Stabilizing unconditional cooperation

Saptarshi Pal[1,*], Mayeul Lambert[2], Martin A. Nowak[1,3]
[1]Department of Mathematics, Harvard University
[2] Polytechnic Institute of Paris
[3]Department of Organismic and Evolutionary Biology, Harvard University
[*] corresponding author

Abstract

The ultimate goal of research on evolution of cooperation could be conceived as finding a method for stabilizing strategies that always cooperate, that never deviate from cooperation, that never exploit, and never retaliate, because all such activities are inherently problematic. The hope of achieving this cooperative utopia seems unjustified, especially in the context of direct reciprocity, which relies on the principle that cooperative partner strategies succeed in defending themselves by using moderate retaliation. Here we propose a dynamic that goes a certain way toward achieving the desirable goal. In the ground state, the population consists of individuals that use always-cooperate (ALLC). Occasionally defectors, in the form of always-defect (ALLD), invade the population. Their presence triggers a mutation from ALLC to another strategy, X, with the aim to avert the take-over of defectors. In absence of X, ALLD dominates ALLC, but in the presence of X, the invasion attempt might fail and subsequently the ALLC ground state can be restored. We study this mutation-selection process in finite and infinite populations. We identify the properties of the ideal rescue strategy. We derive an optimum mutation rate which maximally stabilizes ALLC.

Introduction

Cooperation is a central theme in the theory of evolution[1–5]. Cooperators are individuals who pay costs to benefit others. Defectors, in contrast, pay no costs and distribute no benefits. Mechanisms for the evolution of cooperation are interactions structures that allow natural selection to favor cooperation over defection[2,3,6–12]. Here we focus on the mechanism of direct reciprocity[1,13–19]. There are repeated encounters between the same two individuals. If I cooperate now I might induce your cooperation later. If I defect now I might evoke your retaliation. Therefore in an iterated Prisoner's dilemma defection is no longer the only Nash equilibrium[20].

Much of the research within direct reciprocity deals with the question of how to play the repeated game. The search is for strategies that do well. Recently, strategies have been separated into two large categories: rivals and partners[18,21]. Rivals always want a higher payoff than their co-player. They insist on winning. Examples of rivals are always defect (ALLD) or so called extortionists[22]. Partners accept that in some situations their payoff will be lower than that of the co-player, but they incentivize their co-players to cooperate with them: when playing against a partner strategy you maximize your own payoff when you fully cooperate. Partners are focused on sharing. Examples of partners are generous tit-for-tat[23] and win-stay, lose-shift[24]. Broadly speaking, rivals destroy but partners facilitate evolution of cooperation.

When it comes to human behavior, an inherent problem of direct reciprocity is the need to retaliate against defection. A genuinely good person may be reluctant do to so. But if everyone uses unconditional cooperation then the mechanism of direct reciprocity fails to operate. In this paper we explore the question if it is possible to stabilize a population of always cooperate (ALLC) individuals. We explore the following setting. In the ground state, the population consists only of ALLC players. Occasionally, the ground state is challenged by invasion of defectors. For simplicity but without losing generality, we assume the challengers use the notorious ALLD strategy. When the invasion occurs a "danger signal" triggers the mutation from ALLC to a rescue strategy, X. The hope is that in the presence of X the invasion attempt will fail and the ground state will be restored.

In the following, we study both stochastic and deterministic evolutionary dynamics of the resulting mutation selection system. First, we study the stochastic evolutionary process by considering a finite population model. In this model, ALLC individuals probabilistically generate a mutant offspring X only when the number of defectors surpasses a certain threshold in the population. Our goal is to compute the likelihood that the ALLC population is restored once a single ALLD mutant attempts to invade it. We examine the effect of the conditional mutation probability, μ, on the restoration probability and characterize properties of the best rescue strategy, X. Later, we study deterministic evolutionary dynamics by considering infinite population size. Here we find an optimal mutation rate, μ_1, for which ALLC achieves maximal frequency in stable coexistence with X and ALLD.

The repeated donation game and strategies

Description of the game. The donation game involves two-players, each having two possible actions, cooperation, C, or defection, D. When an individual cooperates, they incur a cost c to provide a benefit b to the co-player. We have $b > c > 0$. When an individual defects, they provide no benefit and thereby incur no cost. The donation game is a special case of the Prisoner's Dilemma. For the four possible outcomes of the game, the payoffs for the row player are given by the payoff matrix

$$\begin{array}{cc} & \begin{array}{cc} \mathrm{C} & \mathrm{D} \end{array} \\ \begin{array}{c} \mathrm{C} \\ \mathrm{D} \end{array} & \begin{pmatrix} b-c & -c \\ b & 0 \end{pmatrix} \end{array} \tag{1}$$

In the repeated donation game, players continue with probability δ after each round. We focus on the limit $\delta \to 1$ which represents the infinitely repeated game. We assume that players make rare implementation errors: they sometimes play D when they intend to play C and *vice-versa*[15].

Strategies. A strategy is defined by the probability to cooperate in the first round and by the probabilities to cooperate after every sequence of past play. Here, we study strategies that base their decision on the outcome of the previous round, which are the so-called memory-1 strategies[15]. Since we consider the infinitely repeated game with implementation errors, we can ignore the initial move[15,25].

Therefore, a memory-1 strategy is defined by four parameters, $(p_{\mathrm{CC}}, p_{\mathrm{CD}}, p_{\mathrm{DC}}, p_{DD})$, which denote the probabilities to cooperate if the outcome of the previous round was $\mathrm{CC}, \mathrm{CD}, \mathrm{DC}$ or DD, respectively. In this notation, the first letter (C or D) denotes the move of the focal player, while the second letter (C or D) denotes the move of the other player. A strategy is called deterministic if all probabilities are binary variables taking the values 0 or 1. There are 16 deterministic memory-1 strategies. They include ALLD $(0, 0, 0, 0)$, GRIM $(1, 0, 0, 0)$, Win-stay, lose-shift $(1, 0, 0, 1)$, TFT $(1, 0, 1, 0)$, Firm-but-fair (FBF) $(1, 0, 1, 1)$, which is also known as Forgiver, and ALLC $(1, 1, 1, 1)$.

Computing expected payoffs. The first step in calculating expected payoff per-round for the two players is to determine how often the four states—CC, CD, DC, and DD—occur based on the strategies that the players adopt. Earlier work[25,26] details the method for computing the stationary distribution of the four states, provided both players employ memory-1 strategies. Due to rare implementation errors, there is a unique stationary distribution, $(v_{\mathrm{CC}}, v_{\mathrm{CD}}, v_{\mathrm{DC}}, v_{\mathrm{DD}})$[25,26]. The expected payoff of the focal player is then given by $\pi_1 = v_{\mathrm{CC}}(b-c) + v_{\mathrm{CD}}(-c) + v_{\mathrm{DC}}(b) + v_{\mathrm{DD}}(0)$. We use $\pi(s_i, s_j)$ to denote the payoff of the player who adopts the strategy s_i against her co-player who adopts s_j.

The payoffs when both players employ deterministic memory-1 strategies is given in Table 1. The effect of implementation error is noticeable when one observes the payoff that the strategy TFT (S_{10} in Table 1) obtains against itself. In a game with no errors, two TFT players, who start with cooperation, continue to cooperate in every round and receive an expected per-round payoff of $b-c$. However, with errors, one can derive that two TFT players visit each of the four states equally often: their stationary distribution is $(0.25, 0.25, 0.25, 0.25)$. As a result, they receive an expected payoff of $(b-c)/2$. Therefore, in the presence of noise, TFT fails to achieve the payoff for full cooperation. Consequently, generous tit-for-tat (GTFT) $(1, q, 1, q)$ for some $q > 0$ is a much better strategy, because it has the ability to forgive[15].

Evolutionary dynamics in a finite population

Consider a finite, well-mixed population of size N, whose individuals interact in a repeated donation game. We study evolutionary dynamics in a birth-death process[27] with mutation. In the beginning, the population is only composed of individuals who adopt ALLC. We call this the 'ground state' of the population. We challenge the ground state by switching the strategy of one individual to ALLD. We are interested in studying the evolutionary dynamics and the final state of the process that is initiated with this alteration. Below, we describe in detail the evolutionary process that unfolds.

Description of the birth-death process with conditional mutations. At each step of the birth-death process, first an individual is chosen randomly from the population, with probability proportional to its current fitness, to produce an offspring. In the second step, a second individual, selected uniformly at random to die. Thereby, the population size remains at a constant size. During the birth event, ALLD individuals reproduce without mutation; they always produce ALLD offspring. However, when an ALLC individual reproduces, the offspring is ALLC with probability $1 - \mu_e$, but adopts strategy X with probability μ_e. Individuals who adopt X reproduce without mutation. We allow X to be any memory-1 strategy. Later we also consider back mutation from X to ALLC.

