Assortment by Group Founders Always Promotes the Evolution of Cooperation Under Global Selection But Can Oppose it Under Local Selection

Éloi Martin[1] · Sabin Lessard[1]

Abstract

We consider assortment of cooperators or defectors within groups by assuming that group founders express a preference for other group members that are identical-by-type to them. Assuming pairwise interactions within groups and payoffs determined by a game matrix as in a repeated Prisoner's dilemma with strategies tit-for-tat (representing cooperation C) and always-defect (representing defection D), we show that the dynamics in an infinite population or a large finite population is described by a replicator equation or a diffusion approximation, respectively, as for random pairwise interactions in a well-mixed population with some effective game matrix involving coefficients of pairwise and tripletwise assortment. We deduce that an increase in the assortment level facilitates the evolution of cooperation based on stability properties in an infinite population as well as fixation probabilities in a finite population if selection is global so that groups contribute proportionally to their average payoffs, or if selection is local so that groups contribute equally, but in this case only if C is risk dominant over D. Actually, in the latter case, an increase in the assortment level is less conducive to cooperation, and even opposes the evolution of cooperation, if D is risk dominant over C. Moreover, we show that stochastic variability in the assortment level and/or the group size always facilitates the evolution of cooperation in the former case but not necessarily in the latter.

Keywords Fixation probability · One-third law of evolution · Diffusion approximation · Repeated Prisoner's Dilemma · Coefficients of assortment · Local replicator dynamics

Mathematics Subject Classification Primary 91A22; Secondary 92D15

✉ Sabin Lessard
  sabin.lessard@umontreal.ca

[1] Département de mathématiques et de statistique, Université de Montréal, Montréal H3C 3J7, Canada

Birkhäuser

1 Introduction

The evolution of cooperation remains one of the main topics studied in evolutionary game theory. The reason may be that, although cooperation is unequivocally widespread in natural populations, its evolution following its introduction as a mutant strategy raises challenging questions. As a matter of fact, while it is generally admitted that cooperation may be beneficial once fixed in a population, it is difficult to imagine that it can be advantageous when rare. This is well illustrated by the Prisoner's Dilemma (PD) in pairwise interactions with cooperation paying less than defection against defection [4]. Actually, in the PD game, cooperation pays less than defection against both cooperation and defection. Then, cooperation is completely dominated by defection and this makes impossible its maintenance at any frequency in the population. For such a maintenance to be possible, the average payoff to cooperation has to be higher and this will be the case if the situation is such that cooperation against cooperation will be used more often, or cooperation against defection less often. In a repeated PD game, for instance, the use of the sequential strategy tit-for-tat from one round to the next starting with cooperation can achieve this goal and it has proved to be very efficient against any other sequential strategy using cooperation or defection in each round (Axelrod [3], see Hofbauer and Sigmund [20], and references there in for further details). Other strategies that can have a similar effect have been studied later on such as the out-for-tat strategy starting with cooperation, also known as the opting-out strategy (see, e.g., Zhang et al. [56], Krivan and Cressman [23], Li and Lessard [36]). See also Nowak [42] and references therein for other mechanisms of reciprocity, direct or indirect involving reputation, reward or punishment, that can all favour the use of cooperation.

Another way to promote cooperation is to increase the frequency of encounters between individuals using cooperation, which will be the case if encounters between individuals using the same strategy occur more often than expected only by chance. This may be the result of constraints on the population structure or assortative pairing of individuals based on relatedness or type [5, 11, 17, 18, 52]. If interactions occur between relatives, for instance, the individuals can use the same strategy because they are *identical-by-descent* [37]. But in other models, interacting individuals can use the same strategy just because they were assorted so by choice. Then, they could be said *identical-by-assortment*. Note that in both cases, the interacting individuals are *identical-by-type* but they can also be so only by chance.

With random pairwise interactions in an infinite well-mixed population, tit-for-tat is locally stable against always-defect in a repeated PD game under the replicator dynamics if the number of rounds is large enough (Taylor and Jonker [50]; see, e.g., Hofbauer and Sigmund [20]). However, always-defect is locally stable too, and therefore tit-for-tat cannot increase in frequency when rare. In a finite population, however, the payoffs at all the possible population states come into play in the determination of fixation probabilities and these can turn out in the advantage of tit-for-tat. Actually, the one-third law of evolution predicts that under the most common updating rules from one time step to the next the probability for tit-for-tat to fix in the population following its introduction as a single mutant exceeds its neutral fixation probability if its average payoff is larger than the average payoff to always-defect when the frequency of tit-for-tat is $1/3$ [21, 28, 33, 41]. Moreover, if this occurs when the frequency of tit-for-tat is $1/2$, then this fixation probability for tit-for-tat exceeds the corresponding one for always-defect. All these are conditions that may contribute to the evolution of cooperation.

A group structure has often been considered in evolutionary game theory to study conditions for the evolution of cooperation in an infinite population. Ohtsuki [43], for instance, assumed pairwise interactions within groups, so that average payoffs are linear functions, and

partial dispersal, so that interacting individuals may be related. On the other hand, Hilbe [19] considered *local selection* within groups (soft selection) instead of *global selection* (hard selection) in the case of complete dispersal. All these assumptions and more were taken into account in Lessard [30] for an infinite population as well as in Lessard [32] for a finite population. Note that, using a generalization of tit-for-tat based on the number of cooperators in a repeated n-player PD game with a linear payoff function, Boyd and Richerson [6] had shown that conditions for the evolution of cooperation in an infinite population are usually less restrictive when groups are small and formed of relatives (e.g., sibs). In general, of course, the average payoffs may nonlinearly depend on the number of cooperators (see, e.g., [1, 7, 51]). Recently, Kristensen et al. [24] studied a threshold public goods game under global selection in an infinite population. Following Ohtsuki [44], three models of assortment within groups based on identity-by-descent were considered: leader-driven, members attract, members recruit. The leader-driven group-formation model was shown to be the most conducive to the evolution of cooperation by promoting both the invasion and persistence of cooperators.

In this paper, we will assume that group founders express a preference for other group members which are identical-by-type to them, which is analogous to the leader-driven assortment model. This will create a bias for groups of individuals of the same type. Assuming pairwise interactions within groups and payoffs determined by a game matrix as in a repeated PD game with tit-for-tat and always-defect as strategies, we will address the effect of the assortment level on conditions that can promote the evolution of cooperation in an infinite population under the replicator dynamics as well as in a finite population using a diffusion approximation. Stochastic variability in the assortment level and the group size will also be considered. Two different assumptions on group contributions will be made: contributions proportional to the average payoffs in the groups as a result of global selection and equal contributions irrespective of the group types as a result of local selection.

2 Distribution of Types Within Groups

We consider groups of k interacting individuals in a large (finite or infinite) population. There are two types of individuals in the population, C and D, and the fitness (or reproductive success) of a focal individual will depend not only on its own type but also on the types of the other group members with whom the focal individual can interact.

We will say that a group is of type i if it contains i individuals of type C and $k - i$ individuals of type D, and use z_i to denote the frequency of groups of type i present in the population, for $i = 0, 1, \ldots, k$. Let Z represent the random variable for the number of individuals of type C in a group chosen at random. If x is the frequency of type C in the whole population, then we have

$$\frac{\mathbb{E}[Z]}{k} = \sum_{i=0}^{k} z_i \frac{i}{k} = x, \tag{1}$$

from which $\mathbb{E}[Z] = kx$.

The perceived distribution of groups by an individual of a given type differs from the distribution of Z (see, e.g., [7, 45]). In fact, the probability for an individual of type C chosen at random to be in a group with i other individuals of type C and $k - 1 - i$ individuals of

type D is

$$p_C(i) = \frac{(i+1)z_{i+1}}{\sum_{l=0}^{k-1}(l+1)z_{l+1}} = \frac{(i+1)z_{i+1}}{kx}, \tag{2}$$

and similarly

$$p_D(i) = \frac{(k-i)z_i}{k(1-x)} \tag{3}$$

is the probability for an individual of type D to be in a group with $k-1-i$ other individuals of type D and i individuals of type C, for $i = 0, 1, \ldots, k-1$.

Note that we have $p_D(i) = p_C(i)$ for $i = 0, 1, \ldots, k-1$ if and only if

$$z_i = \frac{k!}{i!(k-i)!}x^i(1-x)^{k-i} \tag{4}$$

for $i = 0, 1, \ldots, k-1$, which means that Z follows a binomial distribution with parameters k and x. This situation corresponds to the case where groups are formed by random sampling with replacement of k individuals in a population in which the frequency of C is x and the frequency of D is $1-x$. We will refer to this situation as the *neutral-assortment model*. Let us recall that sampling with replacement is equivalent to sampling without replacement in an infinite population, while the difference becomes negligible in a finite population in the limit of a large population size.

In the general case, if X_C represents the number of individuals of type C in the same group as a randomly selected individual of type C, then its expected value defined as $\mathbb{E}[X_C] = \sum_{i=0}^{k-1} i p_C(i)$ is given by

$$\begin{aligned}
\mathbb{E}[X_C] &= \frac{1}{kx}\left(\sum_{i=0}^{k-1}(i+1)^2 z_{i+1} - \sum_{i=0}^{k-1}(i+1)z_{i+1}\right) \\
&= \frac{\mathbb{E}[Z^2] - \mathbb{E}[Z]}{kx} \\
&= \frac{\text{Var}[Z] + \mathbb{E}[Z]^2 - \mathbb{E}[Z]}{kx} \\
&= kx - 1 + \frac{\text{Var}[Z]}{kx}, \tag{5}
\end{aligned}$$

where $\text{Var}[Z]$ denotes the variance of Z. Similarly, we have

$$\mathbb{E}[X_D] = \sum_{i=0}^{k-1} i p_D(i) = kx - \frac{\text{Var}[Z]}{k(1-x)}, \tag{6}$$

where X_D designates the number of individuals of type C in the same group as a randomly selected individual of type D.

