

Article

# Evolution of Groupwise Cooperation: Generosity, Paradoxical Behavior, and Non-Linear Payoff Functions

Shun Kurokawa <sup>1,2,3,4,\*</sup>, Joe Yuichiro Wakano <sup>5</sup> and Yasuo Ihara <sup>1</sup>

<sup>1</sup> Department of Biological Sciences, The University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-0033, Japan; iharay@bs.s.u-tokyo.ac.jp

<sup>2</sup> Division of Natural Resource Economics, Graduate School of Agriculture, Kyoto University, Oiwake-cho, Kitashirakawa, Sakyo-ku, Kyoto 606-8502, Japan

<sup>3</sup> Key Lab of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Science, Datun Road, Chaoyang, Beijing 100101, China

<sup>4</sup> School of Economics and Management, Kochi University of Technology, Kochi 780-8515, Japan

<sup>5</sup> Meiji Institute for Advanced Study of Mathematical Sciences, Meiji University, Nakano 4-21-1, Nakano-ku, Tokyo 164-8525, Japan; joewakano@gmail.com

\* Correspondence: kurokawa.shun@kochi-tech.ac.jp

Received: 29 September 2018; Accepted: 13 November 2018; Published: 10 December 2018



**Abstract:** Evolution of cooperation by reciprocity has been studied using two-player and  $n$ -player repeated prisoner's dilemma games. An interesting feature specific to the  $n$ -player case is that players can vary in generosity, or how many defections they tolerate in a given round of a repeated game. Reciprocators are quicker to detect defectors to withdraw further cooperation when less generous, and better at maintaining a long-term cooperation in the presence of rare defectors when more generous. A previous analysis on a stochastic evolutionary model of the  $n$ -player repeated prisoner's dilemma has shown that the fixation probability of a single reciprocator in a population of defectors can be maximized for a moderate level of generosity. However, the analysis is limited in that it considers only tit-for-tat-type reciprocators within the conventional linear payoff assumption. Here we extend the previous study by removing these limitations and show that, if the games are repeated sufficiently many times, considering non-tit-for-tat type strategies does not alter the previous results, while the introduction of non-linear payoffs sometimes does. In particular, under certain conditions, the fixation probability is maximized for a "paradoxical" strategy, which cooperates in the presence of fewer cooperating opponents than in other situations in which it defects.

**Keywords:** evolutionary game theory; cooperation; generosity; repeated game; prisoner's dilemma

## 1. Introduction

Reciprocity is a key factor in the study of evolution of cooperation [1,2]. Evolution of pairwise and groupwise cooperation by reciprocity has been studied by using the two-player and  $n$ -player repeated prisoner's dilemma (PD) games [2–12]. While a population of reciprocators can be stable against invasion by defectors in repeated PDs, so long as games are repeated sufficiently many times, various additional mechanisms have been proposed to explain the initial emergence of reciprocators in a population of defectors, such as invasion by a mass of reciprocators [13], formation of spatial clusters [14,15], stochastic aggregate payoffs [16], and random drift [17]. Of particular relevance to the present study is Nowak et al. (2004). They developed a model of stochastic evolutionary dynamics for the two-player repeated PD and derived the fixation probability with which a single mutant of reciprocator that appears in a population of defectors will eventually replace the whole

population. Nowak et al.'s (2004) framework has been extended to include the more general  $n$ -player repeated PD [18,19], which has enabled to investigate the evolution of group-wise cooperation [20–22] considering the effect of random drift. In this article, we report our further investigation on the stochastic evolutionary model of the  $n$ -player repeated PD.

An interesting feature of group-wise cooperation, or more specifically, the  $n$ -player repeated PD, is that we can conceive of multiple types of reciprocators that differ in their levels of generosity. Take tit-for-tat (TFT) for example. TFT is a reciprocal strategy in the two-player repeated PD, which cooperates in the first round of a repeated game and, from the second round onward, cooperates if and only if its opponent has cooperated in the previous round. In the  $n$ -player repeated PD,  $n$  different reciprocal strategies analogous to TFT, or  $TFT_a$  for  $a \in \{0, 1, \dots, n-1\}$ , are possible, where  $a$  represents the strategy's level of generosity [23]. Namely,  $TFT_a$  cooperates in the first round and after that cooperates if and only if  $a$  or more of its opponents have cooperated in the previous round. Thus,  $TFT_a$  with a smaller value of  $a$  is more generous toward the group members' defection.

Suppose that a single mutant of the strategy that always defects (ALLD) appears in a population of  $TFT_a$ . Obviously, less generous  $TFT_a$  (i.e., having larger  $a$ ) is quicker to find ALLDs in the group and withdraw further cooperation and, thus, more likely to be stable against their invasion. On the other hand, for  $TFT_a$  to be advantageous over ALLD, it is always necessary that each  $TFT_a$  individual tends to have  $a$  or more other  $TFT_a$  individuals in the same group; otherwise,  $TFT_a$  cooperates only in the first round of each game to be exploited by ALLDs. This means that more generous  $TFT_a$  (i.e., having smaller  $a$ ) may compete better against ALLD, particularly when  $TFT_a$  is rare in the population; in other words, generosity may serve to facilitate the initial emergence of  $TFT_a$ .

This intuition has been confirmed by Kurokawa et al. (2010), who compared the fixation probabilities of  $TFT_a$ s with different values of  $a$  when appears as a single mutant in a population of ALLD [19]. They found that more generous  $TFT_a$  can sometimes have a greater fixation probability than less generous ones. They also specified the optimal level of generosity that maximizes the fixation probability, showing that there is a certain threshold level of generosity above which the largest fixation probability can never be attained. Specifically, it was shown that any such  $TFT_a$  that tolerates defections by a half or more of its opponents can never have the largest fixation probability.

Although Kurokawa et al. (2010) focused on the strategies classified as  $TFT_a$ , they are not the only ones that can potentially establish mutual cooperation in the  $n$ -player repeated PD. In fact, tolerating the presence of some defectors in a group may have a similar enhancing effect on the evolution of reciprocal strategies in general. However, this possibility has not been explored thus far. Here we extend Kurokawa et al.'s (2010) analysis to consider the set of all "reactive strategies" that (i) cooperates with probability  $d_\phi \in \{0, 1\}$  in the first round and (ii) cooperates in the  $k$ th round ( $k \geq 2$ ) with probability  $d_j \in \{0, 1\}$ , which depends on the number of opponents,  $j$ , who have cooperated in the  $k-1$ th round. There are  $2^{n+1}$  such strategies in total, including  $n-1$   $TFT_a$ s and  $2^n$  strategies that defect in the first move [24–28]. The rest of the strategies in the set are "paradoxical" in the sense that each of them has a certain situation in which it cooperates even though there are fewer cooperators in the group than in other situations in which it withdraws cooperation. It is of interest whether the strategies defecting in the first move or exhibiting paradoxical behavior can have the largest fixation probability among the set of all reactive strategies when invading ALLD.

Another limitation of Kurokawa et al.'s (2010) work is that they only considered the conventional linear payoff functions (i.e., linear public goods games). However, there are at least two reasons why non-linear payoff functions should be also considered. First, in the real world, payoffs are not always linear [29–39] and it has been demonstrated that introduction of non-linearity can affect the outcome of evolutionary models (e.g., [40–47]). Second, if payoffs increase non-linearly with the number of cooperating members in the group, it can be said that the efficiency of an act of cooperation varies depending on the number of cooperators. For example, in case payoffs increase more rapidly when cooperators are fewer, an individual's act of cooperation is more efficient in the presence of fewer cooperators. In such a case, more generous strategies may be selectively favored because

generosity promotes cooperation when it is efficient. To examine these possibilities, we also extend Kurokawa et al. (2010) in terms of payoff functions.

In addition, Kurokawa et al. (2010) derived their results under the assumption that selection is sufficiently weak. However, the intensity of selection can affect the outcomes of evolutionary models (e.g., [17,48–50]). Hence, we investigate to what extent our results based on the weak selection assumption may be affected by the selection intensity.

In this paper, we extend Kurokawa et al.'s (2010) analysis by removing these limitations. In particular, taking non-linear payoff functions and paradoxical (i.e., non-tit-for-tat type) strategies into consideration, we ask how non-linear payoffs may facilitate the evolution of excessive generosity and whether any paradoxical strategies can ever attain the highest fixation probability among all reactive strategies considered. Additionally, we examine how our results may depend on the intensity of natural selection. In what follows, we first introduce the  $n$ -player repeated PD and our model of stochastic evolutionary dynamics (Model). Then we derive, based on the weak selection approximation, the best reactive strategies, or the strategies maximizing the fixation probability, and provide detailed analyses for linear and non-linear payoff cases. We also examine to what extent our results may be affected if selection is not sufficiently weak (Results). Finally, we summarize the results and discuss some caveats and possibilities of further investigations (Discussion).

## 2. Model

We investigate stochastic evolutionary dynamics of a population of individuals whose fitness is determined by  $n$ -player repeated PD games. We consider a set of “reactive strategies,” as defined below, and compare the fixation probabilities of different reactive strategies when introduced as a mutant into a population of the strategy that always defects, or ALLD. Our goal is to specify the “best reactive strategies,” which maximize the fixation probability.

### 2.1. The $n$ -Player Repeated Prisoner's Dilemma

In the  $n$ -player repeated PD, a group of  $n$  individuals play a game consisting of one or more rounds, in each of which each individual chooses to either cooperate or defect. After one round is finished, another round will be played with probability  $\delta$ ; otherwise, the group will be dismissed ( $0 < \delta < 1$ ). Thus, the expected number of rounds is  $1/(1 - \delta)$ . Suppose that there are  $k$  cooperating and  $n - k$  defecting individuals in a given round. The payoffs to cooperating (C) and defecting (D) individuals in that round are denoted by  $V(C|k)$  and  $V(D|k)$ , respectively ( $0 \leq k \leq n$ ). Note that  $V(C|0)$  and  $V(D|n)$  are defined arbitrarily and used only for the sake of notational convenience. Following Boyd & Richerson (1988),  $n$ -player PD requires the following four conditions to be satisfied:

$$V(C|k+1) < V(D|k), \quad (1)$$

$$V(C|k) < V(C|k+1), \quad (2)$$

$$V(D|k) < V(D|k+1), \quad (3)$$

$$(n - k - 1)V(D|k+1) + (k + 1)V(C|k+1) > (n - k)V(D|k) + kV(C|k), \quad (4)$$

where  $0 \leq k \leq n - 1$ . Conditions (1) states that an individual always gains more from defecting than cooperating. Conditions (2) and (3) mean that an individual always gains more when there are more cooperators in the group. Conditions (4) indicates that the total payoff of the group is larger when there are more cooperators in the group.

We define a reactive strategy in the  $n$ -player repeated PD as follows: an individual with a reactive strategy cooperates with a certain probability in the first round, and from the second round on, cooperates with a certain probability determined by the number of opponents (i.e., group members other than the self) who have cooperated in the preceding round. Any reactive strategy is described by a vector  $\mathbf{d} = (d_\varphi, d_0, d_1, \dots, d_{n-1})$ , where  $d_\varphi$  denotes the probability with which an individual

obeying this strategy cooperates in the first move and  $d_j$  denotes the probability with which he cooperates in a given round, provided that  $j$  of the  $n - 1$  opponents have cooperated in the preceding round. We focus on the set,  $\Omega$ , of all reactive strategies in which  $d_\varphi$  is either 0 or 1 and  $d_j$  is either 0 or 1, namely,  $\Omega = \{(d_\varphi, d_0, d_1, \dots, d_{n-1}) | d_\varphi \in \{0, 1\}, d_j \in \{0, 1\}, j \in \{0, 1, \dots, n - 1\}\}$ . Obviously, there are  $2^{n+1}$  strategies in  $\Omega$ .

