Journal of Theoretical Biology 510 (2021) 110543

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

The effect of the opting-out strategy on conditions for selection to favor the evolution of cooperation in a finite population

Cong Li, Sabin Lessard *

Département de mathématiques et de statistique, Université de Montréal, Montréal, Québec, Canada

ARTICLE INFO

Article history:
Received 30 June 2020
Revised 11 November 2020
Accepted 12 November 2020
Available online 19 November 2020

Keywords:
Diffusion approximation
Fixation probability
Moran model
Prisoner's Dilemma

ABSTRACT

We consider a Prisoner's Dilemma (PD) that is repeated with some probability $1-\rho$ only between cooperators as a result of an opting-out strategy adopted by all individuals. The population is made of N pairs of individuals and is updated at every time step by a birth–death event according to a Moran model. Assuming an intensity of selection of order $1/N$ and taking $2N^2$ birth–death events as unit of time, a diffusion approximation exhibiting two time scales, a fast one for pair frequencies and a slow one for cooperation (C) and defection (D) frequencies, is ascertained in the limit of a large population size. This diffusion approximation is applied to an additive PD game, cooperation by an individual incurring a cost c to the individual but providing a benefit b to the opponent. This is used to obtain the probability of ultimate fixation of C introduced as a single mutant in an all D population under selection, which can be compared to the probability under neutrality, $1/(2N)$, as well as the corresponding probability for a single D introduced in an all C population under selection. This gives conditions for cooperation to be favored by selection. We show that these conditions are satisfied when the benefit-to-cost ratio, b/c, exceeds some increasing function of ρ that is approximately given by $(1+\sqrt{\rho})/(1-\sqrt{\rho})$. This condition is more stringent, however, than the condition for tit-for-tat (*TFT*) to be favored against always-defect (*AllD*) in the absence of opting-out.

1. Introduction

In a two-player two-strategy game known as the Prisoner's Dilemma (PD), in which cooperation and defection, denoted by C and D, respectively, are used by individuals in pairwise interactions, the payoffs are given by the entries of a 2×2 matrix

$$\begin{pmatrix} \pi_{CC} & \pi_{CD} \\ \pi_{DC} & \pi_{DD} \end{pmatrix} \tag{1}$$

Here, π_{ij} represents the payoff to an individual using strategy i against an individual using strategy j where i, j belong to the set of strategies $\{C, D\}$. In a PD game, cooperation against cooperation pays more than defection against defection, but less than defection against cooperation, while cooperation against defection pays the least. Thus, the entries of the payoff matrix satisfy the inequalities $\pi_{DC} > \pi_{CC} > \pi_{DD} > \pi_{CD}$ (Poundstone, 1992). In particular, if cooperation and defection have additive effects on the payoff with cooperation by an individual incurring a cost c to the individual but providing a benefit b to the opponent, then we have $\pi_{CC} = b - c$, $\pi_{CD} = -c$, $\pi_{DC} = b$ and $\pi_{DD} = 0$. We call this case the additive PD game.

In a one-round PD game with defection paying more than cooperation against both defection and cooperation, defection is the only rational choice and the only Nash equilibrium (NE) (see, e.g., Hofbauer and Sigmund, 1998; Nowak, 2006a). In a repeated PD game, however, with a fixed positive probability of repeating the interaction between the same players from one round to the next that does not depend on the strategies in use, the tit-for-tat (*TFT*) strategy starting with cooperation becomes a Nash equilibrium against the always-defect (*AllD*) strategy (actually against any other strategy) if the expected number of rounds of the game is large enough (Axelrod and Hamilton, 1981; Axelrod, 1984). This can be seen as an effect of direct reciprocity (Trivers, 1971), since *TFT* against *TFT* leads to reciprocal cooperation and *TFT* or *AllD* against *AllD* to reciprocal defection at least after the first round.

In the framework of a repeated PD game, many other strategies than *TFT* and *AllD* can be used and be successful (Bendor and Swistak, 1995; Sandholm, 2010; van Veelen, 2012b; García and van Veelen, 2016). This is even more so if every round of the game,

* Corresponding author.
E-mail address: lessards@dms.umontreal.ca (S. Lessard).

https://doi.org/10.1016/j.jtbi.2020.110543
0022-5193/

FEEDBACK

not only the players can choose to cooperate or to defect, but also to stay or to leave for the next round contingently on the past. In this case, a rational strategy is to repeat the interaction with a cooperator as long as possible, but to end it with a defector as soon as possible. This is known as the opting-out or out-for-tat (*OFT*) strategy (Hayashi, 1993; Schuessler, 1989; Aktipis, 2004; Fujiwara-Greve and Okuno-Fujiwara, 2009; Izquierdo et al., 2010, 2014; Fujiwara-Greve et al., 2015). As a result, an interaction between two cooperators may be continued with some probability, while an interaction between two defectors or between one defector and one cooperator may never be repeated. This can be seen as a mechanism that creates direct reciprocity by which the evolution and maintenance of cooperation can be favored by selection (Zhang et al., 2016, 2017; Kurokawa, 2019).

The opting-out strategy is akin to assortment of cooperative acts in social space, where cooperative behaviour is a repeatable trait of individuals, and cooperative individuals associate and interact with each other disproportionately more than with defectors (see, e.g., Eshel and Cavalli-Sforza, 1982; van Veelen et al., 2012a). There is some evidence that individuals from a range of species show stability in their level of cooperativeness (Bergmüller et al., 2010) and that animal social network structures may show significant within-population heterogeneity in social tie strengths (Krause et al., 2015). Recent empirical investigations in wild Trinidadian guppies (*Poecilia reticulata*) support the hypothesis that assortment by repeatable cooperativeness may be an important feature for the evolution and persistence of non-kin cooperation in real-world populations (Brask et al., 2019; Croft et al., 2015).

In this paper, we will consider a Prisoner's Dilemma (PD) with cooperation (C) or defection (D) as only possible strategies in a single interaction between two individuals at any given time step. Moreover, as a result of an opting-out strategy adopted by all individuals, a pairwise interaction will certainly not be repeated from one time step to the next unless the two interacting individuals are both cooperators, in which case the interaction will be repeated with some fixed probability $1-\rho$. The population will be assumed to be made of N pairs of individuals and be updated at every time step by a birth–death event according to a Moran model. With appropriate scalings of the intensity of selection and time with respect to the population size, we will establish a diffusion approximation in the limit of a large population that shows that two time scales come into play, a fast one for pair frequencies and a slow one for C and D frequencies. This diffusion approximation will be applied to an additive PD game, cooperation by an individual incurring a cost c to the individual but providing a benefit b to the interacting partner, and used to obtain the probability of ultimate fixation of C introduced as a single mutant in an all D population under selection. This probability will be compared to the probability under neutrality, which is $1/(2N)$, as well as the corresponding probability for a single D introduced in an all C population under selection. This will provide conditions for cooperation to be favored by selection.

2. The model

Consider a population of N pairs of interacting individuals in which each individual is either a cooperator, C, or a defector, D. The population state and its changes from time t to time $t+\Delta t$, a time interval of length $\Delta t = 1/\left(2N^2\right)$, are represented in Fig. 1.

At time t, which corresponds to the end of one round of the PD game followed by updating, the number of CC pairs in the population is NP_{CC}, while the number of CD pairs is NP_{CD} and the number of DD pairs NP_{DD}. Then, $x = P_{CC} + P_{CD}/2$ is the frequency of C in the population, and $1-x$ is the frequency of D. Note that pairs are not ordered so that a CD pair means a pair made of one C and one D.

Suppose that all individuals in the population adopt the opting-out strategy so that only the individuals paired with a C partner are interested in continuing their interaction with the same partner in the time interval $[t, t+\Delta t]$. As a result, all CD pairs and DD pairs break apart, while each CC pair breaks apart with some probability ρ and, therefore, stays unbroken with probability $1-\rho$. The parameter ρ is assumed to be a positive constant. Then the number of free D individuals is

$$N_D = NP_{CD} + 2NP_{DD} = 2N(1-x), \tag{2}$$

while the number of free C individuals is

$$N_C = NP_{CD} + 2R = 2R + 2N(x - P_{CC}), \tag{3}$$

where R stands for the number of broken CC pairs. This number is a random variable that follows a binomial distribution with parameters NP_{CC} and ρ.

Now assume that all free individuals form new pairs at random. The number of these is

$$\frac{N_C + N_D}{2} = NP_{CD} + NP_{DD} + R, \tag{4}$$

while the conditional expected frequencies of CC, CD and DD among these are

$$\frac{N_C(N_C-1)}{(N_C+N_D)(N_C+N_D-1)}, \quad \frac{2N_CN_D}{(N_C+N_D)(N_C+N_D-1)}, \quad \frac{N_D(N_D-1)}{(N_C+N_D)(N_C+N_D-1)}, \tag{5}$$

respectively. Besides, there are $N_{CC} = NP_{CC} - R$ unbroken CC pairs. All together, we still have N pairs of individuals for the next round of the PD game.