The variance of Z can be interpreted as a measure of the tendency for individuals of the same type to aggregate in the same groups. Indeed, for two distinct individuals randomly selected from the same (also) randomly selected group, let us define

$$I_i = \begin{cases} 1 \text{ if individual } i \text{ is of type } C, \\ 0 \text{ if individual } i \text{ is of type } D, \end{cases} \tag{7}$$

for $i = 1, 2$. Then, the covariance between I_1 and I_2 is given by

$$\begin{aligned}
\mathrm{Cov}[I_1, I_2] &= \sum_{i=1}^{k} z_i \frac{i}{k} \frac{i-1}{k-1} - x^2 \\
&= \frac{\mathbb{E}[Z^2] - \mathbb{E}[Z]}{k(k-1)} - x^2 \\
&= \frac{\mathrm{Var}[Z]}{(k-1)k} - \frac{x(1-x)}{k-1}.
\end{aligned} \tag{8}$$

If this covariance is negative, individuals of the same type tend to be in different groups. On the other hand, if it is positive, they tend to flock to the same groups. It is equal to 0 in the neutral-assortment model.

3 Assortment of Types by Group Founders

In this section, we deduce a probability distribution for group types (as given by the random variable Z) from reasonable assumptions that produce an assortment of types within groups that is analogous to the leader-driven model in Kristensen et al. [24].

We assume that each group of size k is formed independently of all the others according to the following scheme. A first individual is sampled at random in the population to become the *founder* of the group. Then, each of the $k-1$ additional members of the group, independently of all the others, is selected in the population by the founder as follows. The founder chooses someone of the same type as his own with probability p, and someone at random in the whole population with the complementary probability $1 - p$. Individuals are said to be *assorted* if they are of matching types as a result of the preference exercised by the founder for its own type. Note that each additional member of a group is assorted to the founder with probability p and to each other additional member with probability p^2. Moreover, the case $p = 0$ corresponds to the neutral-assortment model. Note also that the assumptions are such that assortment is symmetric with respect to the types of the individuals so that it does not introduce a bias in favour of one type over the other in the whole population.

Let $Z_k(p)$ be the random variable that gives the type of a group that is formed as above. With probability given by x, the frequency of C in the population, the variable $Z_k(p)$ is distributed like $1 + X$ where X follows a binomial distribution with parameters $k - 1$ and $p + (1 - p)x$, and with probability $1 - x$, it is distributed like Y, a binomial random variable with parameters $k - 1$ and $(1 - p)x$. It is shown in Appendix A that

$$\frac{\mathbb{E}[Z_k(p)]}{k} = x \tag{9}$$

and

$$\mathrm{Var}[Z_k(p)] = v_k(p)x(1-x), \tag{10}$$

where

$$v_k(p) = (1 + (k-1)p)^2 + (k-1)(1 - p^2). \tag{11}$$

Note that the expected frequency of C in a group is the same as the frequency of C in the whole population. From this point forward, we always assume that the distribution of groups is given by the distribution of $Z_k(p)$.

Note that, in this case, the covariance between the types of two individuals chosen at random in the same group as defined in the previous subsection is given by

$$\text{Cov}[I_1, I_2] = r_k(p)x(1 - x), \tag{12}$$

where the coefficient of correlation

$$r_k(p) = \frac{\text{Cov}[I_1, I_2]}{\sqrt{\text{Var}[I_1]\text{Var}[I_2]}} = \frac{2}{k}p + \frac{k-2}{k}p^2 \tag{13}$$

is a measure of *pairwise assortment*. Moreover, the expected values of X_C and X_D take the forms

$$\mathbb{E}[X_C] = (k - 1)(x + (1 - x)r_k(p)) \tag{14}$$

and

$$\mathbb{E}[X_D] = (k - 1)x(1 - r_k(p)), \tag{15}$$

respectively.

Some properties of the pairwise assortment coefficient, easy to show, are worth mentioning:

- $r_k(p)$ is an increasing function with respect to $p \in [0, 1]$;
- $r_k(0) = 0$ and $r_k(1) = 1$;
- $r_k(p)$ is a decreasing function with respect to $k \geq 2$;
- $r_2(p) = p$ and $r_\infty(p) = \lim_{k \to \infty} r_k(p) = p^2$.

4 Payoffs to Interacting Individuals

Individuals interact within groups and receive some payoffs following these interactions. Let $\psi_C(j)$ be the average payoff received by an individual of type C in a group of type $j + 1$, and $\psi_D(j)$ be the average payoff received by an individual of type D in a group of type j, for $j = 0, 1, \ldots, k - 1$. Then, the average payoffs to C and D in the population where the frequency of C is x are

$$\pi_C(x) = \sum_{j=0}^{k-1} p_C(j)\psi_C(j) \tag{16}$$

and

$$\pi_D(x) = \sum_{j=0}^{k-1} p_D(j)\psi_D(j), \tag{17}$$

respectively. This gives

$$\bar{\pi}(x) = x\pi_C(x) + (1 - x)\pi_D(x) \tag{18}$$

as average payoff in the whole population. Note that this average payoff can be expressed as

$$\bar{\pi}(x) = \sum_{j=0}^{k-1} z_j \bar{\psi}(j), \tag{19}$$

where

$$\bar{\psi}(j) = \frac{j}{k}\psi_C(j-1) + \left(1 - \frac{j}{k}\right)\psi_D(j) \tag{20}$$

is the average payoff in a group of type j whose frequency is z_j for $j = 0, 1, \ldots, k$, with the convention that $\psi_C(-1) = \psi_D(k) = 0$.

As an important case, assume that the k individuals in a same group pairwise interact with each other as in Hilbe [19] and Lessard [30]. Let the payoffs following each interaction be given by the game matrix

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}. \tag{21}$$

Here, α and β are the payoffs to C against C and D, respectively, while γ and δ are the corresponding payoffs to D against C and D, respectively. In the Prisoner's dilemma, the strategies C and D stand for cooperation and defection, respectively, and we have the inequalities $\gamma > \alpha > \delta > \beta$. This is the case, for instance, in the additive model with $\alpha = b - c$, $\beta = -c$, $\gamma = b$ and $\delta = 0$, where c represents a cost for a cooperator and b a benefit for an individual in interaction with a cooperator. On the other hand, if the dilemma is repeated enough times and the payoffs are additive from one time to the next, however, then the cumulative payoffs to C and D for tit-for-tat starting with cooperation and always defect, respectively, satisfy the inequalities $\alpha > \gamma > \delta > \beta$ (see, e.g., Lessard [32], and references therein). These inequalities will be assumed throughout. They characterize a two-player stag hunt game (see, e.g., Archetti and Scheuring [1]). Note that the inequalities $\alpha > \beta$ and $\gamma > \delta$ mean that it always pays more to interact with C than with D, while $\alpha > \gamma$ and $\delta > \beta$ entails that it pays more to interact with an individual using the same strategy as its own rather than with an individual using a different strategy.

With the above payoff matrix in pairwise interactions, the average payoffs to C and D according to the number of interactions with C individuals in a group are given by

$$\psi_C(j) = \frac{\alpha j + \beta(k-1-j)}{k-1} \tag{22}$$

and

$$\psi_D(j) = \frac{\gamma j + \delta(k-1-j)}{k-1}, \tag{23}$$

respectively, for $j = 0, 1, \ldots, k-1$. In such a case, the average payoff to C in the population takes the form

$$\pi_C(x) = \frac{\alpha}{k-1}\mathbb{E}[X_C] + \frac{\beta}{k-1}(k-1-\mathbb{E}[X_C]), \tag{24}$$

and likewise

$$\pi_D(x) = \frac{\gamma}{k-1}\mathbb{E}[X_D] + \frac{\delta}{k-1}(k-1-\mathbb{E}[X_D]) \tag{25}$$

is the average payoff to D, where $\mathbb{E}[X_C]$ and $\mathbb{E}[X_D]$ are given in Eqs. (14) and (15), respectively.

A crucial quantity will be the difference between the average payoffs to C and D in the whole population. Under the assortment scheme described in the previous section, we have

$$\pi_C(x) = (\alpha - \beta)(x + (1-x)r_k(p)) + \beta \tag{26}$$

and

$$\pi_D(x) = (\gamma - \delta)x(1 - r_k(p)) + \delta. \tag{27}$$

Then, the difference is a linear function of x that can be expressed as

$$\pi_C(x) - \pi_D(x) = a_k(p)x + b_k(p), \tag{28}$$

where

$$a_k(p) = (\alpha - \beta - \gamma + \delta)(1 - r_k(p)) \tag{29}$$

and

$$b_k(p) = (\beta - \delta) + (\alpha - \beta)r_k(p). \tag{30}$$

Note that $\pi_C(x) - \pi_D(x)$ is an increasing function of x since $0 < r_k(p) < 1$ and $\alpha - \beta - \gamma + \delta > 0$ under the assumption $\alpha > \gamma > \delta > \beta$. Moreover, the zero of this function, namely $-b_k(p)/a_k(p),$ has a derivative with respect to $r_k(p)$ given by

$$-\frac{(\alpha - \delta)}{(\alpha - \beta - \gamma + \delta)(1 - r_k(p))^2} < 0. \tag{31}$$

Therefore, this zero is a decreasing function of $r_k(p)$.