At any point in time, the state of the population is specified by three numbers (N_C, N_D, N_X), denoting the abundance of ALLC players, ALLD players, and X players. The total population size, $N = N_C + N_D + N_X$, is constant.

The mutation probability from ALLC to X, μ_e depends on the number of ALLD players in the population, N_D. In particular, we consider

$$\mu_e = \begin{cases} \mu & \text{if } N_D \geq k \\ 0 & \text{if } N_D < k. \end{cases} \tag{2}$$

Thus, ALLC produces mutants only if N_D is greater or equal to k. We consider $k \geq 1$ in our analysis. ALLC does not mutate if $N_D = 0$. This threshold mutation rate represents a specific choice. Later we consider a mutation rate that is a linear function of N_D.

We illustrate the evolutionary process in Fig. 1. For $\mu = 0$, the stochastic process reduces to the fixation dynamics of an ALLD mutant in an ALLC population (see Fig. 1A). In this case, the population is always composed of at most two strategies, ALLC and ALLD, and there are two absorbing states, $(N, 0, 0)$ and $(0, N, 0)$. For $\mu > 0$, the population is composed of at most three strategies, ALLC, ALLD and X. Since mutations only occur when ALLD is present in the population, the population stabilizes once it returns to the ground state, $(N, 0, 0)$. Likewise, since reproductions by ALLD and X are never subject to mutations, the population also stabilizes when it reaches the homogenous states $(0, N, 0)$ or $(0, 0, N)$. Thus, there are three final states of the stochastic process that begins from $(N-1, 1, 0)$: the ground state, $(N, 0, 0)$, the state where everyone adopts ALLD, $(0, N, 0)$, and the state where everyone adopts X, $(0, 0, N)$ (see Fig. 1B).

Fitness of strategies in a population. In the birth-death process, individuals are selected for reproduction with probability proportional to their current fitness. We assume that fitness is a positive quantity that monotonically increases with average payoff. We compute the average payoff of a strategy i (denoted Π_i), by weighing its expected per-round payoffs against other strategies, $\pi(i, \cdot)$, with the probability of meeting them in the well-mixed population. For the population state $\mathbf{N} := (N_C, N_D, N_X)$, we obtain the payoff values

$$\begin{aligned}
\Pi_C(\mathbf{N}) &= (b-c)\left(\frac{\max\{N_C-1,0\}}{N-1}\right) - c\left(\frac{N_D}{N-1}\right) + \pi(C,X)\left(\frac{N_X}{N-1}\right) \\
\Pi_D(\mathbf{N}) &= b\left(\frac{N_C}{N-1}\right) + \pi(D,X)\left(\frac{N_X}{N-1}\right) \\
\Pi_X(\mathbf{N}) &= \pi(X,C)\left(\frac{N_C}{N-1}\right) + \pi(X,D)\left(\frac{N_D}{N-1}\right) + \pi(X,X)\left(\frac{\max\{N_X-1,0\}}{N-1}\right)
\end{aligned} \tag{3}$$

We assume that an individual's fitness is an exponential function of its average expected payoff[28]. That is, an individual with strategy i in the population state $\mathbf{N}$ has the fitness $\mathrm{F}_i(\mathbf{N}) = e^{\beta \Pi_i(\mathbf{N})}$. Here β is the intensity of selection. We consider $\beta = 1$ from here onwards.

Computing absorption probability into final states. When mutations are absent ($\mu = 0$), the evolutionary process is a discrete-time Markov chain in the state space

$$\mathrm{S}_{\mu=0} = \{(N_C, N-N_C, 0) \mid N_{\mathrm{C}} \in \{0, 1, .., N\}\}. \tag{4}$$

The probability that the process ends in state $(0, N, 0)$, which means all players have adopted

ALLD, after beginning from state $(N-1, 1, 0)$ is given by[27]

$$\rho_D = \frac{1}{1 + \sum_{k=1}^{N-1} \prod_{j=1}^{k} e^{-c-\frac{b}{N-1}}}. \tag{5}$$

With the complementary probability, $\rho_C := 1 - \rho_D$, the process ends in the other final state, $(N, 0, 0)$ wherein everyone adopts ALLC.

When mutations are possible, $\mu > 0$, the Markov process covers the bigger state space

$$\mathrm{S} := \left\{ (N_C, N_D, N_X) \in \{0, 1, 2, ..., N\}^3 \,\middle|\, N_C + N_D + N_X = N \right\}. \tag{6}$$

In addition to $(N, 0, 0)$ and $(0, N, 0)$, the state in which everyone adopts X, $(0, 0, N)$, is also an absorbing state. We define ρ_C, ρ_D and ρ_X as the probabilities that the process finishes in $(N, 0, 0)$,$(0, N, 0)$, and $(0, 0, N)$, respectively after it begins from state $(N-1, 1, 0)$.

To explicitly compute these absorption probabilities, we first derive the transition matrix $\mathbf{T}$ associated with the Markov process (see Eq. (13) in **Methods**). Then, we define sub-matrices $\mathbf{Q}$ and $\mathbf{R}$ of $\mathbf{T}$. The elements of sub-matrix $\mathbf{Q}$ only contain transition probabilities between non-absorbing states in S. It is a square matrix of size $|\mathrm{S}| - 3$. The sub-matrix $\mathbf{R}$ contains the transition probabilities from non-absorbing states to absorbing states. It is a matrix of size $(|\mathrm{S}| - 3) \times 3$. The probability that the process is absorbed in state s_{A}, when it begins from the non-absorbing state, s_{NA} is given by the element corresponding to the state-pair $(s_{\mathrm{NA}}, s_{\mathrm{A}})$ in the matrix

$$\mathbf{B} := (\mathbf{I} - \mathbf{Q})^{-1}\mathbf{R}. \tag{7}$$

Here $\mathbf{I}$ is the identity matrix. The absorption probabilities are thus,

$$\begin{aligned} \rho_C &= \mathrm{B}_{((N-1,1,0),(N,0,0))} \\ \rho_D &= \mathrm{B}_{((N-1,1,0),(0,N,0))}, \\ \rho_X &= \mathrm{B}_{((N-1,1,0),(0,0,N))}. \end{aligned} \tag{8}$$

Results

Rescue by deterministic memory-1 strategies

In order to study whether conditional mutations save a resident ALLC population against an invasion by ALLD, we begin by examining the case where X is a deterministic memory-1 strategy. We first study the dynamics in which mutations are triggered whenever ALLD is present ($k = 1$).

We find that among deterministic memory-1 strategies, TFT provides the best defense against an ALLD invasion (Fig. 2). While conditional mutations to strategies S_2, S_8, S_9 and S_{14} suppress an ALLD invasion, they do not restore the ground state as effectively as TFT because these rescue strategies themselves out-compete ALLC. Once ALLD is eradicated and mutation stops, the population, which is now composed of X and ALLC, may favor fixation of X over ALLC. For example, this is the case when the rescue strategy X is win-stay-lose-shift[24], S_9. In contrast, when X is TFT then the fixation of ALLC is favored, because it weakly dominates TFT in the infinitely repeated game, which we consider.

The only other deterministic strategy that displays comparable rescue property is S_{11}. This strategy behaves identically to TFT with the exception that it cooperates if both players defect in the last round - leading to its name *firm-but-fair*[25], FBF. This strategy is also known under the name Forgiver[29]. ALLC also weakly dominates FBF. But since FBF has a higher self-payoff than TFT (see Table 1), it has higher fitness than TFT in any mixed population with ALLC. As a result, FBF contributes to a weaker rescue effect than TFT.

For both TFT and FBF, we observe that rescue is most successful for intermediate mutation rates. Although a high mutation rate substantially reduces the chances that ALLD takes over, it risks producing a residual population in which the abundance of X is much greater than that of ALLC. In this case a return to a homogeneous ALLC ground state is unlikely.

This particular problem disappears if we consider a simple extension of our stochastic process: Once ALLD is eliminated, a back mutation from X to ALLC offspring with probability $\mu_{\text{back}} (> 0)$ is triggered. Now the process is guaranteed to end up in the ground state of only ALLC once ALLD becomes extinct. Interestingly, we observe that the strategy S_2 — to cooperate only after DC — provides marginally better rescue than TFT in this process as long as the mutation rate μ is not too high (see Supplementary Figure S3).