Let the random variables q_{CC}, q_{CD} and q_{DD} represent the frequencies of CC, CD and DD in the set made of all new pairs and all unbroken CC pairs. Note that $q_{CC} + q_{CD}/2 = x$ and $q_{DD} + q_{CD}/2 = 1-x$, which means that the frequencies of C and D in the population are unchanged. On the other hand, the expected values of q_{CC}, q_{CD} and q_{DD} are given by

$$\mathbf{E}(q_{CC}) = 2x - 1 + \frac{(1-x)^2}{1-(1-\rho)P_{CC}} + O\left(N^{-1/2}\right), \tag{6a}$$

$$\mathbf{E}(q_{CD}) = 2(1-x) - \frac{2(1-x)^2}{1-(1-\rho)P_{CC}} + O\left(N^{-1/2}\right), \tag{6b}$$

$$\mathbf{E}(q_{DD}) = \frac{(1-x)^2}{1-(1-\rho)P_{CC}} + O\left(N^{-1/2}\right), \tag{6c}$$

while their variances are all of order N^{-1}, that is,

$$\mathbf{Var}(q_{CC}) = O\left(N^{-1}\right), \tag{7a}$$

$$\mathbf{Var}(q_{CD}) = O\left(N^{-1}\right), \tag{7b}$$

$$\mathbf{Var}(q_{DD}) = O\left(N^{-1}\right). \tag{7c}$$

These results are shown in Appendix A.

The update of the population at the end of the time interval $[t, t+\Delta t]$ is obtained by a birth–death event according to a Moran model in a context of evolutionary game theory (see, e.g., Hofbauer and Sigmund, 1998; Ewens, 2004; Nowak et al., 2004; Ohtsuki et al., 2006). One individual is chosen with probability proportional to fitness to produce an offspring identical to itself and one individual is chosen at random to be replaced by the offspring.

Here, the fitness of an i-strategist in interaction with a j-strategist is given in the form

$$w_{ij} = 1 + s\pi_{ij}, \tag{8}$$

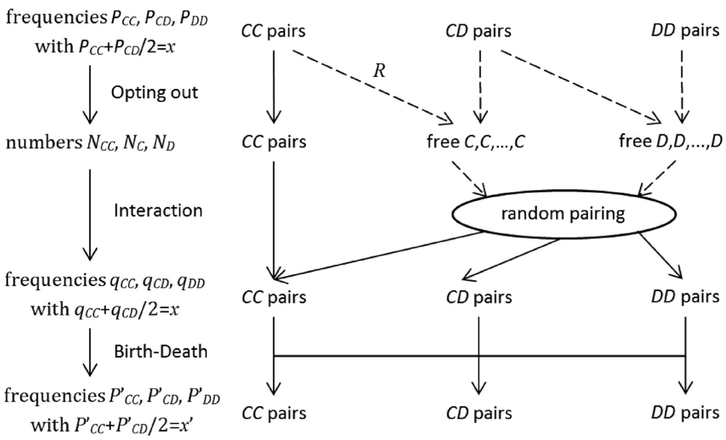

Fig. 1. Changes of the population state in a time interval $[t, t + \Delta t]$.

where 1 stands for a baseline fitness, $s = \sigma N^{-1}$ represents an intensity of selection, and π_{ij} denotes the payoff to i against j for i, j in $\{C, D\}$ as given in (1) for a one-round PD game. Note that CC, CD and DD pairs are then in frequencies q_{CC}, q_{CD} and q_{DD}, respectively, and $w_{DC} \neq w_{CD}$ in CD pairs since $\pi_{DC} \neq \pi_{CD}$. Therefore, the offspring produced is a cooperator with conditional probability

$$\begin{aligned} Pr(C) &= \tfrac{2q_{CC}w_{CC}+q_{CD}w_{CD}}{2q_{CC}w_{CC}+q_{CD}w_{CD}+q_{CD}w_{DC}+2q_{DD}w_{DD}} \\ &= x + \tfrac{\sigma}{N}((1-x)A - xB) + o\left(N^{-1}\right), \end{aligned} \tag{9}$$

and a defector with conditional probability

$$\begin{aligned} Pr(D) &= \tfrac{2q_{DD}w_{DD}+q_{CD}w_{DC}}{2q_{CC}w_{CC}+q_{CD}w_{CD}+q_{CD}w_{DC}+2q_{DD}w_{DD}} \\ &= 1 - x - \tfrac{\sigma}{N}((1-x)A - xB) + o\left(N^{-1}\right), \end{aligned} \tag{10}$$

where

$$A = q_{CC}\pi_{CC} + \frac{q_{CD}}{2}\pi_{CD} \tag{11}$$

and

$$B = q_{DD}\pi_{DD} + \frac{q_{CD}}{2}\pi_{DC}. \tag{12}$$

Note that $Pr(C)$ and $Pr(D)$ are both random variables whose main terms are linear functions of q_{CC}, q_{CD} and q_{DD}.

On the other hand, the offspring produced replaces a cooperator with probability x, and a defector with probability $1 - x$. Actually, it replaces a cooperator in a CC pair with probability q_{CC} or a CD pair with probability $q_{CD}/2$, while it replaces a defector in a DD pair with probability q_{DD} or a CD pair with probability $q_{CD}/2$.

Following the replacement of an individual by the offspring, the frequencies of CC, CD and DD among the N pairs are denoted by P'_{CC}, P'_{CD} and P'_{DD}, respectively, and the frequency of C in these pairs is $P'_{CC} + P'_{CD}/2 = x'$. This gives the population state at the beginning of the next time interval which corresponds to time $t + \Delta t = t + 1/\left(2N^2\right)$ with $2N^2$ time intervals as unit of time. Note that P_{CD} and P_{DD} can be expressed in terms of x and P_{CC}, so that x and P_{CC} can be used to describe the population state.

3. Diffusion approximation

Let $\Delta x = x' - x$ and $\Delta P_{CC} = P'_{CC} - P_{CC}$ be the changes in the frequencies of C and CC, respectively, from time t to time $t + \Delta t$ with $\Delta t = 1/\left(2N^2\right)$. Given these frequencies at time t, the first, second and fourth moments of Δx are approximated as (see Appendix B for details)

$$\mathbf{E}(\Delta x) = \frac{1}{2N^2}m(x, P_{CC}) + o\left(N^{-2}\right), \tag{13}$$

$$\mathbf{E}\left((\Delta x)^2\right) = \frac{1}{2N^2}v(x) + o\left(N^{-2}\right) \tag{14}$$

and

$$\mathbf{E}\left((\Delta x)^4\right) = o\left(N^{-3}\right), \tag{15}$$

respectively, where

$$m(x, P_{CC}) = \sigma\mathbf{E}((1-x)A - xB) \tag{16}$$

and

$$v(x) = x(1-x). \tag{17}$$

Moreover, we have

$$\mathbf{E}(\Delta P_{CC}) = \frac{(x - P_{CC})^2 - \rho P_{CC}(1 - 2x + P_{CC})}{1 - (1-\rho)P_{CC}} + O\left(N^{-1/2}\right) \tag{18}$$

and

$$\mathbf{Var}(\Delta P_{CC}) = O\left(N^{-1}\right) \tag{19}$$

for the mean and variance, respectively, of the change ΔP_{CC}. On the other hand, in an infinite population in the absence of selection, the frequency of C remains constant, while the frequency of CC converges uniformly to an equilibrium value P^*_{CC} in $[0,1]$. This equilibrium value is obtained by solving the equation $\mathbf{E}(\Delta P_{CC}) = 0$, which gives

$$P^*_{CC} = x + \frac{\rho}{2(1-\rho)} - \frac{\sqrt{\rho^2 + 4x(1-x)\rho(1-\rho)}}{2(1-\rho)} \tag{20}$$

for ρ in $[0,1)$, and $P^*_{CC} = x^2$ for $\rho = 1$, where x is the frequency of C (see Appendix C for details). Note that the expression for P^*_{CC} in (20) tends to x^2 as ρ tends to 1 so that we use only this expression in the rest of the paper.

The conditions (13)–(15) and (18)–(20) show that there are two time scales at work in the discrete-time Markov chain for the population state, the variable P_{CC} changing more rapidly than the variable x. Moreover, as $N \to \infty$, these conditions ascertain that the Markov chain converges to a diffusion approximation with $m(x) = m(x, P^*_{CC})$ as drift function, and $v(x) = x(1-x)$ as diffusion function (Ethier and Nagylaki, 1980).

Using (6) with $P_{CC} = P^*_{CC}$ and the equality (see (81) in Appendix C)

$$\frac{(1-x)^2}{1-(1-\rho)P^*_{CC}} = P^*_{CC} - 2x + 1 \tag{21}$$

leads to

$$\mathbf{E}(q_{CC}) = P^*_{CC} + O\left(N^{-1/2}\right), \tag{22a}$$

$$\mathbf{E}(q_{CD}) = 2x - 2P^*_{CC} + O\left(N^{-1/2}\right), \tag{22b}$$

$$\mathbf{E}(q_{DD}) = 1 - 2x + P^*_{CC} + O\left(N^{-1/2}\right). \tag{22c}$$

Let us summarize.

Result 1. *Consider a PD game with payoff matrix (1) for N pairs of individuals so that, as a result of opting-out from one round to the next, all pairs break apart to form new pairs at random but a random proportion of CC pairs whose mean is $1-\rho < 1$. Assume one birth–death event at the end of each round with the probability of giving birth proportional to 1 plus the payoff times σ/N and the probability of dying given by $1/(2N)$. Taking $2N^2$ birth–death events as unit of time and letting $N \to \infty$, the Markov chain of the frequency of C converges to a diffusion with $v(x) = x(1-x)$ as diffusion function and*

$$m(x) = \sigma\big(x(1-x)(\pi_{CC} - \pi_{DD}) - (x - P^*_{CC}) \\ ((1-x)(\pi_{CC} - \pi_{CD}) + x(\pi_{DC} - \pi_{DD}))\big) \tag{23}$$

as drift function, where P^*_{CC} is given by (20).