Note also that we get the same average payoffs, and consequently the same difference between them, if the payoffs in random pairwise interactions in the whole population where the frequencies of C and D are x and $1 - x$, respectively, are given by the game matrix

$$\begin{pmatrix} \alpha & \beta(1 - r_k(p)) + \alpha r_k(p) \\ \gamma(1 - r_k(p)) + \delta r_k(p) & \delta \end{pmatrix} = (1 - r_k(p)) \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} + r_k(p) \begin{pmatrix} \alpha & \alpha \\ \delta & \delta \end{pmatrix}. \tag{32}$$

Such a payoff matrix, which can be traced back to Grafen [16], will be called an *effective game matrix* as in Lessard [30] and Lessard et al. [35] for analogous models.

5 Replicator Dynamics in an Infinite Population

The replicator equation is the basic model for the dynamics of strategy frequencies (Taylor and Jonker [50], see also Hofbauer and Sigmund [20], and references therein). The population is assumed to be infinite, time is continuous, and the instantaneous growth rate of a strategy is given by its current average payoff. In the case of two strategies C and D with frequencies x and $1 - x$, respectively, and corresponding average payoffs $\pi_C(x)$ and $\pi_D(x)$, respectively, the replicator equation takes the form

$$\dot{x} = x(\pi_C(x) - \bar{\pi}(x)) = x(1 - x)(\pi_C(x) - \pi_D(x)). \tag{33}$$

Note that, owing to Eq. (19), this equation for the group-structured population described above assumes *global selection* in the sense that the contributions of the different groups to the growth rate of the population are proportional to their average payoffs in the whole population. This assumption corresponds to *hard selection* in population genetics [9].

The replicator equation admits at least two equilibria, $x = 0$ and $x = 1$, and whether the frequency of type C is increasing or decreasing at any given time depends only on the sign of the difference between the average payoffs.

The fixation state $x = 1$ is *locally stable*, and then, C is an *evolutionarily stable strategy* (ESS) in the sense that it can resist invasion by D once fixed in the population [39], if we have $\pi_C(1) > \pi_D(1)$, which is equivalent here to

$$\alpha > \gamma + (\delta - \gamma)r_k(p). \tag{34}$$

On the other hand, the fixation state $x = 0$ is *locally unstable*, and then, D is not an ESS in the sense that it is invaded by C once fixed in the population, if $\pi_C(0) > \pi_D(0)$, which means here

$$\beta + (\alpha - \beta)r_k(p) > \delta. \tag{35}$$

Finally, if $\pi_C(1/2) > \pi_D(1/2)$, or equivalently here

$$\alpha + \beta - \gamma - \delta + (\alpha - \beta + \gamma - \delta)r_k(p) > 0, \tag{36}$$

then the basin of attraction of $x = 1$ is larger than the basin of attraction of $x = 0$. In this case, the strategy C is said to be *risk dominant* (RD) over D.

Note that, ignoring degenerate cases, the above sufficient conditions are also necessary. Moreover, if $\gamma > \delta$ and $\alpha > \beta$, meaning that it is always more advantageous to play against C than against D, then the above conditions become less stringent as $r_k(p)$ increases. In particular, they are all less stringent than in the neutral-assortment case which corresponds to $p = 0$.

6 Stochastic Dynamics in a Large Finite Population

In this section, we consider that the change in the frequency of C in competition with D in a large finite population is approximated by a continuous-time diffusion whose infinitesimal mean is given by

$$m(x) = sx(1 - x)(\pi_C(x) - \pi_D(x)) \tag{37}$$

and infinitesimal variance by $v(x) = x(1 - x)$. Here, the parameter $s \geq 0$ represents an intensity of selection with $s = 0$ corresponding to neutrality in the absence of selection.

With appropriate scalings of the selection parameters and unit of time with respect to the population size as this size goes to infinity, the diffusion approximation holds for a wide range of discrete-time models that includes in particular the Moran and Wright-Fisher models (see Appendix B). In the former, the updating of the population from one time step to the next takes place according to a birth-death event, while in the latter it occurs according to a binomial sampling scheme.

Owing to diffusion theory (see, e.g., Ewens [12]), the probability of ultimate fixation of C when represented once in a population of size N is then approximated as

$$F_C(1/N) \approx \frac{\int_0^{1/N} \psi(y)\mathrm{d}y}{\int_0^1 \psi(y)\mathrm{d}y}, \tag{38}$$

where

$$\psi(y) = \exp\left(-2\int_0^y \frac{m(x)}{v(x)}\mathrm{d}x\right) = \exp\left(-2s\int_0^y (\pi_C(x) - \pi_D(x))\mathrm{d}x\right). \tag{39}$$

This yields the approximation

$$F_C(1/N) \approx \frac{1}{N} + \frac{2s}{N} \int_0^1 \int_0^y (\pi_C(x) - \pi_D(x)) \mathrm{d}x \mathrm{d}y \qquad (40)$$

for s small enough.

We note that, for the linear case at hand with $\pi_C(x) - \pi_D(x)$ given in Eq. (28), we have

$$
\begin{aligned}
\int_0^1 \int_0^y (\pi_C(x) - \pi_D(x)) \mathrm{d}x \mathrm{d}y &= \int_0^1 \int_0^y (a_k(p)x + b_k(p)) \mathrm{d}x \mathrm{d}y \\
&= \frac{1}{2} \left(\frac{a_k(p)}{3} + b_k(p) \right) \\
&= \frac{1}{2} (\pi_C(1/3) - \pi_D(1/3)).
\end{aligned}
\qquad (41)
$$

This expression for the coefficient of $2s/N$ in the approximation of the fixation probability $F_C(1/N)$ leads us to the *one-third law of evolution*. This law, first introduced for a Moran model [41], was extended later on to the Wright–Fisher model [21, 28] and more generally to a large class of exchangeable selection models [33]. See Lessard [32] for more details.

First, let us say that C is *favoured by selection* to replace D in a population of finite size N if the probability for a single individual using strategy C to invade and take over a population of D players is larger under weak enough selection than what it would be under neutrality, that is, $F_C(1/N) > 1/N$. The one-third law states that this will be the case if $\pi_C(1/3) > \pi_D(1/3)$. In the case at hand, this condition becomes

$$\alpha + 2\beta - \gamma - 2\delta + (2\alpha - 2\beta + \gamma - \delta)r_k(p) > 0. \qquad (42)$$

Analogously, if $\pi_C(2/3) > \pi_D(2/3)$, then D is *disfavoured by selection* to replace C in the sense that the probability for a single D individual to invade a C population up to fixation is smaller under weak enough selection than under neutrality, which means that $F_D(1/N) < 1/N$. Here, this will occur when

$$2\alpha + \beta - 2\gamma - \delta + (\alpha - \beta + 2\gamma - 2\delta)r_k(p) > 0. \qquad (43)$$

Finally, C is *more favoured by selection* than D if the former fixation probability is larger than the latter, that is, $F_C(1/N) > F_D(1/N)$. It turns out that this will be the case here if

$$\alpha + \beta - \gamma - \delta + (\alpha - \beta + \gamma - \delta)r_k(p) > 0, \qquad (44)$$

which is the condition for C to be risk dominant over D. Note that this is also the condition for C to be more abundant than D on the average in the stationary state under symmetric mutation in the limit of a low mutation rate [13, 14].

Again here, all the above conditions become less stringent as $r_k(p)$ increases when $\alpha > \beta$ and $\gamma > \delta$.

7 Local Selection

If all groups contribute equally to the growth of an infinite population and the growth rates of C and D within each group are given by their average payoffs in this group defined in Eqs.

(22) and (23), then the time change in the frequency of C is described by the equation

$$\dot{x} = \sum_{j=0}^{k} z_j \frac{j}{k} (\psi_C(j-1) - \bar{\psi}(j)). \tag{45}$$

We recall that

$$\bar{\psi}(j) = \frac{j}{k} \psi_C(j-1) + \left(1 - \frac{j}{k}\right) \psi_D(j) \tag{46}$$

is the average payoff in a group of type j whose frequency is z_j for $j = 0, 1, \ldots, k$, with the convention that $\psi_C(-1) = \psi_D(k) = 0$. This yields

$$\dot{x} = \sum_{j=0}^{k} z_j \frac{j}{k} \left(1 - \frac{j}{k}\right) (\psi_C(j-1) - \psi_D(j)). \tag{47}$$

This equation has been called the *local replicator dynamics* [19]. This is the version of the replicator equation under the assumption of *local selection*, called *soft selection* in population genetics [9]. While the global replicator dynamics assumes that the growth rates of the individual types are given by their average payoffs in the whole population, the local replicator dynamics assumes that these are given by their average payoffs within groups weighted by their frequencies.