We note that the mechanism of conditional mutation backfires against ALLC when the rescue strategy is poorly chosen. For some X, the odds that ALLD takes over increases in comparison to the case where ALLC never mutates ($\mu = 0$). For the specific parameters we adopt in Figure 2, this occurs when X is S_5 (anti TFT), S_7 (cooperate always, except after mutual cooperation), S_{13} (cooperate always, except after CD), or S_0 (ALLD itself).

We also study the case where rescue response is delayed ($k = 25$). In this case, mutations only occur when the number of ALLD individuals exceed the threshold, $k = 25$. We summarize the result in Supplementary Figure S2. While rescue outcomes are always worse compared to $k = 1$, TFT still acts as the best rescue strategy, among all deterministic memory-1 strategies. Importantly, strategies that performed reasonably well in eradicating ALLD earlier, now do poorly. In fact for most X, outcomes are almost identical to the case where mutations remain absent ($\mu = 0$). For our parameters, TFT is the only strategy that brings the restoration probability close to one.

We also examine how deterministic memory-1 strategies perform as rescue strategies under an alternative mutational scheme, where the mutation probability is not a step function of the number of ALLD individuals, but increases linearly. Specifically, we consider a mutation rate defined by $\mu = N_D/(N-1)$. In this case, the probability of mutation from ALLC to X is zero when ALLD is absent, increases linearly with the number of ALLD in the population, and reaches one when $N-1$ out of N individuals are ALLD. The results are presented in Supplementary Figure S8. Again we find that TFT is the best rescue strategy among deterministic memory-1 strategies (see Figure S8A). We also compare the performance of the linear mutation function with the step function (Figure S8B). For this comparison, we used a step function of $\mu = 0.5$ at $k = 1$. For most rescue strategies, the linear mutation function performed better. But for TFT, there was no notable difference between the two, and for FBF, the step function performed considerably better.

Rescue by reactive strategies

Reactive strategies respond to the co-player's most recent move[15]. They are a two dimensional subspace of memory-1 strategies. A reactive strategy $S(p, q)$ cooperates with probability p after co-player cooperates, and cooperates with probability q after co-player defects. That is, for a reactive strategy $S(p, q)$ we have $p_{\mathrm{CC}} = p_{\mathrm{DC}} = p$ and $p_{\mathrm{CD}} = p_{\mathrm{DD}} = q$. For example, TFT is a reactive strategy with $(p, q) = (1, 0)$. We perform numerical computations to identify the best rescue strategy among the reactive strategies (see Fig. 3A). As before, we identify that rescue is most successful for intermediate mutation rates, μ. In addition, best rescue is performed by a reactive strategy that has $p \approx 1$, and $q \approx 0$. Those strategies are close to TFT. The values of p and q for the best rescue strategy depend on μ. In particular, for some values of μ (these are $\mu =$ 0.4, 0.5 and 1 in Fig. 3A), the best rescue strategy is exactly TFT.

Rescue by memory-1 strategies

We search for the best rescue strategy in the entire memory-1 space, which is given by the hypercube $[0,1]^4$. The details of our search process are described in **Methods**. The findings are summarized in Fig. 3B. We observe that the best rescue strategy has the form $(1, p_{\mathrm{CD}}, p_{\mathrm{DC}}, 0)$. Unlike the best rescue strategy in the reactive space, here the best rescue strategy always fully cooperates with ALLC ($p_{\mathrm{CC}} = 1$) and fully defects with ALLD ($p_{\mathrm{DD}} = 0$). The numerical values of p_{CD} and p_{DC} depend on μ and the remaining model parameters, which are $b = 5, c = 1$ and $N = 50$ (see Fig. 3B). We also plot the self-payoff of the best memory-1 strategy as function of the mutation rate, μ (see Fig. 3B). For low values of μ the self payoff decreases, but for larger values it increases again.

Search for the ideal rescue strategy

Let us now search for the ideal rescue strategy but independent of any constraints that are imposed by specific strategy spaces. The performance of a rescue strategy, X, depends on five numbers: (i) the payoff that X receives from ALLC, (ii) the payoff that X receives from ALLD, (iii) the payoff that ALLC receives from X, (iv) the payoff that ALLD receives from X, and (v) the payoff X receives from itself. It is evident from the search within the space of memory-1 strategies that the optimal rescue strategy, X, fully cooperates with ALLC while fully defecting against ALLD. Hence, the first four numbers must be: (i) $b - c$, (ii) 0, (iii) $b - c$, (iv) 0. The only quantity that needs to be optimized is the payoff X receives from itself which must lie in the interval $[0, b - c]$.

Our results are shown in Fig. 4B. If mutations are triggered by the presence of a single ALLD individual, $k = 1$, the optimal self-payoff of X decreases with mutation rate, μ, up to a certain point, after which it begins to increase. There are two opposing effects. A high self-payoff of X makes it more likely that ALLD becomes extinct. But a high self-payoff of X makes it harder for ALLC to reach fixation after ALLD has become extinct. For small μ, the first effect is more important. For large μ, the second effect is more important. The optimal-r curve aligns with the self-payoff of the best memory-1 rescue strategy of Fig. 3B, indicating agreement between the two searches.

When rescue is delayed, $k = 10$, the optimal self-payoff, r, is higher. In this case, opposing the spread of ALLD is of primary importance.

Evolutionary dynamics in the infinite population model

We now study the evolutionary dynamics of rescue in the limit of infinitely large population size. In the absence of mutation, the standard approach is given by the replicator equation[30–33]. In the presence of mutation, the standard approach is given by the replicator-mutator equation[31,34–37], which we use here. We analyze two versions of the replicator-mutator equation: one in which mutations occur during replication and the other in which mutations occur independently of the replication process[34,38–40]. In the main-text we focus on the latter version. In the Supplementary Information, we demonstrate that these two approaches produce qualitatively similar outcomes for our context.

In particular, we study the dynamics of a three-strategy system with ALLC, TFT and ALLD. Their relative frequencies are x, y and z, respectively. We have $x+y+z = 1$. Mutations from ALLC to TFT occur at a fixed rate μ. Evolutionary dynamics are given by

$$\begin{aligned}
\frac{\mathrm{d}x}{\mathrm{d}t} &= x(f_x - \bar{f}) - \mu x, \\
\frac{\mathrm{d}y}{\mathrm{d}t} &= y(f_y - \bar{f}) + \mu x, \\
\frac{\mathrm{d}z}{\mathrm{d}t} &= z(f_z - \bar{f}).
\end{aligned} \tag{D1}$$

Here f_x, f_y and f_z represent frequency-dependent fitness of the three types. We have

$$\begin{aligned} f_x &= b(x+y)-c \\ f_y &= (b-c)(x+y/2) \\ f_z &= bx \end{aligned} \tag{9}$$

The average fitness is $\bar{f} = xf_x+yf_y+zf_z = (b-c)(x+xy+y^2/2)$. The equation is defined on the simplex Δ_2, which is the set $\{(x,y,z) \in \mathbb{R}^3 \mid x+y+z=1, 0 \leq x,y,z \leq 1\}$. The simplex Δ_2 is invariant under the dynamics given by Eq. (D1). Any point in this set satisfies $\dot{x}+\dot{y}+\dot{z}=0$. Solutions that begin in this set, remain in this set forever. The corners of the simplex, $(1,0,0),(0,1,0)$ and $(0,0,1)$, represent the homogenous populations of ALLC, TFT and ALLD respectively. In Proposition 1 of Supplementary Information, we characterize all rest points of (D1) that lie in the simplex. In Proposition 2, we characterize the conditions under which each rest point is asymptotically stable.

As shown in Fig. 5A, the evolutionary dynamics depend on the value of the mutation rate, μ. For $\mu = 0$ the corners of the simplex are saddle points. In addition, there is an interior center which is surrounded by periodic orbits[41]. In the absence of TFT, ALLC is dominated by ALLD: therefore any mixed population of ALLC and ALLD converges to a homogeneous ALLD population.

We find two critical mutation rates, μ_1 and μ_2. They are given by

$$\mu_1 = b\left(1-\sqrt{1+\frac{c^2}{(b-c)^2}}\right)+\frac{c^2}{b-c} \tag{10}$$

and

$$\mu_2 = \frac{b-c}{2} \tag{11}$$

We show in Supplementary Information that $b > c > 0$ implies that $\mu_1 < \mu_2$.

If $0 < \mu < \mu_1$, there is an unsaturated fixed point on the edge between ALLC and TFT. This saddle point, which attracts all initial conditions where ALLD is absent, can be invaded by ALLD. In addition, there is an interior equilibrium which is asymptotically stable. All initial conditions with $x > 0$ and $z > 0$ converge to the interior equilibrium. All initial conditions with $x = 0$ and $z > 0$ converge to the TFT corner.