In the diffusion approximation, it is known (see, e.g., Kimura, 1964; Risken, 1992; Ewens, 2004) that the probability density function of C evaluated at x at time $t > 0$ given an initial value p at time 0, denoted by $f(x,p,t)$, satisfies the forward Kolmogorov (Fokker–Planck) equation

$$\frac{\partial f(x,p,t)}{\partial t} = -\frac{\partial}{\partial x}(m(x)f(x,p,t)) + \frac{\partial^2}{\partial x^2}\left(\frac{v(x)f(x,p,t)}{2}\right), \tag{24}$$

as well as the backward Kolmogorov equation

$$\frac{\partial f(x;p,t)}{\partial t} = m(p)\frac{\partial f(x;p,t)}{\partial p} + \frac{v(p)}{2}\frac{\partial^2 f(x;p,t)}{\partial p^2}. \tag{25}$$

In the case at hand with no mutation, the two boundaries $x = 0$ and $x = 1$ are absorbing states.

Moreover, if $u(p,t)$ denotes the probability that C is fixed by time $t > 0$ so that $x(t) = 1$ given an initial frequency $x(0) = p$, then it is known that this fixation probability satisfies the backward Kolmogorov equation, that is,

$$\frac{\partial u(p,t)}{\partial t} = m(p)\frac{\partial u(p,t)}{\partial p} + \frac{v(p)}{2}\frac{\partial^2 u(p,t)}{\partial p^2}, \tag{26}$$

with the boundary conditions $u(0,t) = 0$ and $u(1,t) = 1$. By letting $t \to \infty$, the limit

$$u(p) = \lim_{t\to\infty} u(p,t) \tag{27}$$

represents the probability of ultimate fixation of C given an initial frequency $x(0) = p$. As $t \to \infty$, the left-hand side in (26) tends to 0 so that we have

$$0 = m(p)\frac{du(p)}{dp} + \frac{v(p)}{2}\frac{d^2u(p)}{dp^2} \tag{28}$$

with the boundary conditions $u(0) = 0$ and $u(1) = 1$. The solution of this ordinary differential equation is known to be (see, e.g., Ewens, 2004)

$$u(p) = \frac{\int_0^p \psi(y)dy}{\int_0^1 \psi(y)dy}, \tag{29}$$

where

$$\psi(y) = \exp\left(-2\int_0^y \frac{m(x)}{v(x)}dx\right). \tag{30}$$

Note that the probability of ultimate fixation of D is given by $1 - u(p)$, since there is ultimate fixation of C or D with probability 1.

4. Additive PD game

Consider an additive Prisoner's Dilemma (PD) where a cooperator pays a fixed cost $c > 0$ while its partner receives a fixed benefit $b > c$. The payoff matrix (1) takes the form

$$\begin{pmatrix} \pi_{CC} & \pi_{CD} \\ \pi_{DC} & \pi_{DD} \end{pmatrix} = \begin{pmatrix} b-c & -c \\ b & 0 \end{pmatrix}. \tag{31}$$

Substituting the above payoffs into $m(x)$ in Result 1 yields

$$\begin{aligned} m(x) &= \sigma\big(x(1-x)(b-c) - b(x - P^*_{CC})\big) \\ &= \sigma\left(x(1-x)(b-c) - \tfrac{bf(x,\rho)}{2(1-\rho)}\right), \end{aligned} \tag{32}$$

where

$$f(x,\rho) = \sqrt{\rho^2 + 4x(1-x)\rho(1-\rho)} - \rho. \tag{33}$$

This function defines a concave surface on the domain $[0,1]\times[0,1]$ with $f(x,\rho) > 0$ inside this domain and $f(x,\rho) = 0$ on its boundary.

Now, let $F_C = u\left((2N)^{-1}\right)$ be the probability of ultimate fixation of C introduced as a single mutant in an all D population of size $2N$. The corresponding fixation probability for a single D introduced in an all C population is $F_D = 1 - u\left(1 - (2N)^{-1}\right)$. The evolution of cooperation is said to be favored by selection if $F_C > (2N)^{-1}$, where $(2N)^{-1}$ is the fixation probability under neutrality. Similarly, the evolution of defection is said to be unfavored by selection if $F_D < (2N)^{-1}$. On the other hand, the evolution of cooperation is said to be more favored by selection than the evolution of defection if $F_C > F_D$. Finally, if the three conditions are simultaneously satisfied, that is, $F_C > (2N)^{-1} > F_D$, then the evolution of cooperation is said to be fully favored by selection (Nowak et al., 2004; Li and Lessard, 2020).

When the population size $2N$ is big enough, the conditions to have $F_C > (2N)^{-1},\ F_D < (2N)^{-1},\ F_C > F_D$ take the form

$$\int_0^1 \psi(y)dy < 1, \tag{34a}$$

$$\int_0^1 \psi(y)dy > \psi(1), \tag{34b}$$

$$\psi(1) < 1, \tag{34c}$$

respectively (Li and Lessard, 2020). Here, we have

$$\psi(y) = \exp\left(-2\int_0^y g(x)dx\right), \tag{35}$$

where

$$\begin{aligned} g(x) = \tfrac{m(x)}{v(x)} &= \sigma c\left((b/c - 1) - \tfrac{(b/c)f(x,\rho)}{2(1-\rho)x(1-x)}\right) \\ &= \sigma c\left(r\left(1 - \tfrac{2}{\sqrt{1+4ax(1-x)}+1}\right) - 1\right), \end{aligned} \tag{36}$$

with the notation $r = b/c > 1$ for the benefit-to-cost ratio and $a = (1/\rho) - 1 > 0$ for the expected number of times that each CC pair continues to interact.

In the extreme case $a = 0$ (or $\rho = 1$ which means no repeated interactions between cooperators), we have $g(x) = -\sigma c < 0$ which implies that $\psi(y)$ is a strictly increasing function of y in $[0,1]$ with $\psi(0) = 1$. None of the conditions in (34) is satisfied. Cooperation is never favored by selection, while defection always is. This is exactly the case of the classic PD game. On the other hand, in the limit as $a \to \infty$ (or $\rho = 0$ which means permanent CC pairs), we have $g(x) = \sigma(b - c) > 0$ which implies that $\psi(y)$ is a strictly decreasing function of y in $[0,1]$ with $\psi(0) = 1$. All conditions in (34) are satisfied. Cooperation is fully favored by selection. This is easy to understand since, in this case, CC pairs never break apart and their number can only increase.

Analogously, in the extreme case $r = 1$, we get $-\sigma c \leqslant g(x) < 0$ for x in $[0,1]$ and, therefore, $\psi(y)$ is a strictly increasing function of y in $[0,1]$ with $\psi(0) = 1$, which implies that none of the conditions in (34) is satisfied and cooperation can never be favored by selection. In this case, cooperators pay as much as they give and the game is actually no longer a PD game. On the other hand, for any given $a > 0$, we have $g(x) > 0$ for x in $[0,1]$ if $r > 0$ is large enough, which implies that $\psi(y)$ is a strictly decreasing function of y in $[0,1]$ with $\psi(0) = 1$. In this case, the conditions in (34) are all satisfied and cooperation is fully favored by selection.

In the general case $0 < \rho < 1$ and $r > 1$, the expression of $g(x)$ in (36) shows that $\partial g(x)/\partial a > 0$ and $\partial g(x)/\partial r > 0$. This implies that $\psi(y)$ is a strictly decreasing function of a and r for every y in $[0,1]$. This leads to the following conclusion.

Result 2. *Consider an additive PD game with payoff matrix (31) in the framework of Result 1 with $0 < \rho < 1$. In a large enough population, increasing the value of $r = b/c > 1$ or $a = 1/\rho - 1 > 0$ (or decreasing the value of ρ) increases (decreases, respectively) the probability of ultimate fixation of cooperation (defection, respectively) introduced as a single mutant in an all defecting (cooperating, respectively) population, F_C (F_D, respectively).*

The proof of this result is straightforward by using the approximations

$$F_C \approx \left(2N\int_0^1 \exp\left(-2\int_0^y g(x)dx\right)dy\right)^{-1}, \tag{37a}$$

$$F_D \approx \left(2N\int_0^1 \exp\left(2\int_y^1 g(x)dx\right)dy\right)^{-1}, \tag{37b}$$

for N large enough.

Moreover, $g(x)$ is a symmetric function, that is, $g(x) = g(1-x)$ holds for x in $[0,1]$. When $F_C = F_D$, that is, $\psi(1) = 1$, we have $\int_0^1 g(x)dx = 0$ from which

$$\begin{aligned} \textstyle\int_0^{1-y} g(x)dx &= \textstyle\int_0^1 g(x)dx - \int_{1-y}^1 g(x)dx \\ &= -\textstyle\int_y^0 g(1-x)d(1-x) \\ &= -\textstyle\int_0^y g(1-x)dx \\ &= -\textstyle\int_0^y g(x)dx. \end{aligned} \tag{38}$$

In this case, we have

$$\begin{aligned} \textstyle\int_0^1 \psi(y)dy &= \textstyle\int_0^{\frac12} \psi(y)dy + \int_{\frac12}^1 \psi(y)dy = \int_0^{\frac12}(\psi(y) + \psi(1-y))dy \\ &= \textstyle\int_0^{\frac12}\left(\exp\left(-2\int_0^y g(x)dx\right) + \exp\left(-2\int_0^{1-y} g(x)dx\right)\right)dy \\ &= \textstyle\int_0^{\frac12}\left(\exp\left(-2\int_0^y g(x)dx\right) + \exp\left(2\int_0^y g(x)dx\right)\right)dy \\ &\geqslant 2\textstyle\int_0^{\frac12} dy = 1, \end{aligned} \tag{39}$$

with an equality if and only if $g(x) = 0$ on $[0,1]$, which means $\sigma = 0$ (no selection). Otherwise, $\int_0^1 \psi(y)dy > 1 = \psi(1)$, which means that $F_C = F_D < (2N)^{-1}$ owing to (34a,b,c). From the previous analysis in the extreme cases and Result 2, we know that increasing the value of a from 0 to ∞, or the value of r from 1 to ∞, will increase F_C from a value smaller than $(2N)^{-1}$ to a value larger than $(2N)^{-1}$, and will decrease F_D in the opposite direction. But when F_C and F_D are equal, their values are less than $(2N)^{-1}$. This implies that F_D crosses the value $(2N)^{-1}$ first, then equals F_C and finally F_C crosses the value $(2N)^{-1}$. Thus we get the following corollary of Result 2.