Table 1 shows the payoff matrix of the general two-strategy  $n$ -player game, where  $a_l$  denotes the payoff to an individual with strategy A whose opponents are  $n - l$  individuals with strategy A and  $l - 1$  individuals with strategy B; likewise,  $b_l$  denotes the payoff to an individual with strategy B confronting with the same composition of opponents. Let A and B in Table 1 represent a strategy  $\mathbf{d}$  in  $\Omega$  and ALLD, respectively. To describe the game involving  $\mathbf{d}$  and ALLD, we need to write down all  $a_l$  and  $b_l$  in Table 1 in terms of  $\delta$ ,  $V(C|k)$ ,  $V(D|k)$ , and  $\mathbf{d}$ . Consider a group consisting of  $k$  individuals adopting  $\mathbf{d}$  and  $n - k$  individuals adopting ALLD. We denote by  $h_k$  the expected number of rounds in which a  $\mathbf{d}$  individual cooperates before the group is dismissed. Since there are on average  $1/(1 - \delta)$  rounds, we have  $0 \leq h_k \leq 1/(1 - \delta)$ . Further, since we do not consider any errors in behavior,  $\mathbf{d}$  individuals perfectly coordinate their behavior, that is, either all or none of them cooperate in a given round. Hence, we obtain:

$$a_{n-k+1} = h_k V(C|k) + \left(\frac{1}{1 - \delta} - h_k\right) V(D|0), \tag{5}$$

$$b_{n-k} = h_k V(D|k) + \left(\frac{1}{1 - \delta} - h_k\right) V(D|0). \tag{6}$$

For any group composition,  $h_k$  is determined by  $d_\varphi$ ,  $d_{k-1}$ , and  $d_0$  in the following manner. First, suppose that  $d_\varphi = d_{k-1} = 1$ . In this case,  $\mathbf{d}$  never defects and, hence,  $h_k = 1/(1 - \delta)$ . Second, if  $d_\varphi = 1$  and  $d_{k-1} = d_0 = 0$ ,  $\mathbf{d}$  cooperates only in the first round and, thus,  $h_k = 1$ . Third, if  $d_\varphi = d_0 = 1$  and  $d_{k-1} = 0$ ,  $\mathbf{d}$  alternates cooperation and defection indefinitely, in which case  $h_k = 1/(1 - \delta^2)$ . Fourth, if  $d_\varphi = d_0 = 0$ ,  $\mathbf{d}$  defects in every round and hence  $h_k = 0$ . Fifth, if  $d_\varphi = 0$  and  $d_0 = d_{k-1} = 1$ ,  $\mathbf{d}$  defects only in the first round and thus  $h_k = \delta/(1 - \delta)$ . Sixth, if  $d_\varphi = d_{k-1} = 0$  and  $d_0 = 1$ ,  $\mathbf{d}$  alternates defection and cooperation indefinitely, in which case  $h_k = \delta/(1 - \delta^2)$ . In sum,  $h_k$  is given by:

$$h_k = d_\varphi \left[ 1 + \frac{\delta d_{k-1}}{1 - \delta} + \frac{\delta^2 d_0 (1 - d_{k-1})}{1 - \delta^2} \right] + (1 - d_\varphi) d_0 \left[ \frac{\delta}{1 - \delta^2} + d_{k-1} \frac{\delta^2}{1 - \delta^2} \right]. \tag{7}$$

Therefore, an  $n$ -player repeated PD involving  $\mathbf{d}$  and ALLD is specified using Equations (5)–(7).

**Table 1.** The payoff matrix of the general  $n$ -player game.

Strategy of the focal individual	Number of a Individuals among the $n - 1$ Opponents					
	$n - 1$	$n - 2$	$n - 3$	$\dots$	1	0
A	$a_1$	$a_2$	$a_3$	$\dots$	$a_{n-1}$	$a_n$
B	$b_1$	$b_2$	$b_3$	$\dots$	$b_{n-1}$	$b_n$

### 2.2. Evolutionary Dynamics

We use a model of stochastic evolutionary dynamics developed for the general two-strategy  $n$ -player game [18]. A population consisting of  $i$  individuals adopting strategy A and  $N - i$  individuals adopting strategy B is considered, where the population size,  $N$ , is constant ( $0 \leq i \leq N$ ). Groups of  $n$  individuals are formed by choosing individuals at random from the population and a game is played within each of these groups ( $2 \leq n \leq N$ ). The payoffs of the game are as shown in Table 1. Let  $F_i$  and  $G_i$  be the expected payoffs to an A individual and a B individual, respectively, when the number of A individuals in the population is  $i$ . The fitness of A and B individuals are given by  $f_i = 1 - w + wF_i$  and  $g_i = 1 - w + wG_i$ , respectively, where  $w$  is the selection intensity ( $0 < w < 1$ ). The population dynamics are formulated as a Moran process with frequency-dependent fitness: at each time step, an individual is chosen for reproduction with the probability proportional to its fitness, and then

one identical offspring is produced to replace another individual randomly chosen for death with probability  $1/N$  [17,51]. Denote by  $\rho_{A,B}$  the probability that a single individual with strategy A in a population of  $N - 1$  individuals with strategy B will finally take over the whole population (i.e., the fixation probability). Throughout the paper, we assume that the interaction group is smaller than the population ( $N > n$ ). Here, we have:

$$\rho_{A,B} = 1 / \left( 1 + \frac{\sum_{l=1}^{N-1} \prod_{i=1}^l \frac{1 - w + w \sum_{k=1}^n \frac{\binom{i}{n-k} \binom{N-i-1}{k-1}}{\binom{N-1}{n-1}} b_k}{1 - w + w \sum_{k=1}^n \frac{\binom{i-1}{n-k} \binom{N-i}{k-1}}{\binom{N-1}{n-1}} a_k} \right), \tag{8}$$

where  $\binom{s}{t}$  represents a binomial coefficient when  $s \geq t$  and is defined zero when  $s < t$ .

Assuming that selection is weak ( $w \ll 1$ ), [19,52,53] have shown that  $\rho_{A,B}$  is given approximately by:

$$\rho_{A,B} \approx \frac{1}{N} \frac{1}{1 - \alpha w / [n(n+1)]}, \tag{9}$$

where:

$$\alpha = (N - n) \sum_{k=1}^n k(a_k - b_k) + (n + 1) \sum_{k=1}^{n-1} k(a_{k+1} - b_k) \tag{10}$$

(see also [18,54,55]). Since  $\rho_{A,B}$  increases with  $\alpha$ , the strategy with the largest fixation probability maximizes the right-hand side of Equation (10).

### 3. Results

#### 3.1. General Case

We define the best reactive strategy as the strategy (or strategies) that maximizes the fixation probability when introduced as a single mutant in a population of ALLD. The best reactive strategy maximizes the right-hand side of Equation (10) in the context of the  $n$ -player repeated PD specified by Equations (5)–(7). Equation (10) is equivalent to:

$$\alpha = (N - n) \left( \sum_{k=1}^n a_{n-k+1} - n b_n \right) + (N + 1) \sum_{k=1}^n (n - k) (a_{n-k+1} - b_{n-k}). \tag{11}$$

Putting Equations (5) and (6) into Equation (11) and explicitly denoting the dependence on strategy  $\mathbf{d} \in \Omega$ , we obtain:

$$\alpha(\mathbf{d}) = \sum_{k=1}^n h_k(\mathbf{d}) f(k), \tag{12}$$

where:

$$f(k) = (N - n)(V(C|k) - V(D|0)) + (N + 1)(n - k)(V(C|k) - V(D|k)), \tag{13}$$

is independent of strategy  $\mathbf{d}$ . Note that  $h_k(\mathbf{d})$  is given by Equation (7) and is dependent on  $\mathbf{d}$  in the following manner:

$$h_k(\mathbf{d}) = h(d_\varphi, d_0, d_{k-1}, k). \tag{14}$$

The best reactive strategy  $\mathbf{d}^*$  is given by the following equation:

$$\mathbf{d}^* = \operatorname{argmax}_{\mathbf{d} \in \Omega} \left( \sum_{k=1}^n h_k(\mathbf{d}) f(k) \right). \tag{15}$$

As shown in Appendix A, the solution of the maximization problem in Equation (15) is given as follows. On the one hand, when:

$$\frac{1}{1-\delta} \sum_{k=1}^n \max\{f(k), 0\} + \sum_{k=1}^n \min\{f(k), 0\} > 0, \tag{16}$$

the solution of Equation (15) is given by:

$$\begin{cases} d_\varphi^* = 1, \\ d_{k-1}^* = 1 & \text{if } f(k) > 0, \\ d_{k-1}^* = 0 & \text{if } f(k) < 0, \end{cases} \tag{17}$$

for  $1 \leq k \leq n$ . Note that  $d_{k-1}^*$  can be either 0 or 1 in case  $f(k) = 0$ . On the other hand, when the inequality in Inequality (16) is reversed, the solution of Equation (15) is given by:

$$d_\varphi^* = d_0^* = 0, \tag{18}$$

where  $d_{k-1}^*$  can be either 0 or 1 for  $2 \leq k \leq n$ . In sum, Conditions (16)–(18) specify the best reactive strategy,  $\mathbf{d}^*$ .

Our main result Conditions (16)–(18) states that so long as the game is repeated sufficiently many times (i.e.,  $\delta$  is sufficiently large), each bit,  $d_{k-1}$ , of the best reactive strategy is solely determined by the sign of  $f(k)$ , which is determined by the payoff structure of the game (and the group size and the population size. In Equation (13), the term  $V(C|k) - V(D|0)$  is the payoff to a cooperating player in a group where  $k$  players cooperate minus the payoff to a defecting player in a group where no one cooperates. Thus, this term represents the benefit of cooperation. The term  $V(C|k) - V(D|k)$  represents the payoff difference between a cooperating player and a defecting player in a group where  $k$  individuals cooperate. Thus, this negative term represents the cost of cooperation. Weighted sum of them determines whether the best reactive strategy should cooperate if  $k - 1$  players have cooperated in the previous round.

In the following analysis, we consider cases in which Inequality (16) holds unless otherwise stated.

### 3.2. Linear Payoff

Let us begin with a special case where the following conventional linear payoff assumption holds:

$$V(D|k) = \frac{bk}{n}, \tag{19}$$

$$V(C|k) = V(D|k) - c. \tag{20}$$

From Inequalities (1)–(4), it has to be that  $0 < c < b < nc$ . Putting Equations (19) and (20) into Equation (13), we obtain:

$$f(k) = (N - n) \left( \frac{bk}{n} - c \right) - (N + 1)(n - k)c. \tag{21}$$

Since  $f(k)$  increases linearly with  $k$  in this case, there exists a critical value above which  $f(k)$  is positive (Figure 1a. Hence, from Equation (17),  $\alpha$  is maximized when the following condition is met:

$$\begin{cases} d_k = 1 \text{ for } k > k^*, \\ d_k = 0 \text{ for } k < k^*, \end{cases} \tag{22}$$

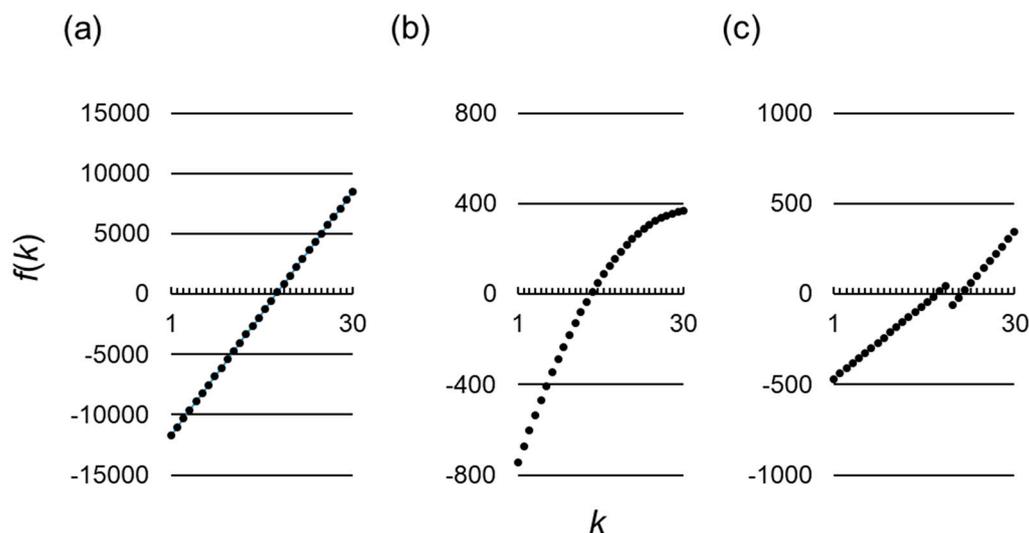
where:

$$k^* = \frac{(n + 1)N}{(N - n)b/(nc) + N + 1} - 1. \tag{23}$$

From Equation (23),  $k^*$  decreases with increasing  $N$ , and for an infinitely large population,  $k^*$  is given by:

$$k^* = \frac{n + 1}{b/(nc) + 1} - 1. \tag{24}$$

Since  $b/(nc) < 1$  by assumption,  $k^* > (n - 1)/2$  holds. Therefore, any strategy  $\mathbf{d}$  that cooperates when a half or less of the opponents have cooperated in the previous round cannot be the best reactive strategy.