If $\mu_1 < \mu < \mu_2$, the fixed point between ALLC and TFT is saturated and asymptotically stable. All initial conditions with $x > 0$ converge to this boundary equilibrium. There is no interior equilibrium. All initial conditions with $x = 0$ and $z > 0$ converge to the TFT corner. For all initial conditions with $z < 1$, ALLD becomes extinct.

If $\mu_2 < \mu$ the corner point of TFT is asymptotically stable. All initial conditions with $x > 0$ or $y > 0$ converge to a population that is homogeneous in TFT. For all initial conditions with $z < 1$, ALLD becomes extinct.

In Fig. 5B, we show the equilibrium frequencies of the three strategies as function of the mutation rate μ. We observe that the equilibrium frequency of ALLC attains a maximum for $\mu = \mu_1$. The maximum equilibrium abundance of ALLC for $\mu > 0$ is given by

$$x_{max} = 1 - \frac{\sqrt{2(b^2+c^2) - 2b(c + \sqrt{(b-c)^2 + c^2})}}{b-c} \tag{12}$$

Therefore the deterministic system has an optimum mutation rate that maximizes the abundance of ALLC. In the stochastic system, we also found that intermediate mutation rates were optimal. We also note that $\mu \geq \mu_1$ guarantees the extinction of ALLD.

Discussion

In this paper, we have shown that a population of ALLC can be protected against invasion of defectors by mutating toward a rescue strategy. We imagine a situation where the presence of defectors triggers a warning signal that activates the mutation. We examine the resulting mutation-selection dynamics for both a stochastic system, which describes finite population size, and a deterministic system, which describes infinite population size. For both systems, we assume that individuals in the population are randomly paired to play infinitely repeated donation games. Our work adds to the larger body of literature, which studies the role of mutation-generated diversity in the context of evolution of cooperation[42–48].

We have focussed on the donation game as it offers the simplest framework for illustrating the role of rescuing strategies. The donation game is widely used for the study of direct reciprocity[16,25]. A donation games is a special case of the Prisoner's Dilemma[2], which is the most stringent game for studying evolution of cooperation[3,9]: while mutual cooperation yields higher payoffs for both players than mutual defection, cooperation is not the individually rational choice. Our framework is not limited to the donation game and can readily be applied to other scenarios. In Supplementary Figures S6 and S7, we present the performance of the 16 deterministic memory-1 strategies as rescuers of an ALLC population, when the underlying game is a Prisoner's Dilemma or a Stag Hunt. In the Prisoner's Dilemma (see Fig. S6), TFT remains the best rescue strategy among deterministic memory-1 strategies. In the Stag-Hunt game (Fig. S7), no deterministic memory-1 strategy reliably restores the ALLC ground state, as the rescue strategy often takes over the population. For example, when GRIM is used, ALLD is suppressed at high mutation rates, but the population is more likely to absorb into the GRIM state than the ALLC state. In such games, employing back-mutations, as previously discussed, may enable the recovery of the ALLC population.

For the stochastic system, we have searched for the optimal rescue strategy in the space of memory-1 strategies. Among deterministic and reactive memory-1 strategies, TFT and TFT-like strategies serve as the best rescuers (Fig. 2, 3A). In the full, stochastic memory-1 space, the best rescue strategy shares key properties with TFT: it fully cooperates with ALLC but fully defects with ALLD (see Fig. 3B). For a comparison between TFT and the optimal rescue strategy see Supplementary Figure S10. For such a strategy, its self payoff determines its efficiency to rescue the ALLC population (see Fig 4).

For the deterministic system, we study a replicator-mutator equation, considering the strategies ALLC, ALLD, and TFT. We examine how the rate of mutation from ALLC to TFT, μ, affects the three-strategy dynamic (see Fig. 5). Our analysis reveals a critical mutation rate μ_1 above which ALLD is guaranteed to perish from any mixed population. Furthermore, μ_1 is also optimal: at $\mu = \mu_1$, ALLC attains maximal abundance at an asymptotically stable equilibrium. While we have thoroughly explored TFT as the rescue strategy in the infinite population model, future work will focus on extending the analysis to a more general class of rescue strategies characterized by the following properties: the strategy fully cooperates with ALLC, fully defects against ALLD, and achieves an arbitrary payoff $r \in [0, b - c]$ when playing against itself in direct competition.

Most approaches to evolution of cooperation by direct or indirect reciprocity conclude that cooperation can only be sustained by populations of conditional cooperators (discriminators)[17,24,35,49–54]. Our paper suggests an alternative role for discriminators in evolution. In the context of direct reciprocity, we demonstrate their success as temporary helpers: they only emerge when unconditional cooperators (ALLC) need help to fight off an invasion of defectors. After the defectors are eliminated, the discriminators naturally become extinct thereby restoring the ALLC population. This mechanism of transient helpers is reminiscent of findings from other models of evolution of cooperation—those not based on direct reciprocity—where type X assists in eliminating type Z to protect type Y, but is ultimately eliminated by Y once its purpose is served[55,56].

For our mechanism to succeed, it is important that ALLC dominates the rescue strategy, as this increases the likelihood that, once defectors are eliminated, ALLC can outcompete the helpers and reclaim the population. In our case, TFT, which is a highly effective rescuer, is dominated by ALLC because of its inability to fully cooperate against itself in the presence of implementation errors. In this sense, errors are not an impediment, as previously argued[35,49,51,57], but rather a critical factor behind TFT's role in helping to re-establish stable cooperation. In future works, it will be valuable to explore how the rescue mechanism operates in the context of indirect reciprocity[53,58–63] and in spatial games or evolutionary graph theory[64–67].

Methods

Computation of transition matrix. For the finite population model, the transition matrix $\mathbf{T}$ collects probabilities of transition between states in S. The probability of transition from state $\mathbf{N}_a := (i, j, N-i-j)$ to $\mathbf{N}_b := (q, r, N-q-r)$ in a step is given by Eq. (13). In the following we denote with μ_e the value that the mutation function $\mu_e(N_D; k)$ attains at state $\mathbf{N}_a$ as per Eq. (2). We express the transition probability in terms of the fitnesses $\mathrm{F}_C, \mathrm{F}_D$ and F_X of the three strategies at population state $\mathbf{N}_a$. We define $\mathrm{F}_w := i\mathrm{F}_C + j\mathrm{F}_D + (N-i-j)\mathrm{F}_X$.

$$\mathrm{T}_{\mathbf{N}_a \to \mathbf{N}_b} = \begin{cases} (1-\mu_e) \cdot \dfrac{i\mathrm{F}_C}{\mathrm{F}_w} \cdot \dfrac{N-i-j}{N} & \text{if} \quad (q,r) = (i+1, j) \\[2ex] (1-\mu_e) \cdot \dfrac{i\mathrm{F}_C}{\mathrm{F}_w} \cdot \dfrac{j}{N} & \text{if} \quad (q,r) = (i+1, j-1) \\[2ex] \dfrac{j\mathrm{F}_D}{\mathrm{F}_w} \cdot \dfrac{N-i-j}{N} & \text{if} \quad (q,r) = (i, j+1) \\[2ex] \dfrac{(N-i-j)\mathrm{F}_X}{\mathrm{F}_w} \cdot \dfrac{j}{N} + \mu_e \cdot \dfrac{i\mathrm{F}_C}{\mathrm{F}_w} \cdot \dfrac{j}{N} & \text{if} \quad (q,r) = (i, j-1) \\[2ex] \dfrac{j\mathrm{F}_D}{\mathrm{F}_w} \cdot \dfrac{i}{N} & \text{if} \quad (q,r) = (i-1, j+1) \\[2ex] \mu_e \cdot \dfrac{i\mathrm{F}_C}{\mathrm{F}_w} \cdot \dfrac{i}{N} + \dfrac{(N-i-j)\mathrm{F}_X}{\mathrm{F}_w} \cdot \dfrac{i}{N} & \text{if} \quad (q,r) = (i-1, j) \\[2ex] 1 - \displaystyle\sum_{\mathbf{s} \neq \mathbf{N}_a} \mathrm{T}_{\mathbf{N}_a \to \mathbf{s}} & \text{if} \quad (q,r) = (i, j) \\[2ex] 0 & \text{otherwise} \end{cases} . \tag{13}$$

Payoffs when TFT faces ALLC, ALLD or itself in the limit of rare implementation error. Consider the case where strategies are subject to implementation errors. An intended cooperation is executed as defection with probability $\alpha\varepsilon$ and an intended defection is executed as cooperation with probability $\beta\varepsilon$. Let us assume that these two types of error are independent. In this setting, TFT or $(1,0,1,0)$ is effectively TFT$(\alpha,\beta,\varepsilon)$:= $(1-\alpha\varepsilon, \beta\varepsilon, 1-\alpha\varepsilon, \beta\varepsilon)$. Similarly, ALLD and ALLC are respectively ALLD$(\alpha,\beta,\varepsilon)$:= $(\beta\varepsilon, \beta\varepsilon, \beta\varepsilon, \beta\varepsilon)$, and ALLC$(\alpha,\beta,\varepsilon)$:= $(1-\alpha\varepsilon, 1-\alpha\varepsilon, 1-\alpha\varepsilon, 1-\alpha\varepsilon)$.