Result 3. *In the setting of Result 2, as the value of $r = b/c > 1$ or $a = 1/\rho - 1 > 0$ increases, the conditions (34b), (34c) and (34a) for $F_D < (2N)^{-1}, F_C > F_D$ and $F_C > (2N)^{-1}$, respectively, are satisfied in this order. In particular, when cooperation is favored by selection, it is necessarily fully favored by selection.*

The fixation probabilities F_C and F_D as functions of ρ and $r = b/c$, respectively, are showed in Fig. 2 for particular values of c, σ and N. This illustrates the effect of decreasing ρ or increasing r on these probabilities.

In order to get explicit conditions on the parameters of the model for cooperation to be favored by selection, we use the inequalities (see Appendix D for details)

$$4x(1-x)(\sqrt{\rho} - \rho) \leqslant f(x,\rho) \leqslant \sqrt{4x(1-x)}(\sqrt{\rho} - \rho),$$

where the lower bound is the limit of $f(x,\rho)$ as $\rho \to 0$ and the upper bound the limit of $f(x,\rho)$ as $\rho \to 1$. Panels (a), (b) and (c) in Fig. 3 illustrate the surfaces determined by $f(x,\rho), 4x(1-x)(\sqrt{\rho} - \rho)$ and $\sqrt{4x(1-x)}(\sqrt{\rho} - \rho)$, respectively. Panels (d) and (e) show the transverse sections where $f(x,\rho)$ approaches the upper bound when ρ is close to 0 and the lower bound when ρ is close to 1, respectively.

Now, substituting $g(x)$ given in (36) in the expression of $\psi(y)$ given in (35) yields

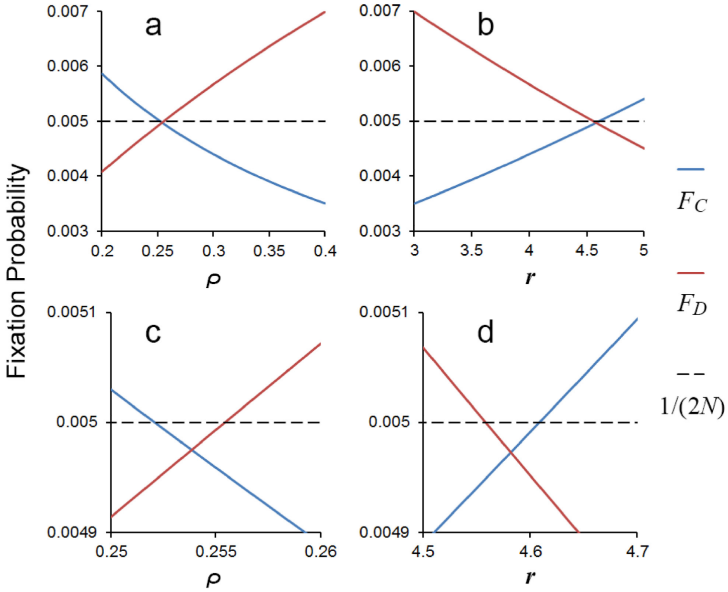

Fig. 2. *Fixation probabilities F_C and F_D as functions of ρ and $r = b/c$.* The numerical values of the expressions given in Eq. (37) are showed in the case $c = 1,\ \sigma = 1,\ N = 100$ with respect to ρ and r on two different scales.

$$\psi(y) = \exp\left(\sigma c\left(-2(r-1)y + \frac{r}{1-\rho}\int_0^y \frac{f(x,\rho)}{x(1-x)}dx\right)\right). \tag{41}$$

Panel (f) in Fig. 3 shows that the lower bound of $f(x,\rho)$ gives a good approximation of the integral in (41) for $y = 1$. Using this lower bound in this equation yields the approximation

$$\psi(y) \approx \exp\left(-2\sigma c\left(r - 1 - r\frac{2\sqrt{\rho}}{1+\sqrt{\rho}}\right)y\right), \tag{42}$$

which is a monotonic function of y starting with the value 1 at $y = 0$. Using this approximation for $\psi(y)$ in (34), this approximation would have to be a strictly decreasing function for cooperation to be favored by selection in any sense, which is the case if and only if

$$r > \frac{1+\sqrt{\rho}}{1-\sqrt{\rho}}, \tag{43}$$

or equivalently

$$\rho < \left(\frac{r-1}{r+1}\right)^2. \tag{44}$$

The right-hand side in (43) is a lower threshold value for r, while the right-hand side in (44) is an upper threshold value for ρ. Analogously, using the upper bound of $f(x,\rho)$ in (41) yields the approximation

$$\psi(y) \approx \exp\left(-2\sigma c\left((r-1)y - r\frac{\sqrt{\rho}}{1+\sqrt{\rho}}(\pi/2 - \arcsin(1-2y))\right)\right). \tag{45}$$

Using this approximation in (34c), this approximation would have to be less than 1 at $y = 1$ for selection to favor more cooperation than defection, which is the case if and only if

$$r > \frac{1+\sqrt{\rho}}{1+\sqrt{\rho}-\pi\sqrt{\rho}}, \tag{46}$$

or equivalently

$$\rho < \left(\frac{r-1}{(\pi-1)r+1}\right)^2. \tag{47}$$

The right-hand side in (46) is an upper threshold value for r, while the right-hand side in (47) is a lower threshold value for ρ. Note that the right-hand side in (46) goes to ∞ as $\rho \to (\pi-1)^{-2} \approx 0.218$. As shown in panel (f) of Fig. 3, the approximation of $f(x,\rho)$ by the upper bound is not that good unless ρ is small enough.

The values of r or ρ such that $F_D = (2N)^{-1}, F_C = F_D$ and $F_C = (2N)^{-1}$, respectively, are illustrated in Fig. 4 in the case $c = 1$ and $\sigma = 1$. The relative positions of the three curves using the exact expression of $f(x,\rho)$ given in (33) are in agreement with Result 3

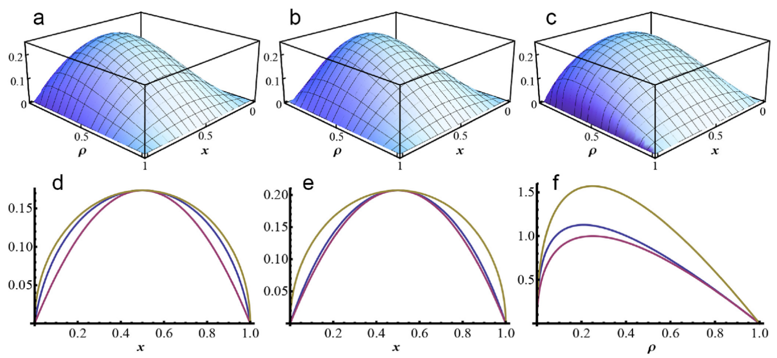

Fig. 3. *Function* $f(x,\rho)$. Panels (a), (b) and (c) show the surfaces determined by the function $f(x,\rho)$, the lower bound $4x(1-x)(\sqrt{\rho}-\rho)$ and the upper bound $\sqrt{4x(1-x)}(\sqrt{\rho}-\rho)$, respectively. Panels (d) and (e) show the transverse sections for $\rho = 0.05$ and $\rho = 0.5$. Panel (f) shows $\int_0^1 \frac{f(x,\rho)}{x(1-x)}dx$ which comes into play in $\psi(1)$. The blue, red and yellow curves in panels (d), (e) and (f) stand for $f(x,\rho)$, $4x(1-x)(\sqrt{\rho}-\rho)$ and $\sqrt{4x(1-x)}(\sqrt{\rho}-\rho)$, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

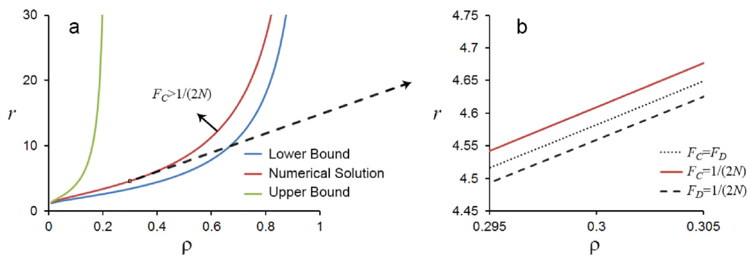

Fig. 4. *Conditions for cooperation to be favored by selection in the additive PD game in the case* $c = 1$ *and* $\sigma = 1$. In panel (a), the red curve stands for the exact numerical solutions of $F_D = (2N)^{-1}$, $F_C = F_D$ and $F_C = (2N)^{-1}$, which are almost identical, while the blue and green curves stand for approximations obtained by using the lower and upper bounds of $f(x,\rho)$ plotted in Fig. 3. Panel (b) is a magnification of the three curves for a small region of the domain in panel (a) at the start of the dashed arrow. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

but very close to each other. These curves are compared to the two curves obtained by using the lower and upper bounds of $f(x,\rho)$ given in (40). These correspond to the boundaries of the regions defined by (43) and (46), respectively. It can be seen that the first one based on the lower bound gives a good approximation. The form of the curves suggests that the conditions $F_D < (2N)^{-1}, F_C > F_D$ and $F_C > (2N)^{-1}$ take all the form r greater than some increasing function of ρ, and the approximation obtained from the lower bound of $f(x,\rho)$ that this function is close to the one given in (43).