**Figure 1.** Functional forms of  $f(k)$  for various payoff assumptions. The horizontal and vertical axes represent  $k$ , the number of cooperating individuals in a given round among the  $n$  group members, and  $f(k)$ , a function of  $k$  as defined by Equation (13), respectively. Parameter values used are  $N = 400$  and  $n = 30$ , (a) The conventional linear payoff assumption, Equations (19) and (20);  $b = 24$  and  $c = 1$ , (b) A non-linear payoff assumption, Equations (42) and (43); (c) A non-linear payoff assumption, Equations (20) and (52).  $b = 1$ .

A similar, but slightly different, result has been obtained by Kurokawa et al. (2010) [19]. They examined the fixation probability of a strategy called  $TFT_a$ , being introduced as a mutant into a population of ALLD, under the linear payoff assumption as given by Equations (19) and (20). An individual adopting  $TFT_a$  cooperates in the first round, and from the second round on, cooperates if and only if  $a$  or more of the opponents have cooperated in the previous round, where  $a \in \{0, 1, \dots, n - 1\}$  measures the level of generosity [23]. Thus, there are  $n$  strategies that are classified as  $TFT_a$  with different levels of generosity. From these, Kurokawa et al. (2010) specified  $a_{opt}$ , the optimal level of generosity, which maximizes the fixation probability. In the terminology of the present study, the  $n$  strategies constitute the subset of  $\Omega$  for which  $d_\varphi = 1$  and  $d_0 \leq d_1 \leq \dots \leq d_{n-1} = 1$ . In addition to these, we consider two kinds of non- $TFT_a$  reactive strategies. One is those that defect in the first round (i.e.,  $d_\varphi = 0$ ). Our analysis shows that some of this kind can maximize the fixation probability if the expected number of rounds is sufficiently small, partially modifying Kurokawa et al.'s (2010)

finding. Another kind is what we call “paradoxical” strategies (see Introduction), in which  $d_k > d_{k+1}$  holds for at least some  $k$ ; for example,  $\mathbf{d} = (1, 0, 1, 1, 0, \dots, 0, 1)$ . It turns out that Equations (22)–(24) is equivalent to the optimal level of generosity obtained by Kurokawa et al. (2010). Thus, their result is not changed by taking paradoxical strategies into consideration.

### 3.3. Non-Linear Payoff

The  $n$ -player PD can be further subcategorized according to characteristics of the fitness functions [39,46]. To illustrate, we define the cost and benefit associated with cooperation as:

$$c(k) = V(D|k) - V(C|k), \tag{25}$$

$$b(k) = V(D|k), \tag{26}$$

respectively. From Inequalities (1) and (2),  $c(k) > 0$  is required for all  $k$ . We can measure how the cost and benefit depend on the number of cooperators by considering:

$$\Delta c(k) = c(k + 1) - c(k), \tag{27}$$

$$\Delta b(k) = b(k + 1) - b(k). \tag{28}$$

For example, when  $\Delta c(k) = 0$  holds for all  $k$ , the cost is constant. Let us also consider:

$$\Delta^2 b(k) = \Delta b(k + 1) - \Delta b(k), \tag{29}$$

for  $0 \leq k \leq n - 2$ . It should be intuitively clear that the  $n$ -player PD requires  $\Delta b(k) > 0$ , which is equivalent to (3), while  $\Delta^2 b(k)$  can be either positive or negative (or zero). If, for instance,  $\Delta^2 b(k) \geq 0$  holds for all  $k$ , the benefit increases in an accelerative manner with the number of cooperators (e.g., the weakest-link game; [56]), whereas if  $\Delta^2 b(k) \leq 0$  always holds, the increase is decelerated (e.g., the volunteer’s dilemma; [57]). The class of  $n$ -player PDs for which  $\Delta^2 b(k) = 0$  always holds represent the public goods games in which the amount of public goods increases linearly with the number of cooperators [47].

The linear payoff assumptions in Equations (19) and (20) mean that the following conditions are satisfied for all  $k$ :

$$\Delta^2 b(k) = 0, \tag{30}$$

$$\Delta c(k) = 0. \tag{31}$$

Below we examine how evolution of generosity and paradoxical behavior may be affected by relaxing Equations (30) and (31).

#### 3.3.1. Evolution of Generosity with Non-Linear Payoff

##### Constant Cost with Non-Linear Benefit

Consider the case when there is no restriction on the benefit function (except Inequalities (1)–(4)), that is, Equation (30) is not necessarily true. For the moment, we assume Equation (31). From Equations (13) and (20), we have:

$$f(k) = (N - n)(V(D|k) - V(D|0)) - [(n + 1)N - k(N + 1)]c. \tag{32}$$

Using Conditions (1) and (20), we obtain:

$$V(D|k) - V(D|0) = \sum_{l=1}^k [V(D|l) - V(C|l) + V(C|l) - V(D|l - 1)] < kc, \tag{33}$$

which gives:

$$f(k) < [(2N - n + 1)k - (n + 1)N]c. \tag{34}$$

Thus, we obtain a necessary condition for  $f(k) > 0$ , namely:

$$k > \frac{(n + 1)N}{2N - n + 1}. \tag{35}$$

The right-hand side of Inequality (35) decreases and approaches  $(n + 1)/2$  with increasing  $N$ . Hence, from Equation (17),  $d_k = 1$  holds in the best reactive strategy only if  $k > (n - 1)/2$ . This means that any strategy  $\mathbf{d}$  who cooperates when a half or less of the opponents have cooperated in the previous round cannot be the best reactive strategy.

#### Variable Cost with Non-Decelerating Benefit

Let us now remove the restriction on the cost function, that is, we consider the case when Equation (31) does not necessarily hold. Instead of Equation (30), we assume:

$$\Delta^2 b(k) \geq 0, \tag{36}$$

which means that the benefit from cooperation increases with the number of cooperators in either linear or accelerative fashion. From Equations (13), (26) and (28), we have:

$$f(k) = (N - n) \left( \sum_{l=0}^{k-1} \Delta b(l) \right) - [(N + 1)(n - k) + (N - n)]c(k). \tag{37}$$

Using Inequality (36), we obtain:

$$\sum_{l=0}^{k-1} \Delta b(l) \leq k\Delta b(k - 1). \tag{38}$$

From Conditions (37) and (38), we have:

$$f(k) \leq (N - n)k\Delta b(k - 1) - [(N + 1)(n - k) + (N - n)]c(k). \tag{39}$$

Meanwhile, using Conditions (1), (25), (26) and (28), we obtain:

$$\Delta b(k - 1) = V(D|k) - V(D|k - 1) < V(D|k) - V(C|k) = c(k). \tag{40}$$

Therefore, combining Equation (39) and (40), we obtain:

$$f(k) < [(2N - n + 1)k - (n + 1)N]c(k). \tag{41}$$

Hence, as in the case of constant cost, Inequality (35) is necessary for  $f(k) > 0$ , which again means that the best reactive strategy must defect when a half or more of the opponents have defected in the previous round.

#### A Numerical Example

The results for the above two special cases are consistent with our finding in the analysis under the linear payoff assumption. However, the result may be altered qualitatively in the absence of any restriction on the benefit or cost functions. As an illustration, consider the following payoff functions:

$$V(D|k) = 1 - \left( \frac{n - k + 1}{n + 1} \right)^2, \tag{42}$$

$$V(C|k) = 1 - \left(\frac{n - k + 2}{n + 1}\right)^2 - \frac{1}{2(n + 1)^2}, \tag{43}$$

which satisfies Inequalities (1)–(4), but is not consistent with Equations (30) or (31). To be specific, let us set  $N = 400$ ,  $n = 30$ , and  $\delta = 0.8$ . Putting them into Equation (13), we obtain:

$$f(k) = -\frac{25400}{31} + \frac{146407k}{1922} - \frac{1172k^2}{961}. \tag{44}$$

In this case, Inequality (16) is met, and as shown in Figure 1b,  $f(k) > 0$  holds if and only if  $k \geq 14$ . Hence, from Equation (17), the best reactive strategy ( $\mathbf{d}^*$ ) satisfies the following:

$$\begin{cases} d_\varphi^* = 1, \\ d_k^* = 1 & \text{for } k \geq 13, \\ d_k^* = 0 & \text{for } k \leq 12. \end{cases} \tag{45}$$

Remember that  $d_k$  is defined as the probability with which he cooperates in a given round, provided that  $k$  of the  $n - 1$  “opponents” have cooperated in the preceding round. This demonstrates that there exist situations in which such a strategy that cooperates when a half or more of the opponents have defected in the previous round can be the best reactive strategy.

For comparison, the second and third best strategies following Equation (45) are given by:

$$\begin{cases} d_\varphi^* = 1, \\ d_k^* = 1 & \text{for } k \geq 14, \\ d_k^* = 0 & \text{for } k \leq 13, \end{cases} \tag{46}$$

$$\begin{cases} d_\varphi^* = 1, \\ d_k^* = 1 & \text{for } k \geq 12, \\ d_k^* = 0 & \text{for } k \leq 11, \end{cases} \tag{47}$$

respectively. For  $w = 0.001$ , the approximated fixation probabilities for Equations (45)–(47), calculated from Equation (9), are 0.00254215, 0.00254207, and 0.00254176, respectively.

### 3.3.2. Evolution of Paradoxical Behavior with Non-Linear Payoff

#### Non-Increasing Cost and Non-Linear Benefit

Consider the case when there is no restriction on the benefit function (except Inequalities (1)–(4)) and, thus, Equation (30) is not necessarily true. As for the cost function, we partially relax Equation (31) by assuming instead:

$$\Delta c(k) \leq 0, \tag{48}$$

which means that the cost of cooperation either is constant or decreases with increasing number of cooperators. Equation (13) changes into:

$$f(k) = (N - n)(V(C|k) - V(D|0)) - (N + 1)(n - k)c(k). \tag{49}$$

Hence, we have:

$$f(k + 1) = (N - n)(V(C|k + 1) - V(D|0)) - (N + 1)(n - k - 1)c(k + 1). \tag{50}$$

From Conditions (48)–(50), we have:

$$f(k + 1) > f(k). \tag{51}$$

This means that if  $f(k) > 0$  then  $f(m) > 0$  for any  $m > k$ . Hence, from Equation (17), a paradoxical strategy cannot be the best reactive strategy.