It can be shown (see Appendix C) that the above equation can be written in the form

$$\dot{x} = x(1-x)(\tilde{\pi}_C(x) - \tilde{\pi}_D(x)), \tag{48}$$

where $\tilde{\pi}_C(x)$ and $\tilde{\pi}_D(x)$ are the average payoffs to C and D, respectively, in the case of random pairwise interactions in the whole population with payoff matrix

$$\tilde{A}_k(p) = \begin{pmatrix} \tilde{\alpha}_k(p) & \tilde{\beta}_k(p) \\ \tilde{\gamma}_k(p) & \tilde{\delta}_k(p) \end{pmatrix}. \tag{49}$$

The entries of this effective game matrix are given by

$$\tilde{\alpha}_k(p) = \frac{k-2}{k}(1 - 2r_k(p) + s_k(p))\alpha, \tag{50a}$$

$$\tilde{\beta}_k(p) = \frac{1}{k}(1 - r_k(p))((k-1)\beta - \gamma) - \frac{k-2}{k}(r_k(p) - s_k(p))(\beta - \alpha + \gamma), \tag{50b}$$

$$\tilde{\gamma}_k(p) = \frac{1}{k}(1 - r_k(p))((k-1)\gamma - \beta) - \frac{k-2}{k}(r_k(p) - s_k(p))(\gamma - \delta + \beta), \tag{50c}$$

$$\tilde{\delta}_k(p) = \frac{k-2}{k}(1 - 2r_k(p) + s_k(p))\delta, \tag{50d}$$

where $r_k(p)$ is as defined in Eq. (13) and

$$s_k(p) = \frac{3}{k}p^2 + \frac{k-3}{k}p^3. \tag{51}$$

The quantity $s_k(p)$ can be interpreted as a measure of *tripletwise assortment* (see subsection 9.1 for more details). As for $\tilde{\pi}_C(x)$ and $\tilde{\pi}_D(x)$, they can be called the *effective average payoffs* to C and D, respectively.

It can also be shown (see Appendix B) that the corresponding discrete-time model in a population of finite size with this size as unit of time is described in the limit of a large

population by a diffusion process with

$$m(x) = sx(1 - x)(\tilde{\pi}_C(x) - \tilde{\pi}_D(x)) \tag{52}$$

as infinitesimal mean and $v(x) = x(1 - x)$ as infinitesimal variance, where $s \geq 0$ represents an intensity of selection.

Note that we have $r_k(p) \to p^2$ and $s_k(p) \to p^3$ as $k \to \infty$. Then, it is easy to check that

$$\lim_{k \to \infty} \tilde{A}_k(p) = (1 - p) \begin{pmatrix} (1 + p - p^2)\alpha & (1 + p - p^2)\beta + p^2(\alpha - \gamma) \\ (1 + p - p^2)\gamma + p^2(\delta - \beta) & (1 + p - p^2)\delta \end{pmatrix}. \tag{53}$$

In this limit, we have

$$\lim_{k \to \infty} (\tilde{\pi}_C(x) - \tilde{\pi}_D(x)) = (1 - p)^2(1 + 2p)(\alpha - \beta - \gamma + \delta)x$$
$$+ (1 - p)(1 + p - p^2)(\beta - \delta) + (1 - p)p^2(\alpha - \gamma), \tag{54}$$

which is an increasing linear function of x under the assumption that $\alpha > \gamma > \delta > \beta$. Therefore, we have the implications

$$\tilde{\pi}_C(1) > \tilde{\pi}_D(1) \tag{55a}$$

$$\Uparrow$$

$$\tilde{\pi}_C(2/3) > \tilde{\pi}_D(2/3) \tag{55b}$$

$$\Uparrow$$

$$\tilde{\pi}_C(1/2) > \tilde{\pi}_D(1/2) \tag{55c}$$

$$\Uparrow$$

$$\tilde{\pi}_C(1/3) > \tilde{\pi}_D(1/3) \tag{55d}$$

$$\Uparrow$$

$$\tilde{\pi}_C(0) > \tilde{\pi}_D(0) \tag{55e}$$

for k large enough. Moreover, the zero of the function in Eq. (54), namely

$$x_0(p) = \frac{(\delta - \beta)(1 - p)(1 + p - p^2) - (\alpha - \gamma)(1 - p)p^2}{(\alpha - \beta - \gamma + \delta)(1 - p)^2(1 + 2p)}, \tag{56}$$

has a first derivative with respect to p given by

$$\frac{\mathrm{d}x_0(p)}{\mathrm{d}p} = -\frac{p(2 + p)(\alpha + \beta - \gamma - \delta)}{(1 - p)^2(1 + 2p)^2(\alpha - \beta - \gamma + \delta)} \tag{57}$$

for $0 < p < 1$. This derivative is negative if $\alpha + \beta > \gamma + \delta$, but positive if $\alpha + \beta < \gamma + \delta$. This means that an increase in the probability of assortment relaxes all the above conditions for the evolution of C only if C is risk dominant over D. It is exactly the opposite if D is risk dominant over C.

Finally, comparing the zero of the function in Eq. (54) with the zero of the function in Eq. (28) as $k \to \infty$, we find that the former is greater than the latter, and then

$$\tilde{\pi}_C(x) > \tilde{\pi}_D(x) \Rightarrow \pi_C(x) > \pi_D(x) \tag{58}$$

for k large enough and $x \in [0, 1]$, if

$$(1 + p)(1 + p - p^2)(\beta - \delta) + (1 + p)p^2(\alpha - \gamma)$$

$$< (1 + 2p)(\beta - \delta) + (1 + 2p)p^2(\alpha - \beta), \tag{59}$$

which is equivalent to

$$p(\delta - \beta) + (1 + p)(\alpha - \gamma) < (1 + 2p)(\alpha - \beta). \tag{60}$$

This is always the case for $0 < p < 1$ under the assumption $\alpha > \gamma > \delta > \beta$, since then $\delta - \beta < \alpha - \beta$ and $\alpha - \gamma < \alpha - \beta$. This means that the conditions for the evolution of cooperation are easier to be satisfied under global selection than under local selection.

8 Randomness in Group Size and Probability of Assortment

If the group size k or the probability of assortment p between the group founder and each additional group member varies randomly from one group to another, then all the above results under global selection hold with the coefficient of pairwise assortment $r_k(p)$ replaced by its expected value

$$\mathbb{E}[r_k(p)] = \mathbb{E}[\mathbb{E}[r_k(p) \mid k, p]] = \mathbb{E}\left[\frac{2}{k}p + \frac{k-2}{k}p^2\right]. \tag{61}$$

When p is random and k constant, we have

$$\mathbb{E}[r_k(p)] = \frac{2}{k}\mathbb{E}[p] + \frac{k-2}{k}\mathbb{E}[p^2] \geq \frac{2}{k}\mathbb{E}[p] + \frac{k-2}{k}\mathbb{E}[p]^2, \tag{62}$$

from which

$$\mathbb{E}[r_k(p)] \geq r_k(\bar{p}), \tag{63}$$

where $\bar{p} = \mathbb{E}[p]$. Similarly, when k is random and p constant, we have

$$\mathbb{E}[r_k(p)] \geq r_{\bar{k}}(p), \tag{64}$$

where $\bar{k} = \mathbb{E}[k]$. This is ensured by Jensen's inequality for convex functions, since

$$\frac{\partial^2}{\partial k^2}r_k(p) = \frac{4p(1-p)}{k^3} \geq 0 \tag{65}$$

for $k > 0$ and $0 < p < 1$. Finally, when k and p are independent random variables, Jensen's inequality for conditional expectation leads to

$$\mathbb{E}[r_k(p)] = \mathbb{E}[\mathbb{E}[r_k(p) \mid p]] \geq \mathbb{E}[r_{\bar{k}}(p)] \geq r_{\bar{k}}(\bar{p}). \tag{66}$$

We conclude that stochastic variability in the group size and/or the probability of assortment increases the coefficient of pairwise assortment and consequently relaxes the conditions for the evolution of C under global selection when $\alpha > \beta$ and $\gamma > \delta$.

In the local selection model, the conditions for the evolution of C in Eq. (55) become

$$\mathbb{E}[\tilde{\pi}_C(x) - \tilde{\pi}_D(x)] > 0 \tag{67}$$

for $x = 1, 2/3, 1/2, 1/3, 0$, respectively, where the expected value is an increasing function of x under the assumption $\alpha > \gamma > \delta > \beta$ that can be expressed as

$$\mathbb{E}[(1-p)^2(1+2p)](\alpha - \beta - \gamma + \delta)x$$
$$+ \mathbb{E}[(1-p)(1+p-p^2)](\beta - \delta) + \mathbb{E}[(1-p)p^2](\alpha - \gamma) \tag{68}$$

in the limit of a large group size k owing to Eq. (54). Assuming that $\mathbb{E}[(p - \bar{p})^n] = o(\sigma^2)$ for $n \geq 3$ where $p = \mathbb{E}[p]$ and $\sigma^2 = \mathbb{E}[(p - \bar{p})^2]$, which holds for a wide range of probability distributions including the normal law, the zero of this function for σ^2 small enough can be approximated as

$$x_0(\bar{p}) - \frac{(\alpha + \beta - \gamma - \delta)(1 - 2\bar{p} - 2\bar{p}^2)}{(\alpha - \beta - \gamma + \delta)(1 - \bar{p})^3(1 + 2\bar{p})^2}\sigma^2 + o(\sigma^2), \tag{69}$$

where $x_0(\bar{p})$ is defined in Eq. (56). Under the assumption $\alpha > \gamma > \delta > \beta$, an increase in σ^2 decreases the value of this zero if C is risk dominant over D (that is, $\alpha + \beta > \gamma + \delta$) and $0 < \bar{p} < (\sqrt{3} - 1)/2$, or if D is risk dominant over C (that is, $\alpha + \beta < \gamma + \delta$) and $(\sqrt{3} - 1)/2 < \bar{p} < 1$. These are the conditions for an increase in σ^2 to relax the conditions for the evolution of C under local selection.