5. Discussion

Direct reciprocity is one of the most important mechanisms that can promote the evolution of cooperation (Trivers, 1971; Axelrod and Hamilton, 1981; Axelrod, 1984). In the case of repeated rounds of a two-player Prisoner's Dilemma (PD) game, for instance, tit-for-tat (*TFT*) starting with cooperation is a form of reciprocity and it is known to be a Nash equilibrium against always-defect (*AllD*), and any other strategy, if the probability of repetition from one round to the next is large enough (Nowak et al., 2004; Nowak and Sigmund, 2007). In this paper, we have considered the situation where the PD game can be repeated from one round to the next only if both players are willing to continue their interaction, which occurs with probability $1-\rho$ only when both cooperate. This is a rational choice when all players practice what is known as the opting-out strategy. Then cooperation (C) is equivalent to out-for-tat (*OFT*) starting with cooperation, which cooperates in the first round and opts out with probability 1 if the other defects and probability $1-\sqrt{1-\rho}$ if the other cooperates, and defection (D) to defect-and-run (*DNR*), which defects in the first round and opts out afterwards. This setting creates a kind of assortment that benefits cooperation over defection and should promote its evolution (Hayashi, 1993; Schuessler, 1989; Aktipis, 2004; Fujiwara-

Greve and Okuno-Fujiwara, 2009; Izquierdo et al., 2010, 2014; Fujiwara-Greve et al., 2015).

A theoretical study of the effects of the opting-out strategy on the continuous-time dynamics for a repeated PD game in an infinite population can be found in Zheng et al. (2017). In that paper, the results are based on an analysis of the replicator equation for the C and D frequencies under the assumption of instantaneous equilibrium pair frequencies. In this paper, we have considered the effects of opting-out in a finite population in discrete time and in the limit of a large population size. We have shown that the dynamics of the population state over successive birth–death events according to a Moran model is approximated by a continuous-time diffusion if the intensity of selection and time are appropriately scaled with respect to the population size (Result 1). This has been ascertained by verifying conditions given in Ethier and Nagylaki (1980) for Markov chains with two time scales, here a fast one for pair frequencies and a slow one for C and D frequencies. Note that the drift function in this diffusion approximation given by $m(x)$ in (32), where x is the frequency of C, corresponds to the growth rate of x given by the replicator equation in Zheng et al. (2017). In the diffusion approximation, however, the boundaries $x = 0$ and $x = 1$ are absorbing states that can be reached from any other state, while a stationary distribution with coexistence of C and D is precluded unless a certain level of mutation is introduced.

Assuming an additive PD game with cooperation by an individual incurring a cost c to the individual but providing a benefit b to the opponent, we have shown that increasing the benefit-to-cost ratio, $r = b/c$, or the expected number of repetitions of the PD game for a CC pair, $a = 1/\rho - 1$, makes it easier for the evolution of cooperation to be favored by selection, or for the evolution of defection to be unfavored by selection, or for the evolution of cooperation to be more favored by selection than the evolution of defection (Result 2). Here, this is understood in the sense that the probability of ultimate fixation of C introduced as a single mutant in an all D population under selection, F_C, exceeds what it would be under neutrality, which is given by its initial frequency, or that the probability of ultimate fixation of D introduced as a single mutant in an all C population under selection, F_D, is less than its initial frequency, or that the former probability exceeds the latter. Note that the first condition is the most stringent one and the second condition the least stringent one (Result 3). Moreover, the three conditions take the form r greater than some increasing function of ρ that has been shown to be approximated by $(1+\sqrt{\rho})/(1-\sqrt{\rho})$ (see Fig. 4).

The condition $r > (1+\sqrt{\rho})/(1-\sqrt{\rho})$ for selection to favor the evolution of cooperation in any sense in the case of an additive PD game with opting-out in a large finite population is equivalent to $\rho < (b-c)^2/(b+c)^2$. This happens to be the condition for the coexistence of C and D at a stable interior equilibrium in an infinite population (Zheng et al., 2017). Note that, in this case, C is not a Nash equilibrium and a stable coexistence with D is the only possibility for its maintenance in an infinite population (Fujiwara-Greve et al., 2015). Moreover, the condition for $F_C > F_D$ in a finite population in the absence of mutation is expected to be the condition for C to be more abundant on average than D in the stationary state of a finite population under symmetric recurrent mutation in the limit of a low mutation rate (Antal et al., 2009).

In addition, let us recall that the condition for selection to favor the evolution of *TFT* starting with cooperation against *AllD* in a large finite population is $r > (1+2\rho)/(1-\rho)$, which is equivalent to $\rho < (r-1)/(r+2)$ (Nowak et al., 2004; see Table 1 in Nowak, 2006b). This condition guarantees convergence to the fixation state of *TFT* in an infinite population as soon as *TFT* frequency is larger than 1/3. Moreover, since we have the inequality $(1+\sqrt{\rho})/(1-\sqrt{\rho}) > (1+2\rho)/(1-\rho)$ for ρ in $(0,1)$, this condition is implied by the condition for selection to favor the evolution of *OFT* starting with cooperation against *DNR* in a large finite population. In other words, the condition for selection to favor the evolution of cooperation in a finite population is more stringent with opting-out than without opting-out. This is somehow in agreement with experimental results (Zhang et al., 2016) showing a higher level of cooperation in groups without the possibility of opting-out than in groups using opting-out. However, this does not necessarily mean that *TFT* is more successful than *OFT* in promoting the evolution of cooperation. When both strategies are available, Monte Carlo simulations (Izquierdo et al., 2010; Zheng et al., 2017) have shown that *OFT* can prevail more often than *TFT*. With the possibility of opting-out, many more strategies are possible and the paths to establish cooperation may be very complex and involve indirect invasion of neutrally stable strategies (Bendor and Swistak, 1995; Sandholm, 2010; van Veelen, 2012b; García and van Veelen, 2016).

The opting-out strategy provides an opportunity not only for cooperators to find cooperative partners but also for defectors who have an even greater advantage to do so. Moreover, ending an interaction with someone might incur a cost since there is a risk of not finding a new partner. In our model, there is no cost for opting-out. A cost could affect cooperators and defectors to different degrees and, therefore, the level of cooperation reached in the population.

Finally, the work in this paper has focussed on a two-player game. Kurokawa (2019) has studied the effect of opting-out on a three-player game in an infinite population. Extensions to n-player public goods game would be of interest.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was supported in part by NSERC of Canada (Grant No. 8833) and Chinese Academy of Sciences President's International Fellowship Initiative (Grant No. 2016VBA039). We thank three anonymous referees for helpful comments to improve the manuscript.

Appendix A. Mean and variance of pair frequencies after repairing of free individuals

The Eqs. (2) and (3) for the numbers of free C and D individuals can be written into the form

$$N_C = N\gamma, \tag{48a}$$
$$N_D = N\beta, \tag{48b}$$
$$N_C + N_D = 2N\alpha, \tag{48c}$$

where

$$\gamma = \frac{2R}{N} + P_{CD} \leqslant 2, \tag{49a}$$
$$\beta = P_{CD} + 2P_{DD} = 2(1-x) \leqslant 2, \tag{49b}$$
$$\alpha = \frac{\beta+\gamma}{2} = \frac{R}{N} + P_{CD} + P_{DD} \leqslant 1, \tag{49c}$$

with R being a binomial random variable with parameters NP_{CC} and ρ.