### A Numerical Example

Our analysis thus far has shown that a paradoxical strategy cannot be the best reactive strategy so long as the cost of cooperation is either unaltered or decreased by increasing number of cooperators. However, this may not be the case if the cost of cooperation sometimes increases with the number of cooperators. Let us provide a numerical example in which the best reactive strategy exhibits a paradoxical behavior. We set  $N = 400$ ,  $n = 30$ ,  $\delta = 0.8$ , and  $b = 1$  and assume that Equation (30) is met. Instead of Equation (31), here we assume:

$$c(k) = \begin{cases} 0.07 & \text{for } k \geq 20, \\ 0.04 & \text{for } k \leq 19, \end{cases} \tag{52}$$

for which Inequalities (1)–(4) are satisfied. This represents a case when the cost of cooperation is larger in the presence of more cooperators in the group. Intuitively, in such a situation, a paradoxical strategy that cooperates only when there are relatively few cooperators may be more advantageous than non-paradoxical strategies that, which cooperate whenever the number of cooperators exceeds a threshold. Indeed, since in this case Inequality (16) is met, and  $f(k)$  is non-monotonic in this case (Figure 1c), the best reactive strategy ( $\mathbf{d}^*$ ) is given from Equation (17) by:

$$\begin{cases} d_\phi^* = 1, \\ d_k^* = 1 & \text{for } k = 17, 18, k \geq 21, \\ d_k^* = 0 & \text{for } k \leq 16, k = 19, 20. \end{cases} \tag{53}$$

That is, the best reactive strategy cooperates when there are 17 or more opponents who have cooperated in the previous round, except when the number of cooperators is 19 or 20, in which case it defects. The relative advantages of cooperation and defection are determined by a balance between the following two factors: On the one hand, obviously, cooperation is more advantageous when the associated cost is smaller. On the other hand, given that the cost is constant, the relative advantage of cooperation increases with the number of cooperators in the group. In the current example, there is a rise of the cost at  $k = 20$  (see Equation (52)), which can be compensated for only when the number of cooperating opponents increases to 21 or more.

The second and third best strategies following Equation (53) are given respectively by:

$$\begin{cases} d_\phi^* = 1, \\ d_k^* = 1 & \text{for } k = 16, 17, 18, k \geq 21, \\ d_k^* = 0 & \text{for } k \leq 15, k = 19, 20, \end{cases} \tag{54}$$

$$\begin{cases} d_\phi^* = 1, \\ d_k^* = 1 & \text{for } k = 18, k \geq 21, \\ d_k^* = 0 & \text{for } k \leq 17, k = 19, 20. \end{cases} \tag{55}$$

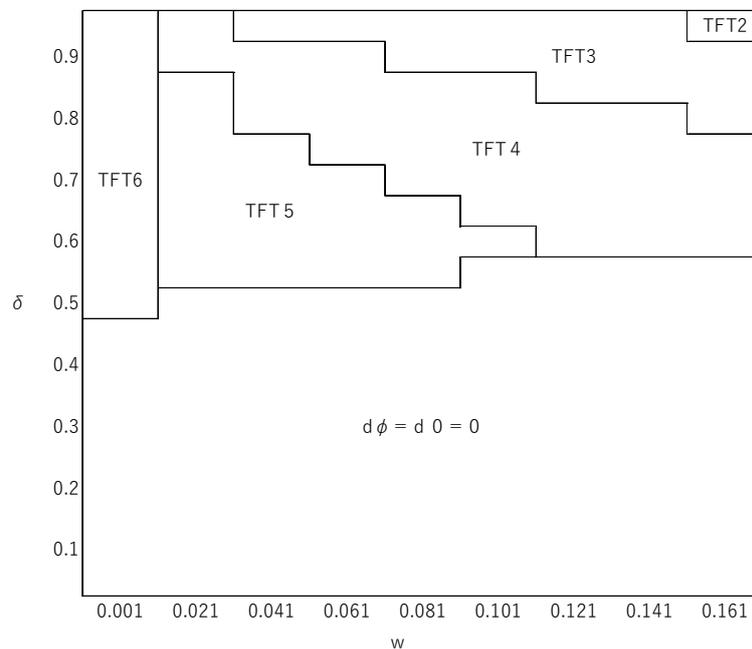
For  $w = 0.001$ , the fixation probabilities of strategies Equations (53)–(55) with the small  $w$  approximation are 0.00251170, 0.00251155, and 0.00251154, respectively.

### 3.4. The Best Reactive Strategy under Moderate Selection Intensity

Thus far we have assumed that selection is sufficiently weak ( $w \ll 1$ ) so that the fixation probabilities can be approximated by Equation (9). Here we numerically obtain the exact fixation

probabilities using (8) without the weak selection assumption to examine how this may affect the identity of the best reactive strategy.

For simplicity, we make the linear payoff assumption, Equations (19) and (20). Figure 2 illustrates how the best reactive strategy ( $\mathbf{d}^*$ ), obtained by using Equation (8), changes depending on the values of  $\delta$  and  $w$  for parameter values  $b = 9$ ,  $c = 1$ ,  $N = 50$ , and  $n = 10$ . For this parameter setting, a comparison of fixation probabilities calculated from Equation (9) would indicate TFT<sub>6</sub> as the best reactive strategy when  $\delta > \delta^*$  holds, where  $\delta^* \approx 0.49$ , and  $d_\phi = d_0 = 0$  when  $\delta < \delta^*$ . This gives a good approximation of the exact numerical solution for  $w = 0.001$  as shown in Figure 2. However, when  $w$  is larger, the approximation is no longer valid. Figure 2 shows that there exist situations in which such a strategy that cooperates when a half or more of the opponents have defected in the previous round can be the best reactive strategy, which cannot be the case in the limit of weak selection. A reactive strategy's fixation probability is determined by its payoffs relative to the unconditional defector at various frequencies ( $i = 1, 2, \dots, N - 1$  in Equation (8)). Among the  $N - 1$  values, a relative payoff at a lower frequency has a greater impact on the fixation probability. Hence, if, as numerically suggested by Kurokawa et al. (2010), the relative payoff at a low frequency tends to be larger for more generous reactive strategies, this is likely to favor the evolution of generosity. This effect might be more pronounced when selection, rather than random drift, plays a greater role. These considerations all in all accord with our numerical example suggesting that more generous reactive strategies are selectively favored when selection is more intense. It was also numerically shown that  $\delta$  can affect the identity of the best reactive strategy, while in the limit of weak selection  $\delta$  does not affect what the best reactive strategy is as far as it satisfies Inequality (16). In addition, within this particular range of parameter values, we did not find any case in which a paradoxical strategy is the best reactive strategy.



**Figure 2.** The effects of  $w$  and  $\delta$  on the identity of the best reactive strategy. The horizontal axis represents  $w$  (0.001 to 0.161 with the interval of 0.02) and the vertical axis is  $\delta$  (0.05 to 0.95 with the interval of 0.05). The fixation probabilities of strategies are obtained without assuming weak selection (i.e., using Equation (8)). The parameter values used are  $N = 50$ ,  $n = 10$ ,  $b = 9$ , and  $c = 1$ .

#### 4. Discussion

We have investigated stochastic evolutionary dynamics of a population in which an individual's fitness is determined by the  $n$ -player repeated PD. We have compared the fixation probabilities of different reactive strategies when they appear as a rare mutant in a population of unconditional

defectors. Reactive strategies in our analysis are described by a vector  $\mathbf{d} = (d_\varphi, d_0, d_1, \dots, d_{n-1})$ , where  $d_\varphi$  represents the probability with which the strategy cooperates in the first round and  $d_j$  represents the probability with which the strategy cooperates in a given round of a repeated game when  $j$  of the  $n - 1$  opponents have cooperated in the previous round. We have considered a set,  $\Omega$ , of all reactive strategies for which  $d_\varphi$  is either 0 or 1 and  $d_j$  is either 0 or 1 to specify the best reactive strategy attaining the maximum fixation probability. In a repeated PD, after one round is played, another round will be played with probability  $\delta$ . Under the assumption of weak selection, we have specified a threshold of  $\delta$  below which all strategies satisfying  $d_\varphi = d_0 = 0$  are the best reactive strategies. We have also found that when  $\delta$  exceeds the threshold,  $d_{k-1}$  of the best reactive strategy is solely determined by the sign of  $f(k)$ , as given by Equation (13), which is interpreted as a weighted sum of the cost and benefit of cooperation when  $k - 1$  opponents have cooperated in the proceeding round.

The present study extends our previous analysis (Kurokawa et al. (2010), [19]) by considering a broader set of reactive strategies and non-linear payoff functions. We also investigate how robust the indicated identity of the best reactive strategy is to deviations from the weak selection assumption. Kurokawa et al. (2010) specified the best reactive strategy from  $n$  reactive strategies classified as  $\text{TFT}_a$ , which constitute a subset of  $\Omega$ , under the conventional linear payoff assumption. One of their findings was that any  $\text{TFT}_a$  that tolerates defection by more than a half of the group members can never be the best reactive strategy. The present study has shown, on the one hand, that under the linear payoff assumption, the previous finding holds true even for all reactive strategies in  $\Omega$ . For non-linear payoff functions, on the other hand, we have shown that the previous finding does not always hold, that is, when the benefit increases in a decelerative manner with the number of cooperators, the best reactive strategy tolerates defection by more than a half of the opponents. We have also found that a strategy tolerating defection by more than a half of the opponents can be the best reactive strategy when selection is not weak.

We have also demonstrated that a paradoxical strategy, in which  $d_k > d_{k+1}$  holds for at least some  $k$ , can be the best reactive strategy when the cost of cooperation increases with the number of cooperators in a group. As far as we know, it has not been pointed out that a conditional cooperator can sometimes do better by behaving paradoxically in competition with defectors. A potentially relevant observation is that human participants in laboratory experiments of the repeated PD sometimes behave as if they were following a paradoxical strategy (e.g., [58–61]), though these studies assume linear payoff functions. The present study suggests that it may be of interest to investigate the effect of the shape of the payoff functions on the occurrence of paradoxical behaviors in the laboratory setting.

So far, we have examined the fixation probability of strategy  $\mathbf{d}$  when appeared as a single mutant in a population of ALLD. In Appendix B we also compare the fixation probability of ALLD when introduced as a single mutant in a population of reactive strategy  $\mathbf{d}$ , across all  $\mathbf{d}$  in  $\Omega$ . A reactive strategy associated with a lower fixation probability of ALLD is regarded as more robust against invasion by ALLD. The reactive strategy that minimizes the fixation probability of ALLD turns out to be  $\text{TFT}_{n-1}$ , that is,  $\mathbf{d} = (1, 0, 0, \dots, 0, 1)$ , when  $\delta$  is larger than the threshold specified by Inequality (16), and the strategies satisfying  $d_\varphi = d_0 = 0$  when otherwise. Further, we investigate the ratio of the fixation probability of  $\mathbf{d}$  in a population of ALLD to the fixation probability of ALLD in a population of  $\mathbf{d}$ , across all  $\mathbf{d}$  in  $\Omega$ . As shown in Appendix C, the reactive strategy maximizing the ratio of the fixation probabilities is  $\text{TFT}_{n-1}$  when  $\delta$  is larger than the threshold, and  $d_\varphi = d_0 = 0$  when otherwise. These results are consistent with our earlier analysis [19], even though the present study considers paradoxical strategies and non-linear payoff functions, which were not considered in the previous study.