We consider $\alpha,\beta \geq 0$ and compute the expected long-run payoffs for strategy $s_1(\alpha,\beta,\varepsilon)$ against strategy $s_2(\alpha,\beta,\varepsilon)$ in the limit $\varepsilon \to 0^+$. Here s_1 and s_2 are TFT, ALLD or ALLC. Since these strategies are effectively purely stochastic, the transition matrix $\mathrm{W}(\alpha,\beta,\varepsilon)$, which collects the probability of transition between the states CC, CD, DC and DD between consecutive rounds is irreducible. The Markov chain therefore has a unique stationary distribution which is given by[68]

$$v(\alpha,\beta,\varepsilon) = (1,1,1,1)\cdot\left(\mathbb{I} + \mathrm{U} - \mathrm{W}(\alpha,\beta,\varepsilon)\right)^{-1}. \tag{14}$$

Where $\mathbb{I}$ is the identity matrix of size 4 and U is a 4×4 matrix with all entries equal to 1. The expected long-run payoff of s_1 and s_2 in the limit of rare implementation errors is given by

$$(\pi_1(\alpha,\beta), \pi_2(\alpha,\beta)) = \left(\lim_{\varepsilon\to 0^+} v(\alpha,\beta,\varepsilon)\right)\cdot\begin{pmatrix} b-c & -c & b & 0 \\ b-c & b & -c & 0 \end{pmatrix}^{\mathsf{T}} \tag{15}$$

We compute the payoffs between ALLC, ALLD and TFT using this method. They are

$$\begin{array}{cc} & \begin{array}{ccc} \text{ALLC} & \text{ALLD} & \text{TFT} \end{array} \\ \begin{array}{c} \text{ALLC} \\ \text{ALLD} \\ \text{TFT} \end{array} & \begin{pmatrix} b-c & -c & b-c \\ b & 0 & 0 \\ b-c & 0 & (b-c)\dfrac{\beta}{\alpha+\beta} \end{pmatrix} \end{array}. \tag{16}$$

For the case $\alpha = \beta$ (i.e., Table 1), TFT earns $(b-c)/2$ against itself. So long as $\alpha+\beta > 0$, TFT can, in principle, earn any self payoff in the interval $[0, b-c]$. In every calculation in this paper, we have assumed $\alpha = \beta$.

Search process for the best rescue strategy in reactive space: For Figure 3A we perform numerical computations to identify the best rescue strategy, (p, q), in the reactive strategy space, $[0,1]^2$. Specifically, we conduct a grid search over 500×500 uniformly spaced points in $[0,1]^2$ and report the strategy that maximizes the restoration probability of the ground state. For comparison, we perform an additional search restricted to rescue strategies with $p = 1$. We report the comparison in Supplementary Figure S9.

Search process for the best rescue strategy in memory-1 space: For Figure 3B we perform numerical computations to identify the best rescue strategy in the memory-1 space, $[0,1]^4$. For our computaitons, we consider the following parameters for our model N (population size) = 50, b, c (benefit and cost of cooperation) = 5, 1 and k (threshold of conditional mutations) = 1. Our search process was three step. In the first step, we performed a $20 \times 20 \times 20 \times 20$ uniformly spaced grid-search in the memory-1 space, and looked for the strategy that yielded highest restoration probability to ALLC. The best strategy from this first step had $p_{\mathrm{CC}} = 1, p_{\mathrm{DD}} = 0$. In the second step, we performed search in a subset of $[0,1]^4$. In this search p_{CC} was restricted to $[0.95, 1]$, p_{DD} to $[0, 0.05]$ while p_{CD} and p_{DC} to $[0,1]^2$. We use 20 uniformly spaced points for p_{CC} and p_{DD} while 50 for p_{CD} and p_{DC}. The best strategy from the second search also had $p_{\mathrm{CC}} = 1$ and $p_{\mathrm{DD}} = 0$. In the third and final search, we fix $p_{\mathrm{CC}} = 1$, $p_{\mathrm{DD}} = 0$ and perform a 500×500 uniformly spaced search in $[0,1]^2$ to find the optimal p_{CD} and p_{DC}.

Appendix: Analysis of the infinite population model

In the Propositions 1 and 2 below, we respectively characterize the fixed points of (D1) and the corresponding condition for the local stability of these fixed point. Proposition 3 describes the asymptotic extinction of ALLD from any mixed population.

Upon adding up the dynamical equations in Eq. (D1), we get $\dot{x} + \dot{y} + \dot{z} = 0$. The triangular simplex $\Delta_2 := \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 1, x, y, z \geq 0\}$ is thus invariant to (D1). If the dynamics begins in Δ_2 it remains there forever. We focus on (D1) in this set. Additionally, subsets $\{(x, y, z) \in \Delta_2 \mid z = 0\}$ and $\{(x, y, z) \in \Delta_2 \mid x = 0\}$ of Δ_2 are also invariant to Eq. (D1). These subsets form two edges of the triangular simplex in which ALLD and ALLC are respectively absent. Finally, the last edge, $\{(x, 0, z) \mid x + z = 1\}$ is only invariant if $\mu = 0$, otherwise not. The proposition below characterizes all rest points of (D1) in Δ_2.

Proposition 1 (Rest points of (D1) in Δ_2)**.** *For the replicator-mutator equation (D1), the following are the rest points in Δ_2 provided associated conditions (if any) are met.*

1. $\mathbf{x}_{\mathrm{TFT}} := (0, 1, 0)$,
2. $\mathbf{x}_{\mathrm{ALLD}} := (0, 0, 1)$,
3. $\mathbf{x}_{\mathrm{ALLC}} := (1, 0, 0)$ *if and only if* $\mu = 0$,
4. $\mathbf{x}_{\mathrm{E}} := \left(1-\sqrt{\frac{2\mu}{b-c}},\ \sqrt{\frac{2\mu}{b-c}},\ 0\right)$, *if and only if* $0 < \mu < \frac{b-c}{2}$,
5. $\mathbf{x}_{\mathrm{I}} := \left(\frac{(b-c)(c+\mu)^2}{2b(c^2-\mu(b-c))},\ \frac{c+\mu}{b},\ 1-\frac{c+\mu}{b}-\frac{(b-c)(c+\mu)^2}{2b(c^2-\mu(b-c))}\right)$ *iff* $0 \le \mu < \mu_1$.

where, μ_1 is given by Eq. (10).

We include proofs of Propositions in the section **Proofs** in Supplementary Information. The above proposition lists five possible rest points of (D1) in Δ_2. The first three, $\mathbf{x}_{\mathrm{TFT}}$, $\mathbf{x}_{\mathrm{ALLD}}$, and $\mathbf{x}_{\mathrm{ALLC}}$ are three corners of the simplex. In these points, the population only consists of TFT, ALLD or ALLC respectively. The fourth rest point, $\mathbf{x}_{\mathrm{E}}$, when it exists, is strictly in the interior of the edge where ALLD is absent (i.e., both ALLC and TFT are present). Finally, the last rest point, $\mathbf{x}_{\mathrm{I}}$ when it exists, is strictly in the interior of Δ_2 (i.e., all strategies coexist). We note the following degenerate cases: a) when $\mu = 0$, $\mathbf{x}_{\mathrm{E}} = \mathbf{x}_{\mathrm{ALLC}}$, b) when $\mu = (b-c)/2$, $\mathbf{x}_{\mathrm{E}} = \mathbf{x}_{\mathrm{TFT}}$. We also note that $\mu_1 < \mu_2 := (b-c)/2$ under the assumption that $b > c > 0$ (see end of Proof of Proposition 1). In the proposition below, we characterize the dynamic stability of these rest points.

Proposition 2 (Asymptotic stability of rest points of (D1))**.** *The following statements describe the asymptotic stability of rest points of (D1) with respect to perturbations in Δ_2. Each statement subsumes that the relevant condition for the fixed point's existence is met (from Proposition 1).*

1. *The rest points, $\mathbf{x}_{\mathrm{ALLD}}$ and $\mathbf{x}_{\mathrm{ALLC}}$, lying at corners of Δ_2, are not stable.*

2. *The third corner rest point,* $\mathbf{x}_{\mathrm{TFT}}$*, is stable if and only if* $\mu > (b-c)/2$.