The number of new pairs formed at random by all free individual is $(N_C + N_D)/2 = \alpha N$. Besides, there are $N_{CC} = NP_{CC} - R = N - \alpha N$

unbroken CC pairs. The new pairs are obtained by pairing the $(2k-1)$-th and $2k$-th free individuals chosen at random without replacement for $k = 1, \ldots, \alpha N$. Let $X_{CC,k}, X_{CD,k}$ and $X_{DD,k}$ be the random variables that take the value 1 if the k-th new pair is of types CC, CD and DD, respectively, and 0 otherwise, for $k = 1, \ldots, \alpha N$. Then the numbers of new CC, CD and DD pairs can be expressed as

$$Y_{CC} = \sum_{k=1}^{\alpha N} X_{CC,k}, \tag{50a}$$

$$Y_{CD} = \sum_{k=1}^{\alpha N} X_{CD,k}, \tag{50b}$$

$$Y_{DD} = \sum_{k=1}^{\alpha N} X_{DD,k}, \tag{50c}$$

respectively. By symmetry, given N_C, N_D and α, the first and second conditional moments of $X_{CC,k}$, $X_{CD,k}$ and $X_{DD,k}$ are given by

$$\mathbb{E}(X_{CC,k}) = \mathbb{E}\left(X_{CC,k}^2\right) = \mathbb{E}(X_{CC,1}) = \frac{N_C}{2\alpha N}\frac{N_C-1}{2\alpha N-1}, \tag{51a}$$

$$\mathbb{E}(X_{CD,k}) = \mathbb{E}\left(X_{CD,k}^2\right) = \mathbb{E}(X_{CD,1}) = 2\frac{N_C}{2\alpha N}\frac{N_D}{2\alpha N-1}, \tag{51b}$$

$$\mathbb{E}(X_{DD,k}) = \mathbb{E}\left(X_{DD,k}^2\right) = \mathbb{E}(X_{DD,1}) = \frac{N_D}{2\alpha N}\frac{N_D-1}{2\alpha N-1}, \tag{51c}$$

$$\mathbb{E}(X_{CC,k}X_{CC,l}) = \mathbb{E}(X_{CC,1}X_{CC,2}) = \frac{N_C}{2\alpha N}\frac{N_C-1}{2\alpha N-1}\frac{N_C-2}{2\alpha N-2}\frac{N_C-3}{2\alpha N-3}, \tag{51d}$$

$$\mathbb{E}(X_{CD,k}X_{CD,l}) = \mathbb{E}(X_{CD,1}X_{CD,2}) = 4\frac{N_C}{2\alpha N}\frac{N_D}{2\alpha N-1}\frac{N_C-1}{2\alpha N-2}\frac{N_D-1}{2\alpha N-3}, \tag{51e}$$

$$\mathbb{E}(X_{DD,k}X_{DD,l}) = \mathbb{E}(X_{DD,1}X_{DD,2}) = \frac{N_D}{2\alpha N}\frac{N_D-1}{2\alpha N-1}\frac{N_D-2}{2\alpha N-2}\frac{N_D-3}{2\alpha N-3}, \tag{51f}$$

where $l \neq k$. In particular, with $\gamma = N_C/N$ and $\beta = N_D/N$, the first conditional moments lead to

$$\mathbb{E}(Y_{CC}) = \frac{N_C(N_C-1)}{2(2\alpha N-1)} = \frac{\gamma N(\gamma N-1)}{2(2\alpha N-1)}, \tag{52a}$$

$$\mathbb{E}(Y_{CD}) = \frac{N_C N_D}{2\alpha N-1} = \frac{\gamma N \beta N}{2\alpha N-1} = \frac{(2\alpha N-\beta N)\beta N}{2\alpha N-1}, \tag{52b}$$

$$\mathbb{E}(Y_{DD}) = \frac{N_D(N_D-1)}{2(2\alpha N-1)} = \frac{\beta N(\beta N-1)}{2(2\alpha N-1)}. \tag{52c}$$

Now, let q_{CC}, q_{CD} and q_{DD} represent the random frequencies of CC, CD and DD, respectively, among the αN new pairs and $N-\alpha N$ unbroken CC pairs. Using the fact that an expected value of a conditional expected value is the expected value, we have

$$\mathbf{E}(q_{DD}) = \mathbf{E}\left(\frac{Y_{DD}}{N}\right) = \frac{1}{2}\mathbf{E}\left(\frac{\beta^2 N-\beta}{2\alpha N-1}\right), \tag{53a}$$

$$\mathbf{E}(q_{CD}) = \mathbf{E}\left(\frac{Y_{CD}}{N}\right) = \mathbf{E}\left(\frac{\beta(2\alpha N-\beta N)}{2\alpha N-1}\right) = \beta - \mathbf{E}\left(\frac{\beta^2 N-\beta}{2\alpha N-1}\right), \tag{53b}$$

$$\mathbf{E}(q_{CC}) = 1 - \mathbf{E}(q_{CD}) - \mathbf{E}(q_{DD}) = 1 - \beta + \frac{1}{2}\mathbf{E}\left(\frac{\beta^2 N-\beta}{2\alpha N-1}\right). \tag{53c}$$

Note that

$$\frac{\beta^2 N-\beta}{2\alpha N-1} = \beta\frac{\beta-1/N}{\beta+\gamma-1/N} \leqslant \beta \leqslant 2. \tag{54}$$

Moreover, the expected value of this random variable is given by

$$\mathbf{E}\left(\frac{\beta^2 N-\beta}{2\alpha N-1}\right) = \frac{\beta^2}{2(\rho P_{CC}+P_{CD}+P_{DD})} + O\left(N^{-1/2}\right). \tag{55}$$

This is obviously true when $\beta = 0$. On the other hand, when $\beta = P_{CD} + 2P_{DD} > 0$, we have $P_{CD} \geqslant 1/N$ or $P_{DD} \geqslant 1/N$, from which $\alpha = R/N + P_{CD} + P_{DD} \geqslant P_{CD} + P_{DD} \geqslant 1/N$. In this case, we have

$$\begin{aligned} 0 &\leqslant \frac{\beta}{\alpha-1/(2N)} \leqslant \frac{P_{CD}+2P_{DD}}{P_{CD}+P_{DD}-1/(2N)} \\ &\leqslant 2\frac{P_{CD}+P_{DD}}{P_{CD}+P_{DD}-1/(2N)} \leqslant 4. \end{aligned} \tag{56}$$

Moreover, $\beta = P_{CD} + 2P_{DD} \leqslant 2(\rho P_{CC} + P_{CD} + P_{DD})$. Using these inequalities and the Cauchy–Schwarz inequality yields

$$\begin{aligned} &\left|\mathbf{E}\left(\tfrac{\beta^2 N-\beta}{2\alpha N-1}\right) - \tfrac{\beta^2}{2(\rho P_{CC}+P_{CD}+P_{DD})}\right| \\ &\leqslant \tfrac{1}{2}\mathbf{E}\left(\left|\tfrac{\beta(\beta-1/N)}{\alpha-1/(2N)} - \tfrac{\beta^2}{\rho P_{CC}+P_{CD}+P_{DD}}\right|\right) \\ &= \tfrac{1}{2}\mathbf{E}\left(\tfrac{\beta}{\alpha-1/(2N)}\left|\beta - \tfrac{1}{N} - \tfrac{\beta(\alpha-1/(2N))}{\rho P_{CC}+P_{CD}+P_{DD}}\right|\right) \\ &\leqslant 2\mathbf{E}\left(\left|\tfrac{\beta(\rho P_{CC}-R/N+1/(2N))}{\rho P_{CC}+P_{CD}+P_{DD}} - \tfrac{1}{N}\right|\right) \\ &\leqslant 2\mathbf{E}\left(\left|\tfrac{\beta(\rho P_{CC}-R/N+1/(2N))}{\rho P_{CC}+P_{CD}+P_{DD}}\right| + \tfrac{1}{N}\right) \\ &\leqslant 4\mathbf{E}\left(\left|\rho P_{CC} - \tfrac{R}{N} + \tfrac{1}{2N}\right|\right) + \tfrac{2}{N} \\ &\leqslant \tfrac{4}{N}\mathbf{E}(|\rho N P_{CC} - R|) + \tfrac{4}{N} \\ &\leqslant \tfrac{4}{N}\left(\mathbf{E}\left((\rho N P_{CC} - R)^2\right)\right)^{1/2} + \tfrac{4}{N} \\ &= 4\left(\tfrac{P_{CC}\rho(1-\rho)}{N}\right)^{1/2} + \tfrac{4}{N} \\ &= O\left(N^{-1/2}\right), \end{aligned} \tag{57}$$

which proves Eq. (55). Using this equation and the equality $\beta^2 = 4(1-x)^2$ in (53) yields

$$\mathbf{E}(q_{DD}) = \frac{(1-x)^2}{1-(1-\rho)P_{CC}} + O\left(N^{-1/2}\right), \tag{58a}$$

$$\mathbf{E}(q_{CD}) = 2(1-x) - \frac{2(1-x)^2}{1-(1-\rho)P_{CC}} + O\left(N^{-1/2}\right), \tag{58b}$$

$$\mathbf{E}(q_{CC}) = 2x - 1 + \frac{(1-x)^2}{1-(1-\rho)P_{CC}} + O\left(N^{-1/2}\right). \tag{58c}$$