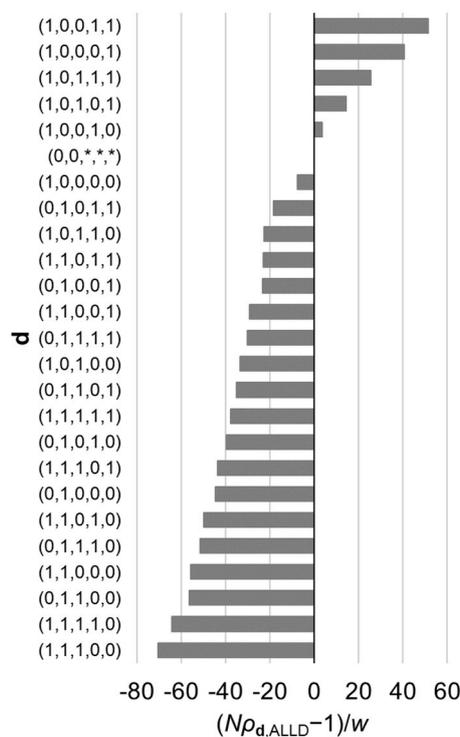
Natural selection is regarded as favoring a mutant reactive strategy replacing a population of ALLD if the fixation probability of the reactive strategy exceeds  $1/N$ , which is the fixation probability for neutral evolution. The reactive strategies that are the best when  $\delta$  is small so that Inequality (16) is not satisfied (i.e.,  $d_\varphi = d_0 = 0$ ) always have the fixation probability  $1/N$ . Thus, when Inequality (16)

is not met, the best reactive strategies and ALLD are selectively neutral. On the other hand, when  $\delta$  is large so that Inequality (16) is satisfied, there are other reactive strategies, including the best one, with the fixation probability larger than  $1/N$ . Hence, in this case, the best reactive strategy is always selectively favored over ALLD. See Appendix D for a further analysis on the case assuming linear payoff functions.

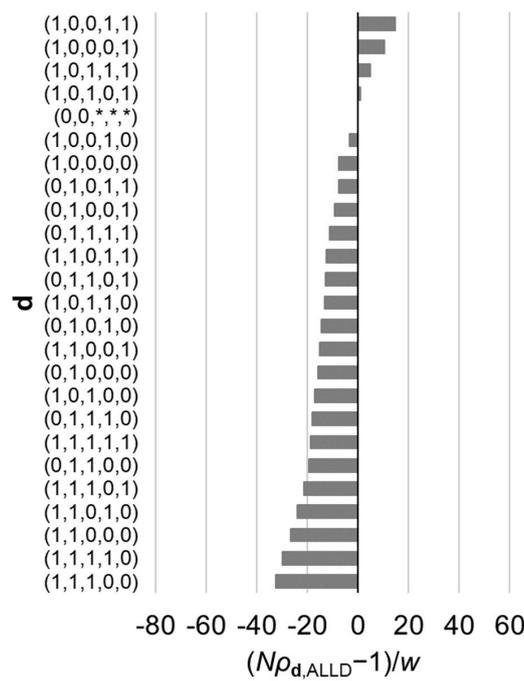
Thus far, our focus has been on the best reactive strategy, which maximizes the fixation probability. However, it is also of interest to examine characteristics of other reactive strategies whose fixation probabilities are relatively high and difference in fixation probabilities among those best strategies. For example, Table 2 gives a list of ten best strategies in the case of  $n = 10$  assuming Equations (19) and (20). Figure 3 illustrates the fixation probabilities of the 32 possible strategies in the case of  $n = 4$  for various values of  $\delta$ , under the assumption of Equations (19) and (20). The value of  $\delta$  makes difference in the order of the strategies. Note that as shown in Appendix A, the value of  $\delta$  does not have an effect on the order of the strategies belonging to subset of  $\Omega^{d_\varphi, d_0}$ .

**Table 2.** A list of the ten best reactive strategies for  $n = 10$  under the linear payoff assumption of Equations (19) and (20). Vector  $\mathbf{d}$  and the corresponding value of  $(N\rho_{\mathbf{d}, \text{ALLD}} - 1)/w$  are given. Parameter values are  $N = 100, b = 8, c = 1$ , and  $\delta = 0.8$ .

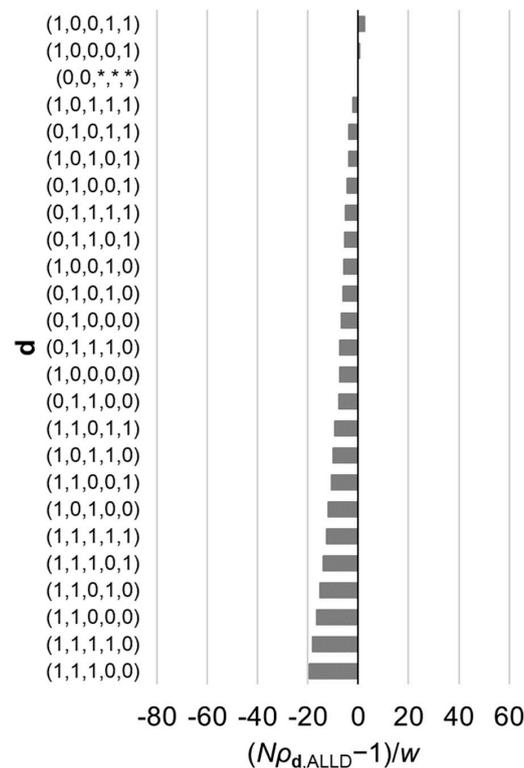
Ranking	$\mathbf{d}$	$(N\rho_{\mathbf{d}, \text{ALLD}} - 1)/w$
1	(1,0,0,0,0,0,1,1,1,1)	40.39
2	(1,0,0,0,0,0,1,1,1,1)	38.14
3	(1,0,0,0,0,0,0,1,1,1)	36.35
4	(1,0,0,0,0,0,1,0,1,1)	34.1
5	(1,0,0,0,0,1,0,1,1,1)	31.85
6	(1,0,0,0,0,0,1,0,1,1)	30.06
7	(1,0,0,0,0,1,1,1,1,1)	29.59
8	(1,0,0,0,0,1,0,0,1,1)	27.81
8	(1,0,0,0,0,1,1,0,1,1)	27.81
10	(1,0,0,0,0,0,0,0,1,1)	26.03



**Figure 3.** Cont.



(b)



(c)

**Figure 3.** Relative fixation probabilities of all reactive strategies for the linear public goods game with  $n = 4$ . The horizontal axis represents  $(N\rho_{d, ALLD} - 1)/w$ , which is positive when  $\rho_{d, ALLD} > 1/N$  (i.e., the fixation probability of a single mutant of strategy  $\mathbf{d}$ , when introduced in a population of ALLD, exceeds what is expected under neutral evolution), and is negative when the inequality is reversed. The strategies are aligned in the order of the fixation probabilities, which are calculated using Equation (9). Results for different value of  $\delta$  are shown: (a)  $\delta = 0.8$ , (b)  $\delta = 0.6$ , and (c)  $\delta = 0.4$ . Other parameter values are  $N = 100, b = 3.5$ , and  $c = 1$ . The asterisks indicate that the corresponding element of  $\mathbf{d}$  can be either 0 or 1. The order of the strategies is affected by  $\delta$ .

Finally, let us conclude by mentioning two issues to be addressed in future research. First, actual human behavior can be more complex than that in our model. Although the present study considers a greater variety of reactive strategies than in a previous study [19], the strategy set can be further broadened. Our strategy set,  $\Omega$ , does not include, for example, continuous strategies that engage in partial cooperation (see [62]), or those that make decisions referring to their own past behaviors ([63–68]. In addition, in our analysis, the benefit,  $b$ , and the cost,  $c$ , of cooperation are assumed to be the same for all individuals. Namely, we assumed that there are no differences among agents in this sense. In reality, however, individuals may have different values of  $b$  and  $c$ , and these variations could affect the model outcome [69]. Secondly, we have assumed that perfect information regarding group members' behaviors is always available. It has been shown, however, that results of related models may be qualitatively altered depending on the accessibility of the relevant information [45,64,70–75].

**Author Contributions:** Every author was involved in all stages of development and writing of the paper.

**Funding:** S.K. is partially funded by Chinese Academy of Sciences President's International Fellowship Initiative. Grant No. 2016PB018 and Monbukagakusho grant 16H06412. J.Y.W. is partly supported by JSPS Kakenhi 16K05283 and 16H06412. Y.I. is partly supported by the JSPS aid for Topic-Setting Program to Advance Cutting-Edge Humanities and Social Sciences Research and Kakenhi 16K07510 and 17H06381.

**Acknowledgments:** We thank Yoshio Kamijo (Kochi University of Technology) for discussions.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

### Solving (15)

To solve the maximization problem Equation (15), we first consider subsets  $\Omega^{u,v}$  of reactive strategies satisfying  $d_\varphi = u$  and  $d_0 = v$ , where:

$$\Omega^{1,0} \cup \Omega^{1,1} \cup \Omega^{0,0} \cup \Omega^{0,1} = \Omega, \quad (\text{A1})$$

we derive the solution,  $d_{k-1}^*$  ( $2 \leq k \leq n$ ), within each subset, after which the solution in  $\Omega$  is obtained by comparing the maxima of  $\alpha$  within the subsets.

First, when  $d_\varphi = 1$  and  $d_0 = 0$ , Equation (7) gives:

$$h_k(\mathbf{d}) = 1 + \frac{\delta d_{k-1}}{1 - \delta}. \quad (\text{A2})$$

Note that in this case  $h_k(\mathbf{d})$  depends only on a single component ( $d_{k-1}$ ) of  $\mathbf{d}$ . Since  $f(k)$  is constant and the right-hand side of Equation (A2) increases with  $d_{k-1}$ , the solution of the maximization problem within subset  $\Omega^{1,0}$  is given by:

$$\begin{cases} d_{k-1}^* = 1 & \text{if } f(k) > 0, \\ d_{k-1}^* = 0 & \text{if } f(k) < 0, \end{cases} \quad (\text{A3})$$

for  $2 \leq k \leq n$ . In case  $f(k) = 0$ , both  $d_{k-1} = 1$  and  $d_{k-1} = 0$  are the solutions. Hence, from Equations (12), (A2) and (A3), we obtain the maximum of  $\alpha$  within  $\Omega^{1,0}$  as follows:

$$\alpha(\mathbf{d}^*) = \frac{1}{1 - \delta} \sum_{k=1}^n \max\{f(k), 0\} + \sum_{k=1}^n \min\{f(k), 0\}. \quad (\text{A4})$$

Note that we use  $f(1) < 0$  (from Conditions (1), (2) and (13)) in the derivation of Equation (A4).

Second, when  $d_\varphi = d_0 = 1$ , we have:

$$h_k(\mathbf{d}) = \frac{1 + \delta d_{k-1}}{1 - \delta^2}, \quad (\text{A5})$$

which increases with  $d_{k-1}$ . Thus, the solution in this case is given as well by Equation (A3) for  $2 \leq k \leq n$ . Accordingly, the maximum of  $\alpha$  within  $\Omega^{1,1}$  is given by:

$$\alpha(\mathbf{d}^*) = \frac{1}{1-\delta} \sum_{k=1}^n \max\{f(k), 0\} + \frac{1}{1-\delta^2} \sum_{k=1}^n \min\{f(k), 0\} + \frac{\delta}{1-\delta^2} f(1). \tag{A6}$$

Third, when  $d_\varphi = d_0 = 0$ , we have:

$$h_k(\mathbf{d}) = 0. \tag{A7}$$

This means that all combinations of  $d_{k-1}$  for  $2 \leq k \leq n$  are the solutions and the maximum of  $\alpha$  within  $\Omega^{0,0}$  is:

$$\alpha(\mathbf{d}^*) = 0. \tag{A8}$$

Fourth, when  $d_\varphi = 0$  and  $d_0 = 1$ , we obtain:

$$h_k(\mathbf{d}) = \frac{\delta(1 + \delta d_{k-1})}{1 - \delta^2}, \tag{A9}$$

which increases with  $d_{k-1}$ . Hence, the solution in this case is also given by Equation (A3) for  $2 \leq k \leq n$ . Therefore, the maximum of  $\alpha$  within  $\Omega^{0,1}$  is given by:

$$\alpha(\mathbf{d}^*) = \frac{\delta}{1-\delta} \sum_{k=1}^n \max\{f(k), 0\} + \frac{\delta}{1-\delta^2} \sum_{k=1}^n \min\{f(k), 0\} + \frac{\delta^2}{1-\delta^2} f(1). \tag{A10}$$

Finally, a comparison of Equations (A4), (A6), (A8) and (A10) shows that the solution of Equation (15) depends on the expected number of rounds,  $1/(1-\delta)$ , as shown in Conditions (16)–(18). Note that as Equations (A2), (A5), (A7) and (A9) indicate,  $\delta$  does not affect the rank order of reactive strategies in the magnitudes of the fixation probabilities within each subset.