9 Discussion

9.1 Coefficients of Assortment

When everyone pairwise interacts with everyone else in a group of k individuals, including the founder of the group, then there are $k - 1$ interactions involving the founder with another member of the group and $(k - 1)(k - 2)/2$ interactions involving two members of the group different from the founder, for a total of $(k - 1)k/2$ interactions. With the former pairs of individuals being assorted with probability p and the latter with probability p^2, the probability for two interacting individuals chosen at random to be assorted is

$$\frac{k - 1}{(k - 1)k/2}p + \frac{(k - 1)(k - 2)/2}{(k - 1)k/2}p^2 = \frac{2}{k}p + \frac{k - 2}{k}p^2 = r_k(p), \tag{70}$$

which is the coefficient of correlation between the frequencies of C in two interacting individuals chosen at random. This is analogous in population genetics to Malécot's [37] coefficient of kinship with respect to identity-by-descent and Wright's [54] coefficient of relationship with respect to the frequency of a given allele. This corresponds to the coefficient of relatedness that was originally used by Hamilton [17, 18] in kin selection theory and later on extended to take into account more general situations such as inbreeding in diploid populations (see Lessard [27], and references therein).

Therefore, an evolutionary game with general payoffs in pairwise interactions within groups corresponds to a kin selection model with $r = r_k(p)$ as coefficient of relatedness between interacting individuals. However, the coefficient that comes into play here, called the *coefficient of pairwise assortment*, does not measure identity-by-descent between interacting individuals, which requires a common ancestor, nor identity-by-type, which may occur only by chance, but rather identity-by-assortment. This coefficient corresponds to the *index of assortativity* in Bergstrom [5]. Of course, positive assortment created by some preference for the same type can promote identity-by-type, but this concept is more general. Moreover, there may be a combination of identity-by-assortment and identity-by-descent. If, for instance, self-interactions are allowed within groups as in [19], then the coefficient of correlation above becomes

$$\frac{k}{k(k + 1)/2} + \frac{k - 1}{k(k + 1)/2}p + \frac{(k - 1)(k - 2)/2}{k(k + 1)/2}p^2, \tag{71}$$

where the first term is for the interactions between identical-by-descent individuals.

The above analysis for $r_k(p)$ can be extended to $s_k(p)$, which represents the probability for three individuals chosen at random in the same group to be simultaneously assorted. Indeed, there are $\binom{k-1}{2}$ triplets that include the group founder, in which case the probability that they are all assorted is p^2, and $\binom{k-1}{3}$ triplets that do not include the group founder, in which case the probability that they are all assorted is p^3, out of a total of $\binom{k}{3}$ triplets. Hence, we get

$$\frac{\binom{k-1}{2}}{\binom{k}{3}} p^2 + \frac{\binom{k-1}{3}}{\binom{k}{3}} p^3 = \frac{3}{k} p^2 + \frac{k-3}{k} p^3 = s_k(p). \tag{72}$$

Therefore, this is a *coefficient of tripletwise assortment*.

9.2 Effective Game Matrix

We have shown that the dynamics with pairwise interactions within groups involving assortment in finite as well as infinite populations is the same as the dynamics with pairwise interactions occurring at random in the whole population, but with the payoff matrix replaced by what has been called an *effective game matrix* by analogy with previous results involving relatedness built by partial dispersal [30].

Here, with groups contributing proportionally to their average payoffs to the time change of the population state, the entries of the effective matrix given in Eq. (32) are linear combinations of payoffs whose coefficients are the coefficient of pairwise assortment and its complement with respect to 1. This is in agreement with early results on the average payoffs in the hawk-dove game played between relatives [16] and a more recent analysis of the Prisoner's dilemma in a stochastic environment with additive cost and benefit for cooperation and a fixed coefficient of assortment between individuals in pairwise interactions (Lessard et al., 2021).

Note that our effective game matrix in the above context is different from the *derived matrix* for a two-player game between relatives given in Maynard Smith ([38], p. 192) which would take here the form

$$\begin{pmatrix} \alpha(1 + r_k(p)) & \beta + \gamma r_k(p) \\ \gamma + \beta r_k(p) & \delta(1 + r_k(p)) \end{pmatrix}. \tag{73}$$

Contrary to our effective matrix, this derived matrix does not yield in general the same difference $\pi_C(x) - \pi_D(x)$ as the one in Eq. (28) for the average payoffs to C and D when their frequencies in the whole population are x and $1 - x$, respectively. When $\alpha = b - c$, $\beta = -c, \gamma = b$ and $\delta = 0$ as in Nowak [42], however, we have $\pi_C(x) - \pi_D(x) = -c + br_k(p)$ for $0 \leq x \leq 1$ in both cases, which then leads to the same dynamics in finite as well as infinite populations. Moreover, this expression can be given an inclusive fitness interpretation [17, 18]. Such an interpretation can be used in linear approximations for two nearby mixed strategies to study convergence stability [10] based on initial increase in frequency of mutant strategies in infinite structured populations or their fixation probability when the structured populations are finite [46, 48, 49]. Inclusive fitness is not necessary for the analysis, however, and limitations of interpretations based on this concept must always be kept in mind (see, e.g., [35, 51]).

Note that the assumption of small differences between mixed strategies is different from the assumption of small differences between payoffs to pure strategies which does not generally lead to the same results (see, e.g., [34, 53]). Moreover, like evolutionary stability, convergence stability is a local property. In the case of cooperation, it may support an evolutionary tendency

toward a given level of cooperation from a neighbouring level but not necessarily from an absence of cooperation. Let us recall that adaptive dynamics relies on the assumptions of gradual changes of rare mutant strategies and global convergence as predicted from initial increase properties (see Avila and Mullon [2], and references therein).

With equal contributions of groups irrespective of their types, we still get an effective game matrix given in Eq. (49) but it is more complicated since it involves the coefficient of tripletwise assortment besides the coefficient of pairwise assortment. As discussed in Lessard [30] for group-structured populations with partial dispersal, this can be explained by competition within groups so that payoffs in pairwise interactions have an effect on individuals in the same group which are identical-by-assortment to the interacting individuals. Note that, in the additive model with $\alpha = b - c$, $\beta = -c$, $\gamma = b$ and $\delta = 0$, a calculation using the effective matrix leads to $\tilde{\pi}_C(x) - \tilde{\pi}_D(x) = -((k-1)c + b)(1 - r_k(p))/k$ for the difference between the average payoffs to C and D. The sign of this difference depends only on $(k-1)c + b$ in agreement with Rousset ([47], p. 125) and Lehman and Rousset [26] in the case of local regulation before dispersal in an island model.

9.3 Effect of Assortment and Randomness on the Evolution of Cooperation

It should be no surprise that assortment creates better conditions for cooperation to evolve than the absence of assortment. While cooperators may perform poorly in a mixed population, groups of cooperators are generally more successful than groups of defectors, and assortment promotes the abundance of homogeneous groups. As a matter of fact, if the contributions of groups are proportional to their average payoffs, then an increase in the assortment level relaxes all the conditions for the evolution of cooperation (C) against defection (D) in pairwise interactions, as in the case of tit-for-tat against always-defect in a Prisoner's dilemma repeated enough times, from the least stringent condition to the most stringent one: C locally stable, D disfavoured by selection to replace C, C more favoured by selection to replace D than the opposite, C favoured by selection to replace D, D locally unstable. Moreover, stochastic variability in the assortment level and/or the group size relaxes even more all these conditions.

When the contributions of groups are equal irrespective of their average payoffs rather than being proportional to those, however, an increase in the assortment level not only relaxes the conditions for the evolution of cooperation if C is risk dominant over D, but also strengthens these conditions if D is risk dominant over C. Contrary to the current belief, assortment may not always be an ally for the evolution of cooperation. As we have shown, *local selection instead of global selection can change the effect of assortment.* This phenomenon is in agreement with results in Rousset ([47], p. 125) and Lehman and Rousset [26] for an additive model. Moreover, we have shown that stochastic variability in the assortment level does not necessarily facilitate the evolution of C when selection is local.

9.4 Comments on the Assumptions

We have assumed that each group is formed and each group member chosen independently of all the other groups and other group members. This is a reasonable assumption if the group size is small compared to the population size, which makes sense if the population is infinite or finite but very large. Preference exercised by a group founder for other group members of the same type as his own is a natural way to model assortment within groups. Note that this implies that the individuals can recognize the types but do not know about the payoffs. The assortment between two members of the same group is lower, however, than the assortment

between each member and the founder. The coefficient of pairwise assortment that comes into play in our study is an average that depends on the group size. If we consider the assumption that the founder is not part of the group, then the coefficient of pairwise assortment within groups would be constant and given by p^2, where p is the probability for the founder to prefer a member of his own type. Similarly, the coefficient of tripletwise assortment within groups would then be p^3. Note that these values correspond to the values in the limit of a large group size without the aforementioned assumption.

Our results based on fixation probabilities in finite populations in the limit of a large population size have been obtained by diffusion approximations as in Lessard [28]. These approximations should hold for general exchangeable selection models as long as at most two lineages out of three coalesce at a time backwards in time with probability 1 in the limit of a large neutral population with updating according to a Cannings model [8], which is the domain of the Kingman coalescent [22] if an appropriate time scale is used (see Möhle [40], Lessard and Ladret [33]). Using a more general coalescent process, it can be shown, for instance, that C is favoured by selection to replace D if $\pi_C(q/3) > \pi_D(q/3)$, where

$$q = \frac{\lambda_{32}}{\lambda_{32} + \lambda_{31}} \tag{74}$$

with λ_{3j} being the rate of coalescence of j lineages out of 3, for $j = 1, 2$, in the limit of a large neutral population [29, 32]. Note that, while $q = 1$ for the Wright-Fisher and Moran models, we may have $q < 1$ when the reproductive success of individuals has a highly skewed probability distribution. Moreover, since $\pi_C(x) - \pi_D(x)$ is an increasing function of $x \in [0, 1]$ in our model, we have

$$\pi_C(q/3) > \pi_D(q/3) \Rightarrow \pi_C(1/3) > \pi_D(1/3). \tag{75}$$

This means that it is more difficult for selection to favour C replacing D when the variability in reproductive success is very high. This is also the case for the other conditions conducive to the evolution of cooperation.