Moreover, using the inequality $2\alpha \geqslant \beta$, we have

$$\begin{aligned} \mathbf{Var}(q_{DD}) &= \mathbf{Var}\left(\tfrac{Y_{DD}}{N}\right) \\ &= \tfrac{1}{N^2}\left(\mathbf{E}\left(Y_{DD}^2\right) - (\mathbf{E}(Y_{DD}))^2\right) \\ &= \tfrac{1}{N^2}\left(\mathbf{E}\left(\left(\sum_{k=1}^{\alpha N} X_{DD,k}\right)^2\right) - \left(\mathbf{E}\left(\sum_{k=1}^{\alpha N} X_{DD,k}\right)\right)^2\right) \\ &= \tfrac{1}{N^2}\mathbf{E}\left(\alpha N(\alpha N-1)\mathbb{E}(X_{DD,1}X_{DD,2}) + \alpha N\mathbb{E}\left(X_{DD,1}^2\right) - (\alpha N\mathbb{E}(X_{DD,1}))^2\right) \\ &= \tfrac{1}{N^2}\mathbf{E}\Big(\alpha N(\alpha N-1)\tfrac{\beta N}{2\alpha N}\tfrac{\beta N-1}{2\alpha N-1}\tfrac{\beta N-2}{2\alpha N-2}\tfrac{\beta N-3}{2\alpha N-3} \\ &\quad + \alpha N\tfrac{\beta N(\beta N-1)}{2\alpha N(2\alpha N-1)} - \alpha^2 N^2\left(\tfrac{\beta N(\beta N-1)}{2\alpha N(2\alpha N-1)}\right)^2\Big) \\ &= \mathbf{E}\left(\alpha^2\tfrac{\beta N(\beta N-1)}{2\alpha N(2\alpha N-1)}\left(\tfrac{(\beta N-2)(\beta N-3)}{(2\alpha N-2)(2\alpha N-3)} - \tfrac{\beta N(\beta N-1)}{2\alpha N(2\alpha N-1)}\right)\right) \\ &\quad + \mathbf{E}\left(\tfrac{\alpha}{N}\tfrac{\beta N(\beta N-1)}{2\alpha N(2\alpha N-1)}\left(1 - \tfrac{(\beta N-2)(\beta N-3)}{(2\alpha N-2)(2\alpha N-3)}\right)\right) \\ &\leqslant \tfrac{1}{N}\mathbf{E}\left(\alpha\tfrac{\beta N(\beta N-1)}{2\alpha N(2\alpha N-1)}\left(1 - \tfrac{(\beta N-2)(\beta N-3)}{(2\alpha N-2)(2\alpha N-3)}\right)\right) \leqslant \tfrac{1}{N}, \end{aligned} \tag{59}$$

which implies that $\mathbf{Var}(q_{DD}) = O\left(N^{-1}\right)$. Analogously, using the inequality $2\alpha \geqslant \gamma$, we have

$$\mathbf{Var}(q_{CD}) = \mathbf{Var}\left(\frac{Y_{CD}}{N}\right) = O\left(N^{-1}\right).$$

Since $q_{CC}+q_{CD}+q_{DD}=1$, we also have

$$\mathbf{Var}(q_{CC})=\mathbf{Var}(1-q_{CD}-q_{DD})\leqslant 2(\mathbf{Var}(q_{CD})+\mathbf{Var}(q_{DD}))=O\left(N^{-1}\right). \tag{61}$$

Appendix B. Moments of changes in C and CC frequencies

From time t to time $t+\Delta t$, the frequencies of CC, CD and DD pairs go from $P_{CC},\ P_{CD}$ and P_{DD}, with $x=P_{CC}+P_{CD}/2$ as frequency of C, to $P'_{CC},\ P'_{CD}$ and P'_{DD}, with $x'=P'_{CC}+P'_{CD}/2$ as frequency of C. This follows first random re-pairing of free individuals, second weighted random sampling of an individual to produce an offspring with weight given by fitness, and third random sampling of an individual to be replaced by the offspring. The random frequencies of CC, CD and DD pairs after the first step are q_{CC}, q_{CD} and q_{DD} with $x=q_{CC}+q_{CD}/2$, while according to (9) and (10), the conditional probabilities to sample C and D at the second step are

$$Pr(C)=x+\frac{\sigma}{N}((1-x)A-xB)+o\left(N^{-1}\right) \tag{62}$$

and

$$Pr(D)=1-x-\frac{\sigma}{N}((1-x)A-xB)+o\left(N^{-1}\right), \tag{63}$$

respectively, where

$$A=q_{CC}\pi_{CC}+\frac{q_{CD}}{2}\pi_{CD} \tag{64}$$

and

$$B=q_{DD}\pi_{DD}+\frac{q_{CD}}{2}\pi_{DC}. \tag{65}$$

Obviously, the probabilities to sample C and D at the third step are x and $1-x$, respectively.

B.1. Change in C frequency

The change in the frequency of C from time t to time $t+\Delta t$, represented by $\Delta x=x'-x$, takes the values $1/(2N)$ and $-1/(2N)$ with conditional probabilities $(1-x)Pr(C)$ and $xPr(D)$, respectively, and 0 otherwise. The expected value of this change is

$$\begin{aligned}\mathbf{E}(\Delta x)&=\mathbf{E}\left(\tfrac{1}{2N}(1-x)Pr(C)-\tfrac{1}{2N}xPr(D)\right)\\&=\tfrac{\sigma}{2N^2}\mathbf{E}((1-x)A-xB)+o\left(N^{-2}\right).\end{aligned} \tag{66}$$

The second moment is given by

$$\begin{aligned}\mathbf{E}\left((\Delta x)^2\right)&=\mathbf{E}\left(\tfrac{1}{4N^2}(1-x)Pr(C)+\tfrac{1}{4N^2}xPr(D)\right)\\&=\tfrac{1}{2N^2}x(1-x)+o\left(N^{-2}\right).\end{aligned} \tag{67}$$

As for the fourth conditional moment, we have

$$\mathbf{E}\left((\Delta x)^4\right)=\mathbf{E}\left(\frac{1}{16N^4}(1-x)Pr(C)+\frac{1}{16N^4}xPr(D)\right)=o\left(N^{-3}\right). \tag{68}$$

B.2. Change in CC frequency

The change in the frequency of CC pairs from time t to time $t+\Delta t$ is given by

$$\Delta P_{CC}=P'_{CC}-P_{CC}=\left(P'_{CC}-q_{CC}\right)+q_{CC}-P_{CC}. \tag{69}$$

Given q_{CC}, q_{CD} and q_{DD}, the difference $P'_{CC}-q_{CC}$ takes the values $1/N$ and $-1/N$ with conditional probabilities $(q_{CD}/2)Pr(C)$ and $q_{CC}Pr(D)$, respectively, and 0 otherwise. Therefore, the difference has an expected value $O\left(N^{-1}\right)$ and a variance $O\left(N^{-2}\right)$. Using this and the expected values and variances of q_{CC}, q_{CD} and q_{DD} given in Appendix A, we get

$$\begin{aligned}\mathbf{E}(\Delta P_{CC})&=\mathbf{E}\left(P'_{CC}-q_{CC}\right)+\mathbf{E}(q_{CC})-P_{CC}\\&=2x-1+\tfrac{(1-x)^2}{1-(1-\rho)P_{CC}}-P_{CC}+O\left(N^{-1/2}\right)\\&=\tfrac{(x-P_{CC})^2-\rho P_{CC}(1-2x+P_{CC})}{1-(1-\rho)P_{CC}}+O\left(N^{-1/2}\right)\end{aligned} \tag{70}$$

and

$$\mathbf{Var}(\Delta P_{CC})=\mathbf{Var}\left(\left(P'_{CC}-q_{CC}\right)+q_{CC}\right)\leqslant 2\left(\mathbf{Var}\left(P'_{CC}-q_{CC}\right)+\mathbf{Var}(q_{CC})\right)=O\left(N^{-1}\right). \tag{71}$$

Appendix C. Convergence of CC frequency in an infinite neutral population

In an infinite population with no selection, the frequency of C in $[0,1]$, represented by $x=P_{CC}+P_{CD}/2$, remains constant since then

$$x'-x=\mathbf{E}(\Delta x)=0 \tag{72}$$

owing to (66) as $N\to\infty$. Moreover, the change in the frequency of CC from time t to time $t+\Delta t$ is given by

$$P'_{CC}-P_{CC}=\mathbf{E}(\Delta P_{CC})=\frac{(x-P_{CC})^2-\rho P_{CC}(1-2x+P_{CC})}{1-(1-\rho)P_{CC}} \tag{73}$$

owing to (70) as $N\to\infty$. After algebraic manipulations, this leads to the recurrence equation

$$P'_{CC}=2x-1+\frac{(1-x)^2}{1-(1-\rho)P_{CC}}=h(P_{CC}). \tag{74}$$

From the facts that $P_{CC}, P_{CD}=2(x-P_{CC})$ and $P_{DD}=1-2x+P_{CC}$ are all in $[0,1]$, we have the constraints

$$\max\{2x-1,0\}\leqslant P_{CC}\leqslant x. \tag{75}$$

Note that $h(0)=x^2\geqslant 0$ and $h(2x-1)\geqslant 2x-1$, so that $h(\max\{2x-1,0\})\geqslant \max\{2x-1,0\}$, while

$$h(x)=x-\frac{\rho x(1-x)}{1-(1-\rho)x}\leqslant x. \tag{76}$$

On the other hand, the first and second derivatives of h are given by

$$\frac{dh(P_{CC})}{dP_{CC}}=\frac{(1-x)^2(1-\rho)}{(1-(1-\rho)P_{CC})^2}\geqslant 0 \tag{77}$$

and

$$\frac{d^2h(P_{CC})}{dP_{CC}^2}=\frac{2(1-x)^2(1-\rho)^2}{(1-(1-\rho)P_{CC})^3}\geqslant 0, \tag{78}$$

respectively. By solving the equation $h(P_{CC})=P_{CC}$, that is,

$$(1-\rho)P_{CC}^2-(2x(1-\rho)+\rho)P_{CC}+x^2=0, \tag{79}$$

the only equilibrium point of h in the interval $[\max\{2x-1,0\},x]$ is found to be

$$P^*_{CC}=x+\frac{\rho}{2(1-\rho)}-\frac{\sqrt{\rho^2+4x(1-x)\rho(1-\rho)}}{2(1-\rho)}. \tag{80}$$

Owing to the above properties, this is a globally stable equilibrium point. At this equilibrium, we get from (74) that

$$\frac{(1-x)^2}{1-(1-\rho)P^*_{CC}} = P^*_{CC} - 2x + 1, \tag{81}$$

which simplifies the expressions for $\mathbf{E}(q_{CC})$, $\mathbf{E}(q_{CD})$ and $\mathbf{E}(q_{DD})$ in (58).