### Appendix B

#### The Fixation Probability of ALLD as a Single Mutant in a Population of $\mathbf{d}$

Under the assumption of weak selection, the fixation probability,  $\rho_{B,A}$ , of a single individual with strategy B in a population of  $N - 1$  individuals with strategy A is given approximately by:

$$\rho_{B,A} \approx \frac{1}{N} \frac{1}{1 - \varphi w / [n(n+1)]}, \tag{A11}$$

where:

$$\varphi = -(N - n) \left( na_1 - \sum_{k=1}^n b_k \right) - (N + 1) \sum_{k=1}^n (n - k)(a_{k+1} - b_k). \tag{A12}$$

Let us specify the reactive strategy  $\mathbf{d}$  that minimizes the fixation probability of ALLD. Putting Equations (5) and (6) into Equation (B2), we obtain:

$$\varphi(\mathbf{d}) = \sum_{k=1}^{n-1} h_k(\mathbf{d})g(k) - (N - n)nh_n(\mathbf{d})(V(C|n) - V(D|0)), \tag{A13}$$

where:

$$g(k) = [(N - n)(V(D|k) - V(D|0)) + (N + 1)k(V(D|k) - V(C|k))], \tag{A14}$$

is independent of strategy  $\mathbf{d}$ . Since  $V(D|k) \geq V(D|0)$  and  $V(D|k) > V(C|k)$  always hold, we have  $g(k) > 0$ .

Firstly, when  $d_\varphi = 1$  and  $d_0 = 0$ , Equations (7) and (A13) show:

$$\varphi = \sum_{k=1}^{n-1} \left[ 1 + \frac{\delta d_{k-1}}{1-\delta} \right] g(k) - (N-n)n \left[ 1 + \frac{\delta d_{n-1}}{1-\delta} \right] (V(C|n) - V(D|0)). \tag{A15}$$

From Equation (A15), since  $g(k) > 0$  and  $V(C|n) > V(D|0)$  always hold,  $\varphi$  is minimized when  $d_\varphi = 1, d_k = 0$  for  $0 \leq k \leq n-2$  and  $d_{n-1} = 1$ . This means that the fixation probability of ALLD is minimized in a population of strategy  $\mathbf{d} = (1, 0, 0, \dots, 0, 1)$ , namely,  $TFT_{n-1}$  in  $\Omega^{1,0}$ . In this case, the minimum of  $\varphi$  is given by:

$$\varphi^* = \sum_{k=1}^{n-1} g(k) - (N-n)n \frac{1}{1-\delta} (V(C|n) - V(D|0)). \tag{A16}$$

Secondly, when  $d_\varphi = d_0 = 1$ , (A13) becomes:

$$\varphi = \sum_{k=1}^{n-1} \left[ \frac{1}{1-\delta^2} + \frac{\delta d_{k-1}}{1-\delta^2} \right] g(k) - (N-n)n \left[ \frac{1}{1-\delta^2} + \frac{\delta d_{n-1}}{1-\delta^2} \right] (V(C|n) - V(D|0)). \tag{A17}$$

Since  $g(k) > 0$  and  $V(C|n) > V(D|0)$ ,  $\varphi$  is minimized when  $d_\varphi = 1, d_k = 0$  for  $1 \leq k \leq n-2$  and  $d_0 = d_{n-1} = 1$ . Hence, the minimum of  $\varphi$  in  $\Omega^{1,1}$  is given by:

$$\begin{aligned} \varphi^* = & \frac{1}{1-\delta} [(N-n)(V(D|1) - V(D|0)) + (N+1)(V(D|1) - V(C|1))] \\ & + \frac{1}{1-\delta^2} \sum_{k=2}^{n-1} g(k) - (N-n)n \frac{1}{1-\delta} (V(C|n) - V(D|0)). \end{aligned} \tag{A18}$$

Thirdly, when  $d_\varphi = d_0 = 0$ , Equations (7) and (A13) indicates that  $\varphi = 0$  always holds. Hence, irrespective of the values of  $d_k$  for  $1 \leq k \leq n-1$ , the minimum of  $\varphi$  in  $\Omega^{0,0}$  is:

$$\varphi^* = 0, \tag{A19}$$

and fourthly, when  $d_\varphi = 0$  and  $d_0 = 1$ , Equation (A13) becomes:

$$\begin{aligned} \varphi = & \sum_{k=1}^{n-1} \left[ \frac{\delta}{1-\delta^2} + d_{k-1} d \frac{\delta^2}{1-\delta^2} \right] g(k) \\ & - (N-n)n \left[ \frac{\delta}{1-\delta^2} + d_{n-1} \frac{\delta^2}{1-\delta^2} \right] (V(C|n) - V(D|0)). \end{aligned} \tag{A20}$$

Since  $g(k) > 0$  and  $V(C|n) > V(D|0)$ ,  $\varphi$  is minimized when  $d_\varphi = 0, d_k = 0$  for  $1 \leq k \leq n-2$  and  $d_0 = d_{n-1} = 1$ . Thus, the minimum of  $\varphi$  in  $\Omega^{0,1}$  is:

$$\begin{aligned} \varphi^* = & \frac{\delta}{1-\delta} [(N-n)(V(D|1) - V(D|0)) + (N+1)(V(D|1) - V(C|1))] \\ & + \frac{\delta}{1-\delta^2} \sum_{k=2}^{n-1} g(k) - (N-n)n \frac{\delta}{1-\delta} (V(C|n) - V(D|0)). \end{aligned} \tag{A21}$$

A comparison of Equations (A16), (A18), (A19) and (A21) shows the following. The strategy that minimizes  $\varphi$  among all strategies in  $\Omega$  is  $TFT_{n-1}$  when:

$$\delta > 1 - \frac{(N-n)n(V(C|n) - V(D|0))}{\sum_{k=1}^{n-1} g(k)}, \tag{A22}$$

and any strategy satisfying  $d_\varphi = d_0 = 0$  when the inequality in Equation (A22) is reversed. This means that the fixation probability of ALLD is minimized in a population of strategy  $\mathbf{d} = (1, 0, 0, \dots, 0, 1)$ ,

namely,  $TFT_{n-1}$  when  $\delta$  exceeds a threshold and otherwise, it is minimized in a population of strategies satisfying  $d_\varphi = d_0 = 0$ .

### Appendix C

#### The Ratio of the Fixation Probabilities

Ratio  $\rho_{A, B} / \rho_{B, A}$  can be viewed as a measure of the selective advantage of A over B for the long-term evolutionary process. From Equations (9) and (A11), the ratio is given approximately by:

$$\frac{\rho_{A, B}}{\rho_{B, A}} \approx 1 + \frac{\sigma\omega}{n}, \tag{A23}$$

where:

$$\sigma = (N - n)(a_1 - b_n) + N \sum_{k=1}^{n-1} (a_{k+1} - b_k). \tag{A24}$$

We regard A and B as strategy  $\mathbf{d}$  and ALLD, respectively, to specify the reactive strategy that maximizes the ratio of the fixation probabilities. Putting Equations (5) and (6) into Equation (A24) yields:

$$\sigma = -N \sum_{k=1}^{n-1} h_k(V(D|k) - V(C|k)) + (N - n)h_n(V(C|n) - V(D|0)). \tag{A25}$$

Firstly, when  $d_\varphi = 1$  and  $d_0 = 0$ , by using Equation (7), Equation (A25) becomes:

$$\begin{aligned} \sigma = -N \sum_{k=1}^{n-1} \left[ 1 + \frac{\delta d_{k-1}}{1 - \delta} \right] (V(D|k) - V(C|k)) \\ + (N - n) \left[ 1 + \frac{\delta d_{n-1}}{1 - \delta} \right] (V(C|n) - V(D|0)). \end{aligned} \tag{A26}$$

Since  $V(D|k) - V(C|k) > 0$  and  $V(C|n) > V(D|0)$  always hold,  $\sigma$  is maximized when  $d_\varphi = 1$ ,  $d_k = 0$  for  $0 \leq k \leq n - 2$  and  $d_{n-1} = 1$ . This means that the largest ratio of the fixation probabilities is attained by strategy  $\mathbf{d} = (1, 0, 0, \dots, 0, 1)$ , namely,  $TFT_{n-1}$ . Hence, the maximum of  $\sigma$  in  $\Omega^{1,0}$  is given by:

$$\sigma^* = -N \sum_{k=1}^{n-1} (V(D|k) - V(C|k)) + (N - n) \frac{1}{1 - \delta} (V(C|n) - V(D|0)). \tag{A27}$$

Secondly, when  $d_\varphi = d_0 = 1$ , Equation (A25) becomes:

$$\begin{aligned} \sigma = -N \sum_{k=1}^{n-1} \left[ \frac{1}{1 - \delta^2} + \frac{\delta}{1 - \delta^2} d_{k-1} \right] (V(D|k) - V(C|k)) \\ + (N - n) \left[ \frac{1}{1 - \delta^2} + \frac{\delta}{1 - \delta^2} d_{n-1} \right] (V(C|n) - V(D|0)). \end{aligned} \tag{A28}$$

Since  $V(D|k) - V(C|k) > 0$  and  $V(C|n) > V(D|0)$ ,  $\sigma$  is maximized when  $d_\varphi = 1$ ,  $d_k = 0$  for  $1 \leq k \leq n - 2$  and  $d_0 = d_{n-1} = 1$ . Thus, the maximum of  $\sigma$  in  $\Omega^{1,1}$  is:

$$\begin{aligned} \sigma^* = -N \left[ \frac{1}{1 - \delta} \right] (V(D|1) - V(C|1)) - N \sum_{k=2}^{n-1} \left[ \frac{1}{1 - \delta^2} \right] (V(D|k) - V(C|k)) \\ + (N - n) \left[ \frac{1}{1 - \delta} \right] (V(C|n) - V(D|0)). \end{aligned} \tag{A29}$$

Thirdly, when  $d_\varphi = d_0 = 0$  holds, Equation (7) and Equation (A25) indicate that  $\sigma = 0$  is always true, irrespective of the values of  $d_k$  for  $1 \leq k \leq n - 1$ . Therefore, the maximum of  $\sigma$  in  $\Omega^{0,0}$  is:

$$\sigma^* = 0. \tag{A30}$$

Fourthly, when  $d_\phi = 0$  and  $d_0 = 1$ , by using Equation (7), Equation (A25) becomes:

$$\begin{aligned} \sigma = & -N \sum_{k=1}^{n-1} \left[ \frac{\delta}{1-\delta^2} + d_{k-1} \frac{\delta^2}{1-\delta^2} \right] (V(D|k) - V(C|k)) \\ & + (N-n) \left[ \frac{\delta}{1-\delta^2} + d_{n-1} \frac{\delta^2}{1-\delta^2} \right] (V(C|n) - V(D|0)). \end{aligned} \tag{A31}$$

Since  $V(D|k) - V(C|k) > 0$  and  $V(C|n) > V(D|0)$  always hold,  $\sigma$  is maximized when  $d_\phi = 0, d_k = 0$  for  $1 \leq k \leq n-2$  and  $d_0 = d_{n-1} = 1$ . Hence, the maximum of  $\sigma$  in  $\Omega^{0,1}$  is given by:

$$\begin{aligned} \sigma^* = & -N \left[ \frac{\delta}{1-\delta} \right] (V(D|1) - V(C|1)) - N \sum_{k=2}^{n-1} \left[ \frac{\delta}{1-\delta^2} \right] (V(D|k) - V(C|k)) \\ & + (N-n) \left[ \frac{\delta}{1-\delta} \right] (V(C|n) - V(D|0)). \end{aligned} \tag{A32}$$

Finally, by comparing Equations (A27), (A29), (A30) and (A32), we obtain the following result. Of all strategies in  $\Omega$ , the strategy maximizing  $\sigma$  is  $TFT_{n-1}$  when:

$$\delta > 1 - \frac{(N-n)(V(C|n) - V(D|0))}{N \sum_{k=1}^{n-1} (V(D|k) - V(C|k))}, \tag{A33}$$

and any strategy satisfying  $d_\phi = d_0 = 0$  when the inequality in Inequality (A33) is reversed. This means that the largest ratio of the fixation probabilities is attained by strategy  $\mathbf{d} = (1, 0, 0, \dots, 0, 1)$ , namely,  $TFT_{n-1}$  when Inequality (A33) is met and, otherwise, the largest ratio of the fixation probabilities is attained by strategies satisfying  $d_\phi = d_0 = 0$ .