9.5 Future Work

There are two obvious ways to extend the present study. The first one is to consider common good games instead of a two-player game. Although linear common good games yield the same dynamic as two-player games, nonlinear common good games (see, e.g., [1, 7, 15, 25, 31, 55]) may interact in new and interesting ways with assortment. As in Peña et al. [46] in the case of assortment built up by partial dispersal, we may expect an analogy with models without assortment but effective payoffs according to the composition of the group. The second one is to consider different schemes for group formation. Assortment by group founders is both natural and mathematically convenient, but assortment between individuals may arise in many other ways and involve other members of the group as described in Kristensen et al. [24].

Appendix A: Mean and Variance of the Group Type

Using the index 0 for the founder of a group and the indices $1, \ldots, k - 1$ for the other members in the same group, the group type can be expressed as

$$Z_k(p) = \sum_{i=0}^{k-1} I_i, \tag{76}$$

where

$$I_i = \begin{cases} 1 & \text{if individual } i \text{ is of type } C, \\ 0 & \text{if individual } i \text{ is of type } D, \end{cases} \tag{77}$$

for $i = 0, 1, \ldots, k-1$. Note that, given I_0, the random variables $I_1, \ldots, I_{k-1}$ are independent. Using the law of total expectation, we find

$$\begin{aligned} \mathbb{E}[Z_k(p)] &= \mathbb{E}[\mathbb{E}[Z_k(p)|I_0]] \\ &= \mathbb{E}[I_0(1 + (k-1)p + (1-p)(k-1)x) + (1-I_0)((1-p)(k-1)x)] \\ &= (1-p)(k-1)x + x(1 + (k-1)p) \\ &= kx. \end{aligned} \tag{78}$$

On the other hand, the law of total variance gives us the identity

$$\mathrm{Var}[Z_k(p)] = \mathrm{Var}[\mathbb{E}[Z_k(p)|I_0]] + \mathbb{E}[\mathrm{Var}[Z_k(p)|I_0]]. \tag{79}$$

The first right-hand side term in this identity is

$$\begin{aligned} \mathrm{Var}[\mathbb{E}[Z_k(p)|I_0]] &= \mathrm{Var}[I_0(1 + (k-1)p + (1-p)(k-1)x) + (1-I_0)((1-p)(k-1)x)] \\ &= (1 + (k-1)p)^2 \mathrm{Var}[I_0] \\ &= (1 + (k-1)p)^2 x(1-x), \end{aligned} \tag{80}$$

while the second one is

$$\mathbb{E}[\mathrm{Var}[Z_k(p)|I_0]] = \mathbb{E}\left[\sum_{j=1}^{k-1} \mathrm{Var}[I_j|I_0]\right] = \mathbb{E}\left[\sum_{j=1}^{k-1} \mathbb{E}[I_j|I_0] - \mathbb{E}[I_j|I_0]^2\right], \tag{81}$$

where

$$\mathbb{E}[I_j|I_0] = I_0(p + (1-p)x) + (1-I_0)(1-p)x = (1-p)x + I_0 p \tag{82}$$

for $j = 1, \ldots, k-1$. Therefore, we arrive at

$$\begin{aligned} &\mathbb{E}[\mathrm{Var}[Z_k(p)|I_0]] \\ &= (k-1)\mathbb{E}\left[(1-p)x - (1-p)^2 x^2 - I_0(p(1-p)x + (1-p)px + p^2 - p)\right] \\ &= (k-1)(1-p^2)x(1-x), \end{aligned} \tag{83}$$

so that we find

$$\mathrm{Var}[Z_k(p)] = \left((1 + (k-1)p)^2 + (k-1)(1-p^2)\right)x(1-x). \tag{84}$$

For moments of any order of $Z_k(p)$, we may resort to its moment-generating function given by

$$M(t) = \mathbb{E}[e^{tZ_k(p)}]$$

$$
\begin{aligned}
&= x e^t \mathbb{E}[e^{t \sum_{j=1}^{k-1} I_j} | I_0 = 1] + (1-x) \mathbb{E}[e^{t \sum_{j=1}^{k-1} I_j} | I_0 = 0] \\
&= x e^t \left(\mathbb{E}[e^{t I_1} | I_0 = 1] \right)^{k-1} + (1-x) \left(\mathbb{E}[e^{t I_1} | I_0 = 0] \right)^{k-1} \\
&= x e^t \left((1-p)(1-x) + (p + (1-p)x)e^t \right)^{k-1} \\
&\quad + (1-x) \left(1 - (1-p)x + (1-p)x e^t \right)^{k-1}.
\end{aligned}
\tag{85}
$$

In particular, this yields

$$
\mathbb{E}[Z_k(p)] = M^{(1)}(0) = kx, \tag{86a}
$$

$$
\mathbb{E}[Z_k(p)^2] = M^{(2)}(0) = [(1 + (k-1)p)^2 + (k-1)(1 - p^2)]x(1-x) + k^2 x^2, \tag{86b}
$$

$$
\begin{aligned}
\mathbb{E}[Z_k(p)^3] = M^{(3)}(0) &= x + (k-1)x(1 + (k-3)(k-2)p^3(1-x)(1-2x) \\
&\quad + 3(k-2)p^2(1-x)(2 + (k-4)x) \\
&\quad + 6p(1-x)(1 + (k-2)x) + kx(3 + (k-2)x)).
\end{aligned}
\tag{86c}
$$

Appendix B: Diffusion Approximation

As in Lessard [28], let the fitness of an individual of type C or D at any current type step in a population of fixed size N be defined as

$$
f_C(x) = 1 + \frac{s}{N} \pi_C(x) \tag{87}
$$

or

$$
f_D(x) = 1 + \frac{s}{N} \pi_D(x), \tag{88}
$$

respectively, where $\pi_C(x)$ and $\pi_D(x)$ are the average payoffs to C and D, respectively, $s \geq 0$ is an intensity of selection and x is the frequency of C at the current time step.

In a Wright-Fisher model, the updating from the current time step to the next, represented by times t and $t + \Delta t$, respectively, is done by making copies of N individuals sampled with replacement such that the probability of sampling each individual is proportional to its fitness. Therefore, the number of C individuals at the next time step, represented by $NX(t + \Delta t)$, given that this frequency of C at the current time step is $X(t) = x$, follows a binomial distribution with parameters N and

$$
\begin{aligned}
x' &= \frac{x f_C(x)}{x f_C(x) + (1-x) f_D(x)} \\
&= x + \frac{s}{N} x(1-x)(\pi_C(x) - \pi_D(x)) + O(s^2/N^2).
\end{aligned}
\tag{89}
$$

Consequently, taking N time steps as the unit of time so that $\Delta t = 1/N$, the change in the frequency of C from time t to time $t + \Delta t$, represented by $\Delta X(t) = X(t + \Delta t) - X(t)$, has conditional mean and variance given by

$$
\mathbb{E}[\Delta X_t | X_t = x] = sx(1-x)(\pi_C(x) - \pi_D(x))\Delta t + o(\Delta t) \tag{90}
$$

and

$$
\text{Var}[\Delta X_t | X_t = x] = x(1-x)\Delta t + o(\Delta t), \tag{91}
$$

respectively, while $\mathbb{E}[(\Delta X_t)^4 \mid X_t = x] = o(\Delta t)$. These conditions ensure convergence to a diffusion process for the frequency of C in the limit of a large population size with

$$m(x) = sx(1 - x)(\pi_C(x) - \pi_D(x)) \tag{92}$$

as infinitesimal mean and

$$v(x) = x(1 - x) \tag{93}$$

as infinitesimal variance (see, e.g., Ewens [12]).

The same result holds for a Moran model in which only one individual chosen at random is replaced from the current time step to the next, and this individual is replaced by a copy of one individual sampled in the population with probability proportional to its fitness. Then, the conditional expected change in the frequency of C is

$$\begin{aligned} \mathbb{E}[\Delta X_t \mid X_t = x] &= \frac{1}{N}((1 - x)x' - x(1 - x')) \\ &= \frac{s}{N^2}x(1 - x)(\pi_C(x) - \pi_D(x)) + O(s^2/N^3), \end{aligned} \tag{94}$$

while

$$\begin{aligned} \mathrm{Var}[\Delta X_t \mid X_t = x] &= \frac{1}{N^2}((1 - x)x' + x(1 - x')) - (\mathbb{E}[\Delta X_t \mid X_t = x])^2 \\ &= \frac{2}{N^2}x(1 - x) + O(s/N^3) \end{aligned} \tag{95}$$

and $\mathbb{E}[(\Delta X_t)^4 \mid X_t = x] = O(1/N^4)$. Taking $N^2/2$ time steps as the unit of time, we get the same diffusion as previously in the limit of a large population but with $\tilde{s} = s/2$ instead of s.