3. *The rest point* $\mathbf{x}_{\mathrm{E}}$*, that lies on the boundary where ALLD is absent, is stable if and only if* $\mu > \mu_1$.

4. *The rest point,* $\mathbf{x}_{\mathrm{I}}$*, lying strictly in the interior of* Δ_2*, is stable if and only if* $\mu > 0$.

Furthermore, if the fixed point $\mathbf{x}_{\mathrm{E}}$ *exists, any solution starting from the set* $\{(x, 1-x, 0) \mid 0 < x \leq 1\}$*, asymptotically converges to* $\mathbf{x}_{\mathrm{E}}$ *(the set includes* $x = 1$ *if* $\mu > 0$*, otherwise not). Additionally, at* $\mu = 0$*, the interior fixed point* $\mathbf{x}_{\mathrm{I}}$ *acts as a center for periodic orbits in* Δ_2.

First, the Proposition claims that rest points corresponding to homogenous populations of ALLC and ALLD are not stable. We show that introducing TFT in small amounts to an ALLD population destabilizes it. Similarly, introducing ALLD destabilizes an ALLC population. When $\mu \leq (b-c)/2$, a homogenous TFT population is also not stable since a small introduction of ALLC grows. However, if mutations are too frequent, $\mu > (b-c)/2$, a homogenous TFT population is stable. Second, the Proposition claims that the value of μ determines the stability of the ALLC-TFT coexistence. This coexistence is stable when $\mu_1 < \mu < (b-c)/2$. When $\mu \leq \mu_1$, this coexistence is not stable; a small introduction of ALLD into this mixed population grows. Finally, the Proposition claims that the three types — ALLC, TFT and ALLD — can stably co-exist, if conditions allow for such a coexistence in the first place *and* the mutation rate is positive (that is, $0 < \mu < \mu_1$). At $\mu = 0$, this coexistence acts as a center for periodic orbits. We also show that any population containing ALLC but no ALLD, asymptotically reaches this ALLC-TFT coexistence. This also implies that this coexistence is always stable to perturbations that do not involve ALLD.

Corollary 1. *The strategy ALLC attains maximal frequency at a stable coexistence when* $\mu = \mu_1$.

Proposition 3 (Extinction of ALLD from any mixed population when μ is sufficiently high)**.** *If* $\mu \geq \mu_1$*, any solution* $(x(t), y(t), z(t))$ *of (D1) with* $(x(0), y(0), z(0)) \in \Delta_2$ *such that* $z(0) \in [0, 1)$ *has* $\lim_{t\to\infty} z(t) = 0$.

This Proposition claims that the dynamics (D1) eliminates ALLD from any starting population that has ALLD and other strategies (either TFT, ALLC or both), provided the mutation rate, μ is atleast μ_1.

Data availability: The necessary data and code for this manuscript has been archived. They can be found in the referenced link[69].

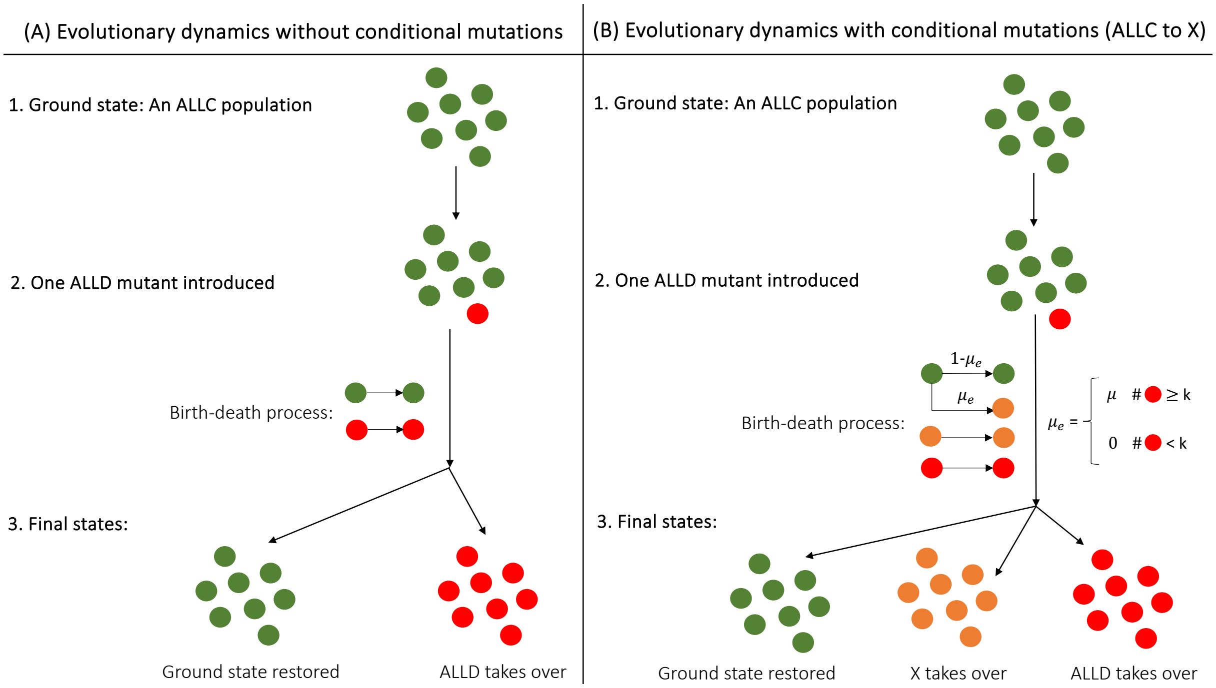


Figure 1: Two evolutionary dynamics involving strategies of the repeated donation game. For both evolutionary dynamics illustrated in this Figure, we consider the same starting point. The population which evolves is finite, well-mixed and initially composed of individuals who use the strategy to always cooperate (ALLC) in an infinitely repeated donation game. We call this population the ground state. In this population, a single individual switches to the strategy to always-defect (ALLD). This is the shared starting point of the two distinct dynamics we study from here onwards. (A) **The simple Birth-death process**, In this process, first, an individual is randomly selected from the population to give birth, with selection probability proportional to its fitness. Fitness of an individual is defined as the exponential of the expected per-round payoff of its strategy. Next, an individual is selected uniformly at random from the remaining population members to die. The process is repeated from first step. This dynamic results in two possible final states for the population; the initial ground state and the state in which all individuals adopt ALLD. (B) **Birth-death process with condtional mutations**, In an alternate dynamic we study the Birth-death process but with the exception that mutations may happen during the birth event. In particular, with probability μ_e, an ALLC individual gives birth to an offspring that adopts strategy X in the repeated game. With probability $1 - \mu_e$ it gives birth to an individual that adopts ALLC. The exact mutation probability depends on the number of ALLD individuals in the current population. In this model, mutations occur only if number of ALLD individuals exceed a certain threshold, k (≥ 1). Reproductions from ALLD and X are not subject to mutations. This dynamic has three final states; the ground state, the state in which everyone adopts X, and the state in which everyone adopts ALLD.