Actually, P^*_{CC} is a uniformly globally stable equilibrium point. As a matter of fact, applying the mean value theorem, there exists $\widetilde{P}_{CC}$ between P_{CC} and P^*_{CC} such that

$$P'_{CC} - P^*_{CC} = h(P_{CC}) - h(P^*_{CC}) = (P_{CC} - P^*_{CC})\frac{dh\left(\widetilde{P}_{CC}\right)}{dP_{CC}} \tag{82}$$

with

$$\frac{dh\left(\tilde{P}_{CC}\right)}{dP_{CC}} = \frac{(1-x)^2(1-\rho)}{\left(1-(1-\rho)\tilde{P}_{CC}\right)^2} \leqslant \frac{(1-x)^2(1-\rho)}{(1-(1-\rho)x)^2} \leqslant 1-\rho < 1. \tag{83}$$

Iterating (82), we have

$$|P^{(n)}_{CC} - P^*_{CC}| \leqslant |P_{CC} - P^*_{CC}|(1-\rho)^n \leqslant (1-\rho)^n \tag{84}$$

for all integers $n \geqslant 1$, with $(1-\rho)^n \to 0$ as $n \to \infty$.

Appendix D. Bounds of $f(x, \rho)$

For the additive PD game, the drift function $m(x)$ is in the form

$$m(x) = \sigma\left(x(1-x)(b-c) - \frac{bf(x,\rho)}{2(1-\rho)}\right), \tag{85}$$

where

$$\begin{aligned} f(x,\rho) &= \sqrt{\rho^2 + 4x(1-x)\rho(1-\rho)} - \rho \\ &= \phi(4x(1-x),\rho)(\sqrt{\rho}-\rho) \end{aligned} \tag{86}$$

with

$$\phi(u,\rho) = \frac{\sqrt{\rho + u(1-\rho)} - \sqrt{\rho}}{1-\sqrt{\rho}} \tag{87}$$

for u, ρ in $[0,1]$. We have $\phi(u,0) = \sqrt{u}$ and $\phi(u,1) = \lim_{\rho\to 1}\phi(u,\rho) = u$ by applying L'Hôpital's rule. Moreover,

$$\frac{\partial\phi(u,\rho)}{\partial\rho} = \frac{(1-u)\sqrt{\rho} + u - \sqrt{\rho + u(1-\rho)}}{2(1-\sqrt{\rho})^2\sqrt{\rho}\sqrt{\rho + u(1-\rho)}} \leqslant 0, \tag{88}$$

since

$$(1-u)\sqrt{\rho} + u \leqslant \sqrt{(1-u)\rho + u} = \sqrt{\rho + u(1-\rho)} \tag{89}$$

by Jensen's inequality for the concave square root function on $[0,1]$. Therefore, $\phi(u,\rho)$ is a decreasing function of ρ from $\sqrt{u}$ at $\rho = 0$ to u at $\rho = 1$ for every u in $[0,1]$.

We conclude that

$$\begin{aligned} \sqrt{4x(1-x)}(\sqrt{\rho}-\rho) &\geqslant f(x,\rho) = \phi(4x(1-x),\rho)(\sqrt{\rho}-\rho) \\ &\geqslant 4x(1-x)(\sqrt{\rho}-\rho), \end{aligned} \tag{90}$$

the upper bound being the limit of $f(x,\rho)$ as $\rho \to 0$, and the lower bound the limit of $f(x,\rho)$ as $\rho \to 1$ (see Fig. 3).

References

Aktipis, C.A., 2004. Know when to walk away: contingent movement and the evolution of cooperation. Journal of Theoretical Biology 231, 249–260.

Antal, T., Nowak, M.A., Traulsen, A., 2009. Strategy abundance in 2×2 games for arbitrary mutation rates. Journal of Theoretical Biology 257, 340–344.

Axelrod, R., 1984. The Evolution of Cooperation. Basic Books, New York.

Axelrod, R., Hamilton, W.D., 1981. The evolution of cooperation. Science 211, 1390–1396.

Bendor, J., Swistak, P., 1995. Types of evolutionary stability and the problem of cooperation. Proceedings of the National Academy of Sciences USA 92, 3596–3600.

Bergmüller, R., Schürch, R., Hamilton, I.M., 2010. Evolutionary causes and consequences of consistent individual variation in cooperative behaviour. Philosophical Transactions of the Royal Society B 365, 2751–2764.

Brask, J.B., Croft, D.P., Edenbrow, M., James, R., Bleakley, B.H., Ramnarine, I.W., Heathcote, R.J.P., Tyler, C.R., Hamilton, P.B., Dabelsteen, T., Darden, S.K., 2019. Evolution of non-kin cooperation: social assortment by cooperative phenotype in guppies. Royal Society Open Science 6. https://doi.org/10.1098/rsos.181493. 181493.

Croft, D.P., Edenbrow, M., Darden, S.K., 2015. Assortment in social networks and the evolution of cooperation. In: Krause, J., James, R., Franks, D.W., Croft, D.P. (Eds.), Animal Social Networks. (Oxford University Press, Oxford, UK, pp. 13–23.

Eshel, I., Cavalli-Sforza, L.L., 1982. Assortment of encounters and evolution of cooperativeness. Proceedings of the National Academy of Sciences USA 79, 1331–1335.

Ethier, S.N., Nagylaki, T., 1980. Diffusion approximations of Markov chains with two time scales and applications to population genetics. Advances in Applied Probability 12, 14–49.

Ewens, W.J., 2004. Mathematical Population Genetics: I Theoretical Introduction. Springer, New York.

Fujiwara-Greve, T., Okuno-Fujiwara, M., 2009. Voluntarily separable repeated Prisoner's Dilemma. The Review of Economic Studies 76, 993–1021.

Fujiwara-Greve, T., Okuno-Fujiwara, M., Suzuki, N., 2015. Efficiency may improve when defectors exist. Economic Theory 60, 423–460.

García, J., van Veelen, M., 2016. In and out of equilibrium I: Evolution of strategies in repeated games with discounting. Journal of Economic Theory 161, 161–189.

Hayashi, N., 1993. From TIT-for-TAT to OUT-for-TAT. Sociological Theory and Methods 8, 19–32.

Hofbauer, J., Sigmund, K., 1998. The Theory of Evolution and Dynamical Systems. Cambridge University Press, Cambridge.

Izquierdo, S.S., Izquierdo, L.R., Vega-Redondo, F., 2010. The option to leave: Conditional dissociation in the evolution of cooperation. Journal of Theoretical Biology 267, 76–84.

Izquierdo, S.S., Izquierdo, L.R., Vega-Redondo, F., 2014. Leave and let leave: A sufficient condition to explain the evolutionary emergence of cooperation. Journal of Economic Dynamics and Control 46, 91–113.

Kimura, M., 1964. Diffusion models in population genetics. Journal of Applied Probability 1, 177–232.

Krause, J., James, R., Franks, D.W., Croft, D.P., 2015. Animal Social Networks. Oxford University Press, Oxford, UK.

Kurokawa, S., 2019. Three-player repeated games with an opt-out option. Journal of Theoretical Biology 480, 13–22.

Li, C., Lessard, S., 2020. Randomized matrix games in a finite population: effect of stochastic fluctuations in the payoffs on the evolution of cooperation. Theoretical Population Biology 134, 77–91.

Nowak, M.A., 2006a. Evolutionary Dynamics. Harvard University Press, Cambridge, MA.

Nowak, M.A., 2006b. Five rules for the evolution of cooperation. Nature 314, 1560–1563.

Nowak, M.A., Sasaki, A., Taylor, C., Fudenberg, D., 2004. Emergence of cooperation and evolutionary stability in finite populations. Nature 428, 646–650.

Nowak, M.A., Sigmund, K., 2007. How populations cohere: Five rules for cooperation. In: May, R.M., McLean, A. (Eds.), Theoretical Ecology – Principles and Applications, Oxford University Press, Oxford, pp. 7–16.

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.

Poundstone, W., 1992. Prisoner's Dilemma. Doubleday, New York.

Risken, H., 1992. The Fokker-Planck Equation, Methods of Solution and Applications. Springer, Berlin.

Sandholm, W.H., 2010. Population Games and Evolutionary Dynamics. The MIT Press, Cambridge, MA.

Schuessler, R., 1989. Exit threats and cooperation under anonymity. Journal of Conflict Resolution 33, 728–749.

Trivers, R.L., 1971. Evolution of reciprocal altruism. Quarterly Review of Biology 46, 35–57.

van Veelen, M., García, J., Rand, D.G., Nowak, M.A., 2012a. Direct reciprocity in structured populations. Proceedings of the National Academy of Sciences USA 109 (25), 9929–9934.

van Veelen, M., 2012b. Robustness against indirect invasions. Games and Economic Behavior 74, 382–393.

Zhang, B.-Y., Fan, S.-J., Li, C., Zheng, X.-D., Bao, J.-Z., Cressman, R., Tao, Y., 2016. Opting out against defection leads to stable coexistence with cooperation. Scientific Reports 6, 35902.

Zheng, X.-D., Li, C., Yu, J.-R., Wang, S.-C., Fan, S.-J., Zhang, B.-Y., Tao, Y., 2017. A simple rule of direct reciprocity leads to the stable coexistence of cooperation and defection in the Prisoner's Dilemma game. Journal of Theoretical Biology 420, 12–17.