More generally, suppose that the numbers of descendants of the individuals having the same fitness from one time step to the next, perhaps including the individuals themselves, are exchangeable random variables [29, 32, 33]. Let the frequency of C at time $t + \Delta t$ be expressed as

$$X(t + \Delta t) = \frac{1}{N}\sum_{i=1}^{k} \nu_i \tag{96}$$

if there are n individuals of type C at time t and these leave $\nu_1, \ldots, \nu_n$ descendants at time $t + \Delta t$, respectively. Assuming the fitnesses of C and D as defined in the beginning of this appendix and an expected proportion ρ_N of the population replaced from one time step to the next irrespective of the types of the individuals, Eq. (89) implies that

$$\mathbb{E}[\nu_i] = 1 + \frac{s\rho_N}{N}(1 - x)(\pi_C(x) - \pi_D(x)) + O(s^2\rho_N/N^2) \tag{97}$$

for $i = 1, \ldots, n$. Note that 1 is the expected value under neutrality. As for the variances and covariances, let us assume that

$$\mathrm{Var}[\nu_i] = \sigma_N^2 + O(s\rho_N/N) \text{ and } \mathrm{Cov}[\nu_i, \nu_j] = -c_N + O(s\rho_N/N^2), \tag{98}$$

where $\sigma_N^2 = \mathrm{Var}_0[\nu_i] = \mathbb{E}_0[\nu_i(\nu_i - 1)] = O(1)$ and $c_N = \mathrm{Cov}_0[\nu_i, \nu_j] = O(1/N)$ are the corresponding values under neutrality ($s = 0$), for $i, j = 1, \ldots, n$ with $i \neq j$. Using the assumption that the numbers of descendants of the N individuals are exchangeable in the neutral model and the fact that they sum up to the constant N, we must have

$$N\sigma_N^2 - N(N - 1)c_N = 0, \tag{99}$$

from which

$$c_N = \frac{\sigma_N^2}{N-1}. \tag{100}$$

This leads us to a change $\Delta X(t) = X(t + \Delta t) - X(t)$ whose conditional mean and variance given that $X(t) = n/N = x$ take the forms

$$\mathbb{E}[\Delta X_t \mid X_t = x] = \frac{s\rho_N}{N} x(1-x)(\pi_C(x) - \pi_D(x)) + O(s^2\rho_N/N^2) \tag{101}$$

and

$$\text{Var}[\Delta X_t \mid X_t = x] = \frac{\sigma_N^2}{N} x(1-x) + O(s\rho_N/N^2), \tag{102}$$

respectively. With N/σ_N^2 time steps as unit of time, which defines the *effective population size*, we get the same infinitesimal mean and variance as previously in the limit of a large population but with intensity of selection

$$\tilde{s} = s \lim_{N \to \infty} \frac{\rho_N}{\sigma_N^2} \tag{103}$$

instead of s. Note that $\rho_N = 1$ and $\sigma_N^2 = 1 - 1/N$ in the Wright-Fisher model, while $\rho_N = 1/N$ and $\sigma_N^2 = 2/N$ in the Moran model. Note also that we assume the supplementary condition $\mathbb{E}[(\Delta X_t)^4 \mid X_t = x] = o(\sigma_N^2/N)$ in order to ensure the diffusion approximation. This condition is expected to hold if

$$\lim_{N \to \infty} \frac{\sum_{i=1}^N \mathbb{E}_0\left[\frac{v_i(v_i-1)(v_i-2)}{N(N-1)(N-2)}\right]}{\sum_{i=1}^N \mathbb{E}_0\left[\frac{v_i(v_i-1)}{N(N-1)}\right]} = \lim_{N \to \infty} \frac{\mathbb{E}_0[v_1(v_1-1)(v_1-2)]}{N\mathbb{E}_0[v_1(v_1-1)]} = 0, \tag{104}$$

which is the necessary and sufficient condition for a diffusion approximation in the neutral model [40]. This condition says that the probability for three distinct individuals randomly chosen in the population at any given time step to have the same parent one time step back is negligible compared to the corresponding probability for two individuals in the limit of a large neutral population. This is the case if the probability distribution for the number of descendants of a single individual is not too highly skewed.

In the case of competition within groups, the fitnesses of C and D in a group of type j are given by

$$f_C(j-1) = 1 + \frac{s}{N} \psi_C(j-1) \tag{105}$$

and

$$f_D(j) = 1 + \frac{s}{N} \psi_D(j), \tag{106}$$

respectively, where $\psi_C(j-1)$ and $\psi_D(j)$ are the average payoffs to C and D, respectively, defined in Eqs. (22) and (23) for $j = 0, 1, \ldots, k$ with the convention that $\psi_C(-1) = \psi_D(k) = 0$. Then, the probability for an offspring to be of type C is

$$\sum_{j=0}^k z_j \frac{\frac{j}{k} f_C(j-1)}{\frac{j}{k} f_C(j-1) + \left(1 - \frac{j}{k}\right) f_D(j)}, \tag{107}$$

where z_j is the frequency of groups with j individuals of type C and $k - j$ individuals of type D for $j = 0, 1, \ldots, k$. The above probability takes the form

$$\sum_{j=0}^{k} z_j \frac{j}{k} + \frac{s}{N} \sum_{j=0}^{k} z_j \frac{j}{k} \left(1 - \frac{j}{k}\right) (\psi_C(j-1) - \psi_D(j)) + O(s^2/N^2). \tag{108}$$

This can be expressed as

$$x + \frac{s}{N} x(1-x)(\tilde{\pi}_C(x) - \tilde{\pi}_D(x)) + O(s^2/N^2), \tag{109}$$

where

$$x = \sum_{j=0}^{k} z_j \frac{j}{k} = \frac{\mathbb{E}[Z_k(p)]}{k} \tag{110}$$

is the current frequency of C in the whole population, while $\tilde{\pi}_C(x)$ and $\tilde{\pi}_D(x)$ are effective average payoffs to C and D, respectively, in this population (see Appendix C). Moreover, the above diffusion approximation in the limit of a large population holds with

$$m(x) = \tilde{s}x(1-x)(\tilde{\pi}_C(x) - \tilde{\pi}_D(x)) \tag{111}$$

as infinitesimal mean.

Appendix C: Effective Average Payoffs Under Local Selection

Using the expression

$$\psi_C(j-1) - \psi_D(j) = \alpha \left(\frac{j-1}{k-1}\right) + \beta \left(\frac{k-j}{k-1}\right) - \gamma \left(\frac{j}{k-1}\right) - \delta \left(\frac{k-1-j}{k-1}\right) \tag{112}$$

from Eqs. (22) and (23) for $j = 1, \ldots, k-1$, the local replicator dynamics given in Eq. (47) can be written in the form

$$\dot{x} = \alpha S_\alpha + \beta S_\beta - \gamma S_\gamma - \delta S_\delta, \tag{113}$$

where

$$S_\alpha = \sum_{j=0}^{k} z_j \frac{j-1}{k-1} \frac{j}{k} \left(1 - \frac{j}{k}\right)$$
$$= -\frac{1}{k^2(k-1)} \mathbb{E}[Z_k(p)^3] + \frac{k+1}{(k-1)k^2} \mathbb{E}[Z_k(p)^2] - \frac{1}{(k-1)k} \mathbb{E}[Z_k(p)], \tag{114a}$$

$$S_\beta = \sum_{j=0}^{k} z_j \frac{k-j}{k-1} \frac{j}{k} \left(1 - \frac{j}{k}\right)$$
$$= \frac{1}{(k-1)k^2} \mathbb{E}[Z_k(p)^3] - \frac{2}{k(k-1)} \mathbb{E}[Z_k(p)^2] + \frac{1}{k-1} \mathbb{E}[Z_k(p)], \tag{114b}$$

$$S_\gamma = \sum_{j=0}^{k} z_j \frac{j}{k-1} \frac{j}{k} \left(1 - \frac{j}{k}\right)$$

$$= -\frac{1}{(k-1)k^2}\mathbb{E}[Z_k(p)^3] + \frac{1}{(k-1)k}\mathbb{E}[Z_k(p)^2], \tag{114c}$$

$$S_\delta = \sum_{j=0}^{k} z_j \frac{k-1-j}{k-1} \frac{j}{k}\left(1 - \frac{j}{k}\right)$$

$$= \frac{1}{(k-1)k^2}\mathbb{E}[Z_k(p)^3] - \frac{2k-1}{(k-1)k^2}\mathbb{E}[Z_k(p)^2] + \frac{1}{k}\mathbb{E}[Z_k(p)]. \tag{114d}$$

From Eq. (86) in Appendix A, we know that $\mathbb{E}[Z_k(p)] = kx$, while the second and third moments given there can be expressed as

$$\mathbb{E}[Z_k(p)^2] = kx(1-x)(1 + (k-1)r_k(p)) + x^2 k^2 \tag{115}$$

and

$$\mathbb{E}[Z_k(p)^3] = k\,(x + (k-1)x(3 + (k-2)x) - (k-1)x(1-x)((k-2)(2x-1)s_k(p)$$
$$-3(1 + (k-2)x)r_k(p) + (3 + (k-2)x))), \tag{116}$$

respectively, where $r_k(p)$ and $s_k(p)$ are given in Eqs. (13) and (51), respectively. Then, straightforward calculations lead to

$$S_\alpha = x(1-x)\frac{k-2}{k}\left((1 - 2r_k(p) + s_k(p))x + (r_k(p) - s_k(p))(1-x)\right), \tag{117a}$$

$$S_\beta = x(1-x)\frac{1}{k}\left((1 + (k-3)r_k(p) - (k-2)s_k(p))x \right.$$
$$\left. + (k - 1 - (2k-3)r_k(p) + (k-2)s_k(p))(1-x)\right), \tag{117b}$$

$$S_\gamma = -x(1-x)\frac{1}{k}\left((k - 1 + (2k-3)r_k(p) - (k-2)s_k(p))x \right.$$
$$\left. - (1 + (k-3)r_k(p) - (k-2)s_k(p))(1-x)\right), \tag{117c}$$

$$S_\delta = -x(1-x)\frac{k-2}{k}((s_k(p) - r_k(p))x - (1 - 2r_k(p) + s_k(p))(1-x)). \tag{117d}$$