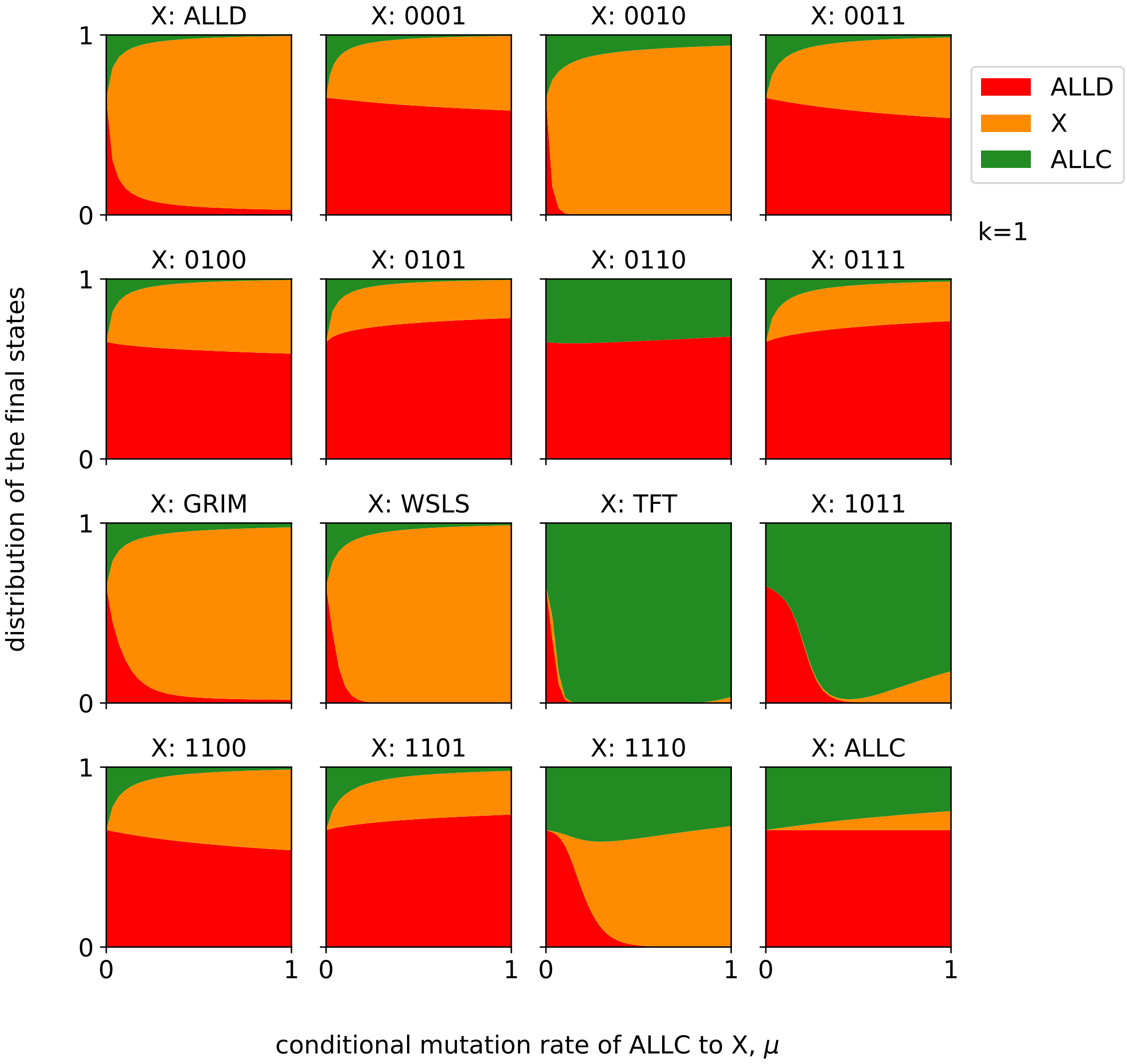


Figure 2: Among all deterministic memory-1 strategies, mutations to TFT result in maximal stabilization of ALLC. We consider $k = 1$ for the stochastic process described in Fig. 1(B). That is, mutations from ALLC to X begin when there is a single ALLD mutant in the population and stops only when the population reaches a state in which ALLD is absent. For each subplot, we consider a fixed memory-1 strategy as X, the strategy to which ALLC conditionally mutates with probability μ. We plot the distribution of the population's three possible final states—where everyone adopts ALLC (green), X (orange), or ALLD (red)—as we vary μ. We present this for all 16 possible deterministic memory-1 strategies as X. Here, a deterministic memory-1 strategy is represented by a four-digit string of 0s and 1s in the form: $p_{\mathrm{CC}}p_{\mathrm{CD}}p_{\mathrm{DC}}p_{\mathrm{DD}}$. These elements denote the probability to cooperate in a round when the outcome of the previous round is CC, CD, DC or DD respectively. The likelihood of restoring the ground state is the highest when ALLC conditionally mutates to TFT (row 3, column 3) at a rate which is optimally high. For this figure, we take a population with 100 individuals. The benefit, b and cost of cooperation, c in the donation game are 5 and 1 respectively.

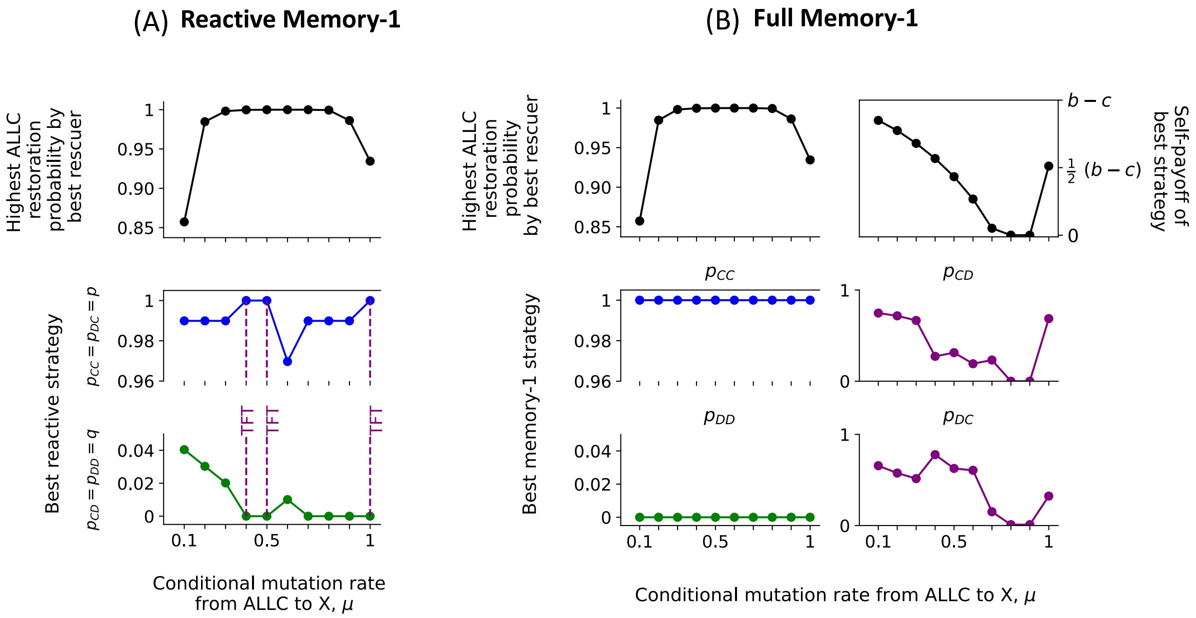


Figure 3: In the reactive memory-1 space, the best rescuers are geometrically close to TFT, whereas in the full memory-1 space, best rescuers have the discriminating property: they fully cooperate with always cooperate and fully defect with always defect. We perform numerical computations to identify the best rescue strategy in the space of reactive memory-1 strategies (panel A) and in the space of all memory-1 strategies (in panel B) (A) A reactive memory-1 strategy (p, q) responds solely to co-player's most recent action. It cooperates with probability p after co-player cooperates, or with probability q after co-player defects. That is, $p_{\mathrm{CC}} = p_{\mathrm{DC}} = p$ and $p_{\mathrm{CD}} = p_{\mathrm{DD}} = q$. We examine 500×500 reactive memory-1 strategies uniformly spaced in $[0, 1]^2$. For each X, we compute the restoration probability of the ALLC population. As we vary μ, we plot the maximum restoration probability obtained from 500×500 choices for X, along with the corresponding strategy X that yields this maximum. We find that ALLC restoration probability is highest when μ is neither too high nor too low. Furthermore, the best rescue strategy is geometrically close to $(p, q) = (1, 0)$, TFT. In fact, at $\mu = 0.4, 0.5$ and 1, the best strategy is exactly TFT. (B) We perform a search, similar to panel A, for finding the best rescue strategy in the entire memory-1 space (for more details on the search process, see **Methods**). The best rescue strategy fully cooperates with ALLC ($p_{\mathrm{CC}} = 1$) and fully defects with ALLD ($p_{\mathrm{DD}} = 0$). The strategy components p_{CD} and p_{DC} of the best rescue strategy determines its self-payoff, which is crucial in the rescue dynamic. We consider $k = 1$, i.e., rescue mutations occur whenever ALLD is present, and a population of size $N = 50$. All other parameters remain consistent with those in Figure 2.