### Appendix D

#### The Conditions for Reactive Strategies to Be Selectively Favored over ALLD

We investigate whether natural selection favors strategy  $\mathbf{d}$  over ALLD under the assumption that the payoffs are given by linear functions in Equations (19) and (20). First, let us specify the condition for  $\mathbf{d}$  replacing a population of ALLD to be selectively favored, which is true when the fixation probability of  $\mathbf{d}$  exceeds  $1/N$ . Unless  $d_\phi = d_0 = 0$  holds true, this condition is given by:

$$\frac{b}{nc} > \frac{\frac{\sum_{k=0}^{n-1} d_k(n-k)}{\sum_{k=0}^{n-1} (n-k)} + d_0 - \left(1 - \frac{1}{\delta}\right)}{\frac{\sum_{k=0}^{n-1} d_k(k+1)}{\sum_{k=0}^{n-1} (k+1)} + d_0 - \left(1 - \frac{1}{\delta}\right)}. \tag{A34}$$

To make a geometric representation of the condition, let strategy  $\mathbf{d}$  be represented by a point,  $P_{\mathbf{d}}$ , on the  $xy$ -plane, whose coordinates are given by:

$$(x, y) = \left( \frac{\sum_{k=0}^{n-1} d_k(k+1)}{\sum_{k=0}^{n-1} (k+1)} + d_0, \frac{\sum_{k=0}^{n-1} d_k(n-k)}{\sum_{k=0}^{n-1} (n-k)} + d_0 \right). \tag{A35}$$

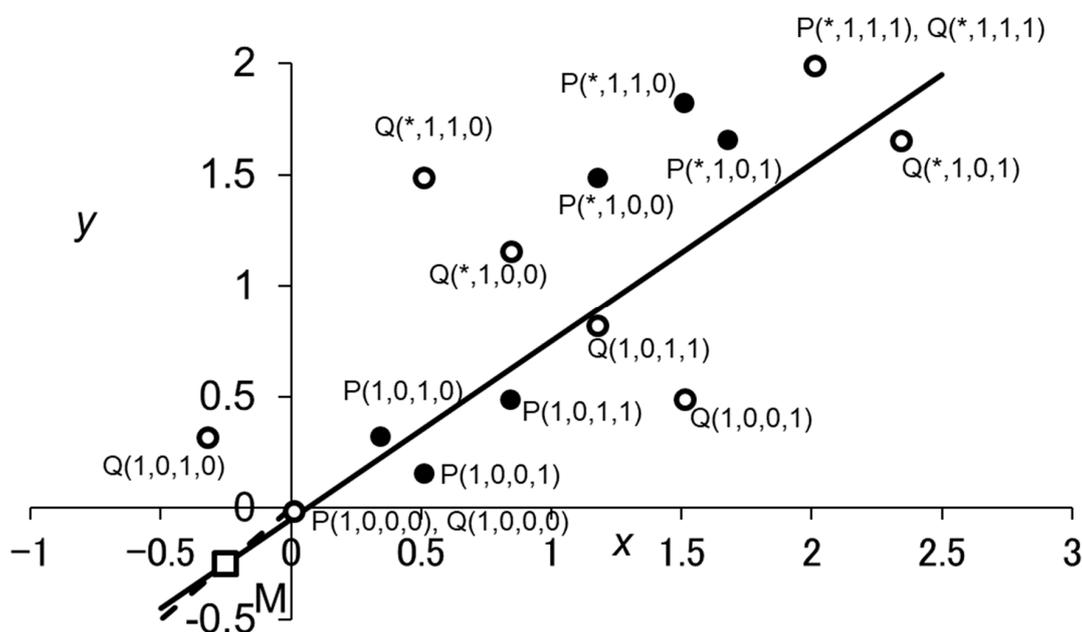
This formulation makes the right-hand side of Equation (A23) equivalent to the slope of the line through points  $P_{\mathbf{d}}$  to  $M$ , the latter of which are given by  $x = y = 1 - 1/\delta$ . Then, Inequality (A34) is satisfied if and only if  $P_{\mathbf{d}}$  is located below the “critical line,” which passes through  $M$  with slope  $b/(nc)$ , where  $1/n < b/(nc) < 1$ . Figure A1 provides such a geometric representation for the case of  $n = 3$ , showing  $\mathbf{d} = (1, 0, 0, 1)$  and  $(1, 0, 1, 1)$ , but none of the others are selectively favored when replacing a population of ALLD. If  $d_\phi = d_0 = 0$  holds true, the fixation probability of  $\mathbf{d}$  in a population of ALLDs is equal to  $1/N$ .

Second, we specify the condition for ALLD replacing a population of  $\mathbf{d}$  to be selectively disfavored, or the condition under which the fixation probability of ALLD is smaller than  $1/N$ . When either  $d_\varphi > 0$  or  $d_0 > 0$  is satisfied, the condition is given by:

$$\frac{b}{nc} > \frac{\frac{\sum_{k=0}^{n-1} d_k(k+1)}{\sum_{k=0}^{n-1} (k+1)} + d_0 - \left(1 - \frac{1}{\delta}\right)}{2d_{n-1} - \frac{\sum_{k=0}^{n-1} d_k(k+1)}{\sum_{k=0}^{n-1} (k+1)} + d_0 - \left(1 - \frac{1}{\delta}\right)}. \tag{A36}$$

Let strategy  $\mathbf{d}$  be represented by point  $Q_{\mathbf{d}}$ , whose coordinates are:

$$(x, y) = \left( 2d_{n-1} - \frac{\sum_{k=0}^{n-1} d_k(k+1)}{\sum_{k=0}^{n-1} (k+1)} + d_0, \frac{\sum_{k=0}^{n-1} d_k(k+1)}{\sum_{k=0}^{n-1} (k+1)} + d_0 \right). \tag{A37}$$



**Figure A1.** Geometric representations of Inequalities (A34) and (A36) for  $n = 3$ . Each reactive strategy  $\mathbf{d} = (d_\varphi, d_0, d_1, d_2)$  is represented by two points on the  $xy$ -plane,  $P(d_\varphi, d_0, d_1, d_2)$  and  $Q(d_\varphi, d_0, d_1, d_2)$ . The empty square represents point  $M$ , which is given by  $x = y = 1 - 1/\delta$  and on the line  $y = x$  ( $x < 0$ ) (the broken line). The thick line is the “critical line,” which passes  $M$  and has the slope  $b/(nc)$ . For given strategy  $\mathbf{d} = (d_\varphi, d_0, d_1, d_2)$ , Inequality (A34) is satisfied if and only if  $P(d_\varphi, d_0, d_1, d_2)$  is below the critical line, showing that  $\mathbf{d} = (1, 0, 0, 1)$  and  $(1, 0, 1, 1)$  are the only reactive strategies that are selectively favored when replacing a population of ALLD. Similarly, Inequality (A36) holds true if and only if  $Q(d_\varphi, d_0, d_1, d_2)$  is below the critical line. Parameter values used are  $\delta = 0.8$  and  $b/c = 2.4$ . The asterisks indicate that the corresponding element of  $\mathbf{d}$  can be either 0 or 1.

The slope of the line through points  $Q_{\mathbf{d}}$  and  $M$  gives the right-hand side of Inequality (A36). Thus, Inequality (A36) is met if and only if  $Q_{\mathbf{d}}$  is below the critical line as mentioned above. From Equation (A37), it is clear that  $x < y$  holds for any  $Q_{\mathbf{d}}$  with  $d_{n-1} = 0$ , which means that  $Q_{\mathbf{d}}$  with  $d_{n-1} = 0$  is above the critical line, for the slope of the critical line is always smaller than one. Hence, the fixation probability of ALLD replacing a population of  $\mathbf{d}$  is larger than  $1/N$  for any  $\mathbf{d}$  having  $d_{n-1} = 0$ . Figure A1 illustrates the case of  $n = 3$ , showing that ALLD is selectively disfavored when replacing a population of  $\mathbf{d} = (1, 0, 0, 1), (1, 0, 1, 1), (1, 1, 0, 1)$ , or  $(0, 1, 0, 1)$ . When  $d_\varphi = d_0 = 0$  holds true, the fixation probability of ALLD is equal to  $1/N$ .

## References

1. Trivers, R. The evolution of reciprocal altruism. *Q. Rev. Biol.* **1971**, *46*, 35–57. [[CrossRef](#)]
2. Axelrod, R.; Hamilton, W.D. The evolution of cooperation. *Science* **1981**, *211*, 1390–1396. [[CrossRef](#)]
3. Joshi, N.V. Evolution of cooperation by reciprocation within structured demes. *J. Genet.* **1987**, *6*, 69–84. [[CrossRef](#)]
4. Boyd, R.; Richerson, P.J. The evolution of reciprocity in sizable groups. *J. Theor. Biol.* **1988**, *132*, 337–356. [[CrossRef](#)]
5. Nowak, M.A. Stochastic strategies in the prisoner's dilemma. *Theor. Popul. Biol.* **1990**, *38*, 93–112. [[CrossRef](#)]
6. Nowak, M.A.; Sigmund, K. The evolution of stochastic strategies in the prisoner's dilemma. *Acta Appl. Math.* **1990**, *20*, 247–265. [[CrossRef](#)]
7. Press, W.H.; Dyson, F.J. Iterated prisoner's dilemma contains strategies that dominate any evolutionary opponent. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 10409–10413. [[CrossRef](#)]
8. Stewart, A.J.; Plotkin, J.B. Extortion and cooperation in the prisoner's dilemma. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 10134–10135. [[CrossRef](#)]
9. Stewart, A.J.; Plotkin, J.B. From extortion to generosity, evolution in the iterated prisoner's dilemma. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 15348–15353. [[CrossRef](#)]
10. Stewart, A.J.; Plotkin, J.B. Collapse of cooperation in evolving games. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 17558–17563. [[CrossRef](#)]
11. Hilbe, C.; Traulsen, A.; Sigmund, K. Partners or rivals? strategies for the iterated prisoner's dilemma. *Games Econ. Behav.* **2015**, *92*, 41–52. [[CrossRef](#)] [[PubMed](#)]
12. Hilbe, C.; Chatterjee, K.; Nowak, M.A. Partners and rivals in direct reciprocity. *Nature Hum. Behaviour.* **2018**, *2*, 469–477. [[CrossRef](#)]
13. Nowak, M.A.; Sigmund, K. Tit for tat in heterogeneous populations. *Nature* **1992**, *355*, 250–253. [[CrossRef](#)]
14. Nowak, M.A.; May, R.M. Evolutionary games and spatial chaos. *Nature* **1992**, *359*, 826–829. [[CrossRef](#)]
15. Killingback, T.; Doebeli, M. Self-organized criticality in spatial evolutionary game theory. *J. Theor. Biol.* **1998**, *191*, 335–340. [[CrossRef](#)] [[PubMed](#)]
16. Fudenberg, D.; Harris, C. Evolutionary dynamics with aggregate shocks. *J. Econ. Theory* **1992**, *57*, 420–441. [[CrossRef](#)]
17. Nowak, M.A.; Sasaki, A.; Taylor, C.; Fudenberg, D. Emergence of cooperation and evolutionary stability in finite populations. *Nature* **2004**, *428*, 646–650. [[CrossRef](#)]
18. Kurokawa, S.; Ihara, Y. Emergence of cooperation in public goods games. *Proc. R. Soc. Lond. B Biol. Sci.* **2009**, *276*, 1379–1384. [[CrossRef](#)]
19. Kurokawa, S.; Wakano, J.Y.; Ihara, Y. Generous cooperators can outperform non-generous cooperators when replacing a population of defectors. *Theor. Popul. Biol.* **2010**, *77*, 257–262. [[CrossRef](#)]
20. Kolllock, P. Social dilemmas: The anatomy of cooperation. *Annu. Rev. Sociol.* **1998**, *24*, 183–214. [[CrossRef](#)]
21. Milinski, M.; Semmann, D.; Krambeck, H.J.; Marotzke, J. Stabilizing the Earth's climate is not a losing game: Supporting evidence from public goods experiments. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 3994–3998. [[CrossRef](#)] [[PubMed](#)]
22. Hauert, C.; Schuster, H.G. Effects of increasing the number of players and memory size in the iterated Prisoner's Dilemma: A numerical approach. *Proc. R. Soc. Lond. B Biol. Sci.* **1997**, *264*, 513–519. [[CrossRef](#)]
23. Taylor, M. *Anarchy and Cooperation*; Wiley: New York, NY, USA, 1976.
24. Boyd, R.; Lorberbaum, J.P. No pure strategy is evolutionarily stable in the repeated prisoner's dilemma game. *Nature* **1987**, *327*, 58–59. [[CrossRef](#)]
25. Boyd, R. Mistakes allow evolutionary stability in the repeated prisoner's dilemma game. *J. Theor. Biol.* **1989**, *136*, 47–56. [[CrossRef](#)]
26. van Veelen, M. Robustness against indirect invasions. *Games Econ. Behav.* **2012**, *74*, 382–393. [[CrossRef](#)]
27. van Veelen, M.; García, J.; Rand, D.G.; Nowak, M.A. Direct reciprocity in structured populations. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 9929–9934. [[CrossRef](#)]
28. García, J.; van Veelen, M. In and out of equilibrium I: Evolution of strategies in repeated games with discounting. *J. Econ. Theor.* **2016**, *161*, 161–189. [[CrossRef](#)]
29. Bonner, J.T. *The Social Amoeba*; Princeton University Press: Princeton, NJ, USA, 2008.

