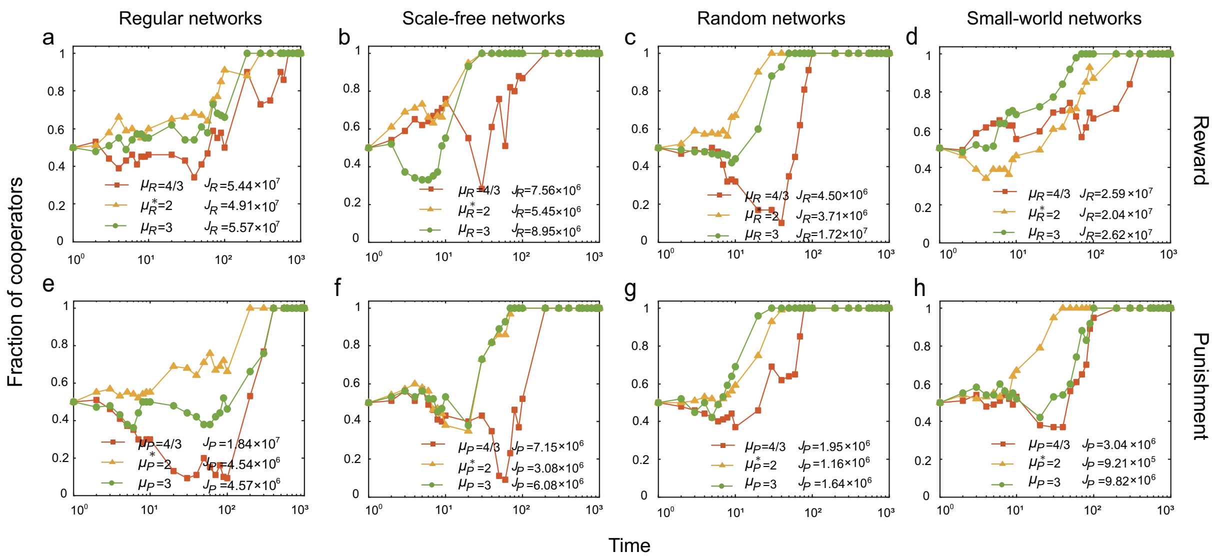
Supplementary Figure 1: **Time evolution of the fraction of cooperators for three different protocols of incentives on four different networks under DB updating.** The applied protocols are marked by the legend, where the optimal one is indicated by $*$. We have also plotted the cumulative cost values for each incentive protocol. The results of Monte Carlo simulations for reward (punishment) are shown on top (bottom) row. Parameters: $N = 100$, $L = 10$, $b = 2$, $c = 1$, $\delta = 0.01$, $\omega = 0.01$, and $p_0 = 0.5$. For proper comparison the average degree is set to 4 for all graphs.



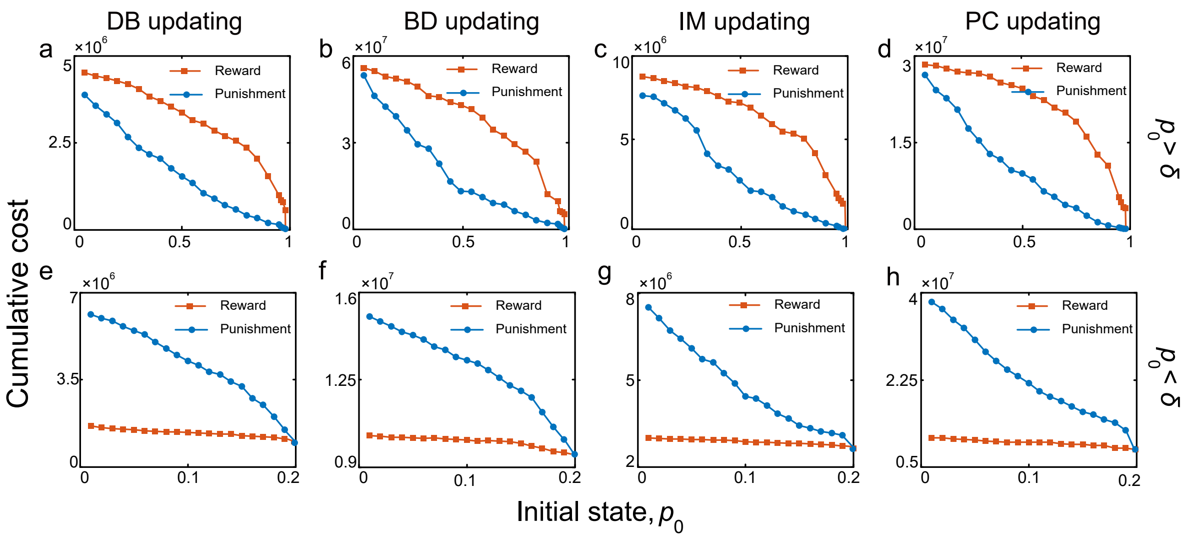
Supplementary Figure 2: **Time evolution of the fraction of cooperators for three different protocols of incentives on four different networks under BD updating.** The applied protocols are marked by the legend, where the optimal one is indicated by $*$. We have also plotted the cumulative cost values for each incentive protocol. The results of Monte Carlo simulations for reward (punishment) are shown on top (bottom) row. Other parameter values are the same as those in figure S1.



Supplementary Figure 3: **Time evolution of the fraction of cooperators for three different protocols of incentives on four different networks under IM updating.** The applied protocols are marked by the legend, where the optimal one is indicated by $*$. We have also plotted the cumulative cost values for each incentive protocol. The results of Monte Carlo simulations for reward (punishment) are shown on top (bottom) row. Other parameter values are the same as those in figure S1.



Supplementary Figure 4: **Time evolution of the fraction of cooperators for three different protocols of incentives on four different networks under PC updating.** The applied protocols are marked by the legend, where the optimal one is indicated by $*$. We have also plotted the cumulative cost values for each incentive protocol. The results of Monte Carlo simulations for reward (punishment) are shown on top (bottom) row. Other parameter values are the same as those in figure S1.



Supplementary Figure 5: Cumulative cost needed for reaching the expected terminal state in dependence of the p_0 initial portion of cooperators for the optimal rewarding and punishing protocols. Each column of panels represents a strategy update rule as indicated. Top row represents the results of Monte Carlo simulations by averaging over 200 independent simulation runs on regular networks with degree k in the condition of $p_0 > \delta = 0.01$, while bottom row represents the results obtained from Monte carlo simulations by averaging over 200 independent simulation runs on regular networks with degree k in the condition of $p_0 < \delta = 0.2$. Other parameters: $N = 100$, $L = 10$, $b = 2$, $c = 1$, $\omega = 0.01$, and $k = 4$.

Supplementary References

[1] Ohtsuki H, Hauert C, Lieberman E, Nowak M A. 2006 A simple rule for the evolution of cooperation on graphs and social networks. *Nature* **441**, 502–505. (doi: 10.1038/nature04605)

[2] Ohtsuki H, Nowak M A. 2006 The replicator equation on graphs. *J. Theor. Biol.* **243**, 86–97. (doi: 10.1016/j.jtbi.2006.06.004)

[3] Evans L C. 2005 *An introduction to mathematical optimal control theory.* Berkeley, USA: University of California Press.

[4] Geering H P. 2007 *Optimal control with engineering applications.* Berlin, Germany: Springer.

[5] Lenhart S, Workman J T. 2007 *Optimal control applied to biological models.* Boca Raton, USA: Chapman and Hall/CRC.