Algebraic manipulations show that Eq. (113) can be expressed as

$$\dot{x} = x(1-x)(\tilde{\pi}_C(x) - \tilde{\pi}_D(x)), \tag{118}$$

where

$$\tilde{\pi}_C(x) = \tilde{\alpha}_k(p)x + \tilde{\beta}_k(p)(1-x), \tag{119a}$$

$$\tilde{\pi}_D(x) = \tilde{\gamma}_k(p)x + \tilde{\delta}_k(p)(1-x), \tag{119b}$$

with

$$\tilde{\alpha}_k(p) = \frac{k-2}{k}(1 - 2r_k(p) + s_k(p))\alpha, \tag{120a}$$

$$\tilde{\delta}_k(p) = \frac{k-2}{k}(1 - 2r_k(p) + s_k(p))\delta, \tag{120b}$$

and then necessarily

$$\tilde{\beta}_k(p) = \frac{1}{k}(1 - r_k(p))((k-1)\beta - \gamma) - \frac{k-2}{k}(r_k(p) - s_k(p))(\beta - \alpha + \gamma), \tag{121a}$$

$$\tilde{\gamma}_k(p) = \frac{1}{k}(1 - r_k(p))((k-1)\gamma - \beta) - \frac{k-2}{k}(r_k(p) - s_k(p))(\gamma - \delta + \beta). \tag{121b}$$

References

1. Archetti M, Scheuring I (2012) Review: game theory of public goods in one-shot social dilemmas without assortment. J Theor Biol 299:9–20
2. Avila P, Mullon C (2023) Evolutionary game theory and the adaptive dynamics approach: adaptation where individuals interact. Philos Trans Roy Soc B Biol Sci 378:3456. https://doi.org/10.1098/rstb.2021.0502
3. Axelrod R (1984) The evolution of cooperation. Basic Books, New York
4. Axelrod R, Hamilton WD (1981) The evolution of cooperation. Science 211:1390–1396
5. Bergstrom T (2003) The algebra of assortative encounters and the evolution of cooperation. Int Game Theory Rev 5:1–18
6. Boyd R, Richerson PJ (1988) The evolution of reciprocity in sizable groups. J Theor Biol 132:337–356
7. Broom M, Pattni K, Rychtář J (2019) Generalized social dilemmas: the evolution of cooperation in populations with variable group size. Bull Math Biol 81:4643–4674
8. Cannings C (1974) The latent root of certain Markov chains arising in genetics: a new approach. 1. Haploid models. Adv Appl Probab 6:260–290
9. Christiansen FB (1975) Hard and soft selection in a subdivided population. Am Nat 109:1–16
10. Christiansen FB (1991) On conditions for evolutionary stability for a continuously varying character. Am Nat 138:37–50
11. Eshel I, Cavalli-Sforza LL (1982) Assortment of encounters and evolution of cooperativeness. Proc Natl Acad Sci USA 79:1331–1335
12. Ewens WJ (2004) Mathematical population genetics: I. Theoretical introduction. Springer, New York
13. Fudenberg D, Imhof LA (2006) Imitation processes with small mutations. J Econ Theory 131:251–262
14. Fudenberg D, Nowak MA, Taylor C, Imhof LA (2006) Evolutionary game dynamics in finite populations with strong selection and weak mutation. Theor Popul Biol 70:352–363
15. Gokhale CS, Traulsen A (2010) Evolutionary games in the multiverse. Proc Natl Acad Sci USA 107:5500–5504
16. Grafen A (1979) The hawk-dove game played between relatives. Anim Behav 27:905–907
17. Hamilton WD (1964) The genetical evolution of social behaviour I. J Theor Biol 7:1–16
18. Hamilton WD (1970) Selfish and spiteful behaviour in an evolutionary model. Nature 228:1218–1220
19. Hilbe C (2011) Local replicator dynamics: a simple link between deterministic and stochastic models of evolutionary game theory. Bull Math Biol 73:2068–2087
20. Hofbauer J, Sigmund K (1998) The theory of evolution and dynamical systems. Cambridge University Press, Cambridge
21. Imhof LA, Nowak MA (2006) Evolutionary game dynamics in a Wright–Fisher process. J Math Biol 52:667–681
22. Kingman JFC (1982) The coalescent. Stoch Process Appl 13:235–248
23. Krivan V, Cressman R (2017) Interaction times change evolutionary outcomes: two-player matrix games. J Theor Biol 416:199–207
24. Kristensen NP, Ohtsuki H, Chisholm RA (2022) Ancestral social environments plus nonlinear benefits can explain cooperation in human societies. Sci Rep 12:20252
25. Kurokawa S, Ihara Y (2009) Emergence of cooperation in public goods games. Proc R Soc B 276:1379–1384
26. Lehmann L, Rousset F (2010) How life history and demography promote or inhibit the evolution of helping behaviours. Philos Trans R Soc B 365:2599–2617
27. Lessard S (1992) Relatedness and inclusive fitness with inbreeding. Theor Popul Biol 42:284–307
28. Lessard S (2005) Long-term stability from fixation probabilities in finite populations: new perspectives for ESS theory. Theor Popul Biol 68:19–27
29. Lessard S (2007) Cooperation is less likely to evolve in a finite population with a highly skewed distribution of family size. Proc R Soc B 274:1861–1865
30. Lessard S (2011) Effective game matrix and inclusive payoff in group-structured populations. Dyn Games Appl 1:301–318
31. Lessard S (2011) On the robustness of the extension of the one-third law of evolution to the multi-player game. Dyn Games Appl 1:408–418
32. Lessard S (2011) Evolution of cooperation in finite populations. In: Sigmund K (ed) Evolutionary game dynamics. American Mathematical Society, Providence, pp 143–171
33. Lessard S, Ladret V (2007) The probability of fixation of a single mutant in an exchangeable selection model. J Math Biol 54:721–744
34. Lessard S, Rocheleau G (2004) Kin selection and coefficients of relatedness in family-structured populations with inbreeding. Theor Popul Biol 66:287–306

35. Lessard S, Li C, Zheng XD, Tao Y (2021) Inclusive fitness and Hamilton's rule in a stochastic environment. Theor Popul Biol 142:91–99
36. Li C, Lessard S (2021) The effect of the opting-out strategy on conditions for selection to favor the evolution of cooperation in a finite population. J Theor Biol 510:110543
37. Malécot G (1948) Les mathématiques de l'hérédité. Masson, Paris
38. Maynard Smith J (1982) Evolution and the theory of games. Cambridge University Press, Cambridge
39. Maynard Smith J, Price GR (1973) The logic of animal conflict. Nature 246:15–18
40. Möhle M (2001) Forward and backward diffusion approximations for haploid exchangeable population models. Stoch Process Appl 95:133–149
41. Nowak MA, Sasaki A, Taylor C, Fudenberg D (2004) Emergence of cooperation and evolutionary stability in finite populations. Nature 428:646–650
42. Nowak MA (2006) Five rules for the evolution of cooperation. Science 314:1560–1563
43. Ohtsuki H (2010) Evolutionary games in Wright's island model: kin selection meets evolutionary game theory. Evolution 64:3344–3353
44. Ohtsuki H (2014) Evolutionary dynamics of n-player games played by relatives. Philos Trans Roy Soc B Biol Sci 369:20130359
45. Peña J, Nöldeke G (2016) Variability in group size and the evolution of collective action. J Theor Biol 389:72–82
46. Peña J, Nöldeke G, Lehmann L (2015) Evolutionary dynamics of collective action in spatially structured populations. J Theor Biol 382:122–136
47. Rousset F (2004) Genetic structure and selection in subdivided populations. Princeton University Press, Princeton
48. Rousset F, Billiard S (2000) A theoretical basis for measures of kin selection in subdivided populations: finite populations and localized dispersal. J Evol Biol 13:814–825
49. Taylor PD, Day T, Wild G (2007) From inclusive fitness to fixation probability in homogeneous structured populations. J Theor Biol 249:101–110
50. Taylor PD, Jonker L (1978) Evolutionary stable strategies and game dynamics. Math Biosci 40:145–156
51. van Veelen M (2009) Group selection, kin selection, altruism and cooperation: when inclusive fitness is right and when it can be wrong. J Theor Biol 259:589–600
52. vanVeelen M, García J, Rand DG, Nowak MA (2012) Direct reciprocity in structured populations. Proc Natl Acad Sci USA 109:9929–9934
53. Wild G, Traulsen A (2007) The different limits of weak selection and the evolutionary dynamics of finite populations. J Theor Biol 247:382–390
54. Wright S (1922) Coefficient of inbreeding and relationship. Am Nat 56:330–338
55. Wu B, Traulsen A, Gokhale CS (2013) Dynamic properties of evolutionary multi-player games in finite populations. Games 4:182–199
56. Zhang BY, Fan SJ, Li C, Zheng XD, Bao JZ, Cressman R, Tao Y (2016) Opting out against defection leads to stable coexistence with cooperation. Sci Rep 6:35902