30. Yip, E.C.; Powers, K.S.; Aviles, L. Cooperative capture of large prey solves scaling challenge faced by spider societies. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 11818–11822. [[CrossRef](#)]
31. Packer, C.; Scheel, D.; Pusey, A.E. Why lions form groups: Food is not enough. *Am. Nat.* **1990**, *136*, 1–19. [[CrossRef](#)]
32. Creel, S. Cooperative hunting and group size: Assumptions and currencies. *Anim. Behav.* **1997**, *54*, 1319–1324. [[CrossRef](#)]
33. Stander, P.E. Foraging dynamics of lions in semi-arid environment. *Can. J. Zool.* **1991**, *70*, 8–21. [[CrossRef](#)]
34. Bednarz, J.C. Cooperative hunting Harris' hawks *Parabuteo unicinctus*. *Science* **1988**, *239*, 1525–1527. [[CrossRef](#)]
35. Rabenold, K.N. Cooperative enhancement of reproductive success in tropical wren societies. *Ecology* **1984**, *65*, 871–885. [[CrossRef](#)]
36. Pacheco, J.M.; Santos, F.C.; Souza, M.O.; Skyrms, B. Evolutionary dynamics of collective action in N-person stag hunt dilemmas. *Proc. R. Soc. Lond. B Biol. Sci.* **2009**, *276*, 315–321. [[CrossRef](#)]
37. Bach, L.A.; Helvik, T.; Christiansen, F.B. The evolution of *n*-player cooperation—Threshold games and ESS bifurcations. *J. Theor. Biol.* **2006**, *238*, 426–434. [[CrossRef](#)] [[PubMed](#)]
38. Souza, M.O.; Pacheco, J.M.; Santos, F.C. Evolution of cooperation under *n*-person snowdrift games. *J. Theor. Biol.* **2009**, *260*, 581–588. [[CrossRef](#)] [[PubMed](#)]
39. De Jaegher, K. Harsh environments and the evolution of multi-player cooperation. *Theor. Popul. Biol.* **2017**, *113*, 1–12. [[CrossRef](#)]
40. Taylor, C.; Nowak, M.A. Transforming the dilemma. *Evolution* **2007**, *61*, 2281–2292. [[CrossRef](#)]
41. Nowak, M.A.; Tarnita, C.E.; Wilson, E.O. The evolution of eusociality. *Nature* **2010**, *466*, 1057–1062. [[CrossRef](#)]
42. Allen, B.; Nowak, M.A.; Wilson, E.O. Limitations of inclusive fitness. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 20135–20139. [[CrossRef](#)]
43. Allen, B.; Nowak, M.A. Games among relatives revisited. *J. Theor. Biol.* **2015**, *378*, 103–116. [[CrossRef](#)] [[PubMed](#)]
44. Kurokawa, S. Payoff non-linearity sways the effect of mistakes on the evolution of reciprocity. *Math. Biosci.* **2016**, *279*, 63–70. [[CrossRef](#)] [[PubMed](#)]
45. Kurokawa, S. Imperfect information facilitates the evolution of reciprocity. *Math. Biosci.* **2016**, *276*, 114–120. [[CrossRef](#)] [[PubMed](#)]
46. Peña, J.; Lehmann, L.; Nöldeke, G. Gains from switching and evolutionary stability in multi-player matrix games. *J. Theor. Biol.* **2014**, *346*, 23–33. [[CrossRef](#)] [[PubMed](#)]
47. Archetti, M.; Scheuring, I. Review: Game theory of public goods in one-shot social dilemmas without assortment. *J. Theor. Biol.* **2012**, *299*, 9–20. [[CrossRef](#)] [[PubMed](#)]
48. Bomze, I.; Pawlowitsch, C. One-third rules with equality: Second-order evolutionary stability conditions in finite populations. *J. Theor. Biol.* **2008**, *254*, 616–620. [[CrossRef](#)] [[PubMed](#)]
49. Wu, B.; García, J.; Hauert, C.; Traulsen, A. Extrapolating weak selection in evolutionary games. *PLoS Comput. Biol.* **2013**, *9*, e1003381. [[CrossRef](#)] [[PubMed](#)]
50. Slade, P.F. On risk-dominance and the '1/3—Rule' in 2X2 evolutionary games. *IJPAM* **2017**, *113*, 649–664. [[CrossRef](#)]
51. Moran, P.A.P. Random processes in genetics. *Math. Proc. Camb.* **1958**, *54*, 60–71. [[CrossRef](#)]
52. Deng, K.; Li, Z.; Kurokawa, S.; Chu, T. Rare but severe concerted punishment that favors cooperation. *Theor. Popul. Biol.* **2012**, *81*, 284–291. [[CrossRef](#)]
53. Zhang, C.; Chen, Z. The public goods game with a new form of shared reward. *J. Stat. Mech. Theor. Exp.* **2016**, *10*, 103201. [[CrossRef](#)]
54. Kurokawa, S.; Ihara, Y. Evolution of social behavior in finite populations: A payoff transformation in general *n*-player games and its implications. *Theor. Popul. Biol.* **2013**, *84*, 1–8. [[CrossRef](#)] [[PubMed](#)]
55. Gokhale, C.S.; Traulsen, A. Evolutionary games in the multiverse. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 5500–5504. [[CrossRef](#)] [[PubMed](#)]
56. Hirshleifer, J. From weakest-link to best-shot: The voluntary provision of public goods. *Public Choice* **1983**, *41*, 371–386. [[CrossRef](#)]
57. Diekmann, A. Volunteer's dilemma. *J. Confl. Resolut.* **1985**, *29*, 605–610. [[CrossRef](#)]
58. Fischbacher, U.; Gächter, S.; Fehr, E. Are people conditionally cooperative? Evidence from a public goods experiment. *Econ. Lett.* **2001**, *71*, 397–404. [[CrossRef](#)]

59. Kocher, M.G.; Cherry, T.; Kroll, S.; Netzer, R.J.; Sutter, M. Conditional cooperation on three continents. *Econ. Lett.* **2008**, *101*, 175–178. [[CrossRef](#)]
60. Herrmann, B.; Thöni, C. Measuring conditional cooperation: A replication study in Russia. *Exp. Econ.* **2009**, *12*, 87–92. [[CrossRef](#)]
61. Martinsson, P.; Pham-Khanh, N.; Villegas-Palacio, C. Conditional cooperation and disclosure in developing countries. *J. Econ. Psychol.* **2013**, *34*, 148–155. [[CrossRef](#)]
62. Takezawa, M.; Price, M.E. Revisiting “the evolution of reciprocity in sizable groups”: Continuous reciprocity in the repeated  $n$ -person prisoner’s dilemma. *J. Theor. Biol.* **2010**, *264*, 188–196. [[CrossRef](#)]
63. Nowak, M.A.; Sigmund, K. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner’s dilemma game. *Nature* **1993**, *364*, 56–58. [[CrossRef](#)] [[PubMed](#)]
64. Kurokawa, S. Persistence extends reciprocity. *Math. Biosci.* **2017**, *286*, 94–103. [[CrossRef](#)] [[PubMed](#)]
65. Hayden, B.Y.; Platt, M.L. Gambling for gatorade: Risk-sensitive decision making for fluid rewards in humans. *Anim. Cogn.* **2009**, *12*, 201–207. [[CrossRef](#)] [[PubMed](#)]
66. Scheibehenne, B.; Wilke, A.; Todd, P.M. Expectations of clumpy resources influence predictions of sequential events. *Evol. Hum. Behav.* **2011**, *32*, 326–333. [[CrossRef](#)]
67. Wang, Z.; Xu, B.; Zhou, H.-J. Social cycling and conditional responses in the rock-paper-scissors game. *Sci. Rep.* **2014**, *4*, 5830. [[CrossRef](#)] [[PubMed](#)]
68. Tamura, K.; Masuda, N. Win-stay lose-shift strategy in formation changes in football. *EPJ Data Sci.* **2015**, *4*, 9. [[CrossRef](#)]
69. Kurokawa, S. Evolution of cooperation: The analysis of the case wherein a different player has a different benefit and a different cost. *Lett. Evol. Behav. Sci.* **2016**, *7*, 5–8. [[CrossRef](#)]
70. Bowles, S.; Gintis, H. *A Cooperative Species: Human Reciprocity and its Evolution*; Princeton University Press: Princeton, NJ, USA, 2011. [[CrossRef](#)]
71. Kurokawa, S. Does imperfect information always disturb the evolution of reciprocity? *Lett. Evol. Behav. Sci.* **2016**, *7*, 14–16. [[CrossRef](#)]
72. Kurokawa, S. Evolutionary stagnation of reciprocators. *Anim. Behav.* **2016**, *122*, 217–225. [[CrossRef](#)]
73. Kurokawa, S. Unified and simple understanding for the evolution of conditional cooperators. *Math. Biosci.* **2016**, *282*, 16–20. [[CrossRef](#)]
74. Kurokawa, S. The extended reciprocity: Strong belief outperforms persistence. *J. Theor. Biol.* **2017**, *421*, 16–27. [[CrossRef](#)] [[PubMed](#)]
75. Kurokawa, S.; Ihara, Y. Evolution of group-wise cooperation: Is direct reciprocity insufficient? *J. Theor. Biol.* **2017**, *415*, 20–31. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).