(A)

$$\begin{array}{c} \\ \text{ALLC} \\ \text{ALLD} \\ \text{X} \end{array} \begin{array}{c} \begin{array}{ccc} \text{ALLC} & \text{ALLD} & \text{X} \end{array} \\ \begin{pmatrix} b-c & -c & b-c \\ b & 0 & 0 \\ b-c & 0 & r \end{pmatrix} \end{array}$$

(B)

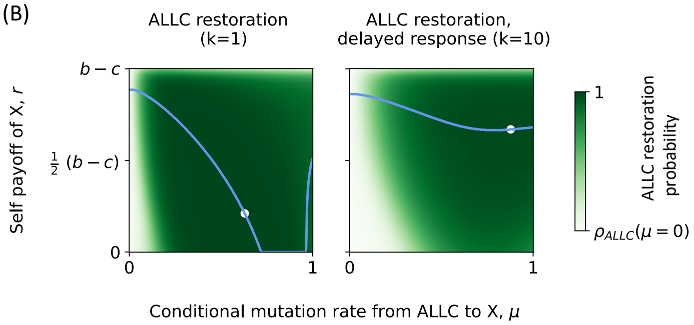


Figure 4: Role of self-payoff of discriminating strategies in the rescue of ALLC. (A) We consider a strategy X that fully cooperates with ALLC, fully defects with ALLD, and receives a payoff of r against itself in an infinitely repeated game, where $r \in [0, b-c]$. In **Methods**, we demonstrate that an example of X is TFT in an environment where two types of implementation errors—specifically, cooperating instead of defecting, and defecting instead of cooperating—are rare, independent, and have different probabilities. We are interested to study how r, the self payoff of X affects the rescue of ALLC. (B) To this end, we numerically compute the probability that the ALLC ground state is restored as we vary r and the conditional mutation probability from ALLC to X, μ. We study two cases, $k = 1$ and $k = 10$. In $k = 1$, mutations occur whenever ALLD is present. In $k = 10$, mutations occur only when the number of ALLD in the population is atleast 10. A blue line denotes the values of r that maximize the probability of ground state restoration for each value of μ. The white dot in each subplot denotes the pair (μ, r) that attains the highest restoration probability of ALLC. The lower end of the color scale denotes the probability of ALLC restoration when mutations do not occur. All other parameters remain consistent with those in Figure 3.

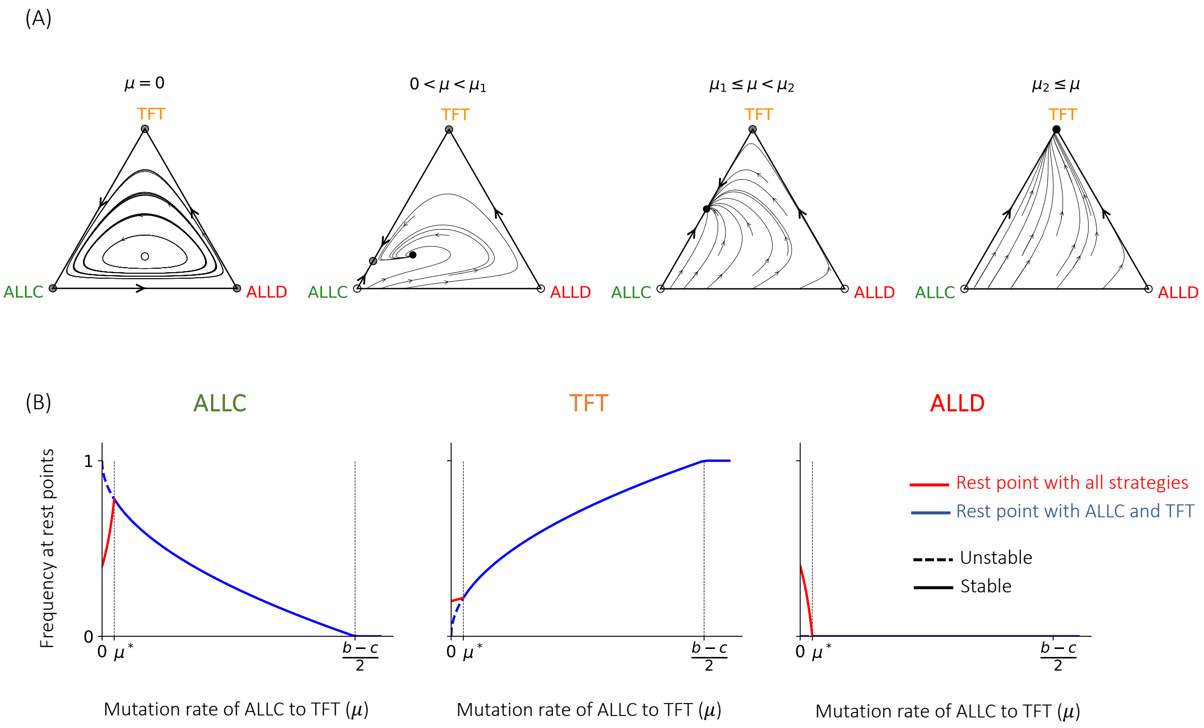


Figure 5: Analysis of the infinite population model with TFT as a rescuer shows an optimal μ ($= \mu_1$) at which ALLC is stable in maximal frequency. (A) The infinite population model demonstrates four unique phase portraits depending on the value of μ. At $\mu = 0$, the dynamic demonstrates a cyclic dominance between ALLC, TFT and ALLD. There is an interior fixed point that acts as center for periodic orbits. When $0 < \mu < \mu_1$, the dynamic produces a stable coexistence between all three strategies and an unstable coexistence between ALLC and TFT (see Eq. (10) for an expression of μ_1 in terms of b and c). Next, when $\mu \geq \mu_1$ but $\mu < \mu_2 =: (b - c)/2$, the dynamic has a single stable rest point, and in this rest point only TFT and ALLC are present. Finally for $\mu \geq \mu_2$, the only stable rest point is the homogenous TFT population. In the last two cases, the stable rest point is the final dynamical fate of all populations that are composed of ALLC and ALLD. (B) We show the frequency of ALLC, TFT and ALLD at rest points which are either in the interior of the simplex or at the ALLC-TFT edge of the simplex. We mark the frequencies with dashed lines if it corresponds to an unstable rest point and a solid line if it corresponds to a stable rest point. The maximum frequency of ALLC at a stable rest point is at $\mu = \mu_1$. (C) We show how the optimal mutation rate μ_1 varies with cost of cooperation c while we keep the benefit of cooperation b fixed. For this Figure, we take the benefit, b and cost of cooperation, c in the donation game as 5 and 1 respectively (for comparison, we make a similar figue in Supplementary Figure S4 for $b = 5, c = 3$).

	S_0	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}	S_{12}	S_{13}	S_{14}	S_{15}
S_0	$\mathbf{0}$	$\frac{b}{2}$	0	$\frac{b}{2}$	$\frac{b}{3}$	b	$\frac{b}{2}$	b	0	$\frac{b}{2}$	0	$\frac{b}{2}$	$\frac{b}{2}$	b	$\frac{2b}{3}$	b
S_1	$-\frac{c}{2}$	$\mathbf{\frac{b-c}{2}}$	$\frac{b-c}{3}$	$\frac{b-c}{2}$	$\frac{b-2c}{5}$	$\frac{2b-c}{3}$	b	$\frac{3b-c}{4}$	$-\frac{c}{2}$	$\frac{2b-c}{3}$	$\frac{b-c}{3}$	$\frac{2b-c}{3}$	$\frac{2b-c}{4}$	b	b	b
S_2	0	$\frac{b-c}{3}$	$\mathbf{\frac{b-c}{4}}$	$\frac{b-c}{2}$	0	$\frac{2b-c}{3}$	0	$\frac{2b-c}{3}$	0	$\frac{b-c}{3}$	$\frac{b-c}{3}$	$\frac{b-c}{2}$	$\frac{2b-c}{4}$	$\frac{2b-c}{2}$	$\frac{4b-2c}{5}$	$\frac{2b-c}{2}$
S_3	$-\frac{c}{2}$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$\mathbf{\frac{b-c}{2}}$	$-\frac{c}{2}$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$-\frac{c}{2}$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$\frac{2b-c}{2}$	$\frac{2b-c}{2}$	$\frac{2b-c}{2}$
S_4	$-\frac{c}{3}$	$\frac{2b-c}{5}$	0	$\frac{b}{2}$	$\mathbf{\frac{b-c}{4}}$	b	$\frac{b}{3}$	b	$-\frac{c}{3}$	$\frac{2b-c}{5}$	0	$\frac{b}{2}$	$\frac{3b-c}{6}$	b	$\frac{2b}{3}$	b
S_5	$-c$	$\frac{b-2c}{3}$	$\frac{b-2c}{3}$	$\frac{b-c}{2}$	$-c$	$\mathbf{\frac{b-c}{2}}$	$\frac{b-c}{2}$	$\frac{2b-c}{3}$	$-c$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$\frac{2b-c}{3}$	$\frac{b-c}{2}$	b	b	b
S_6	$-\frac{c}{2}$	$-c$	0	$\frac{b-c}{2}$	$-\frac{c}{3}$	$\frac{b-c}{2}$	$\mathbf{0}$	$\frac{2b-c}{3}$	$-\frac{2c}{3}$	$-c$	$\frac{b-c}{2}$	$\frac{2b-2c}{3}$	$\frac{b-c}{2}$	$\frac{4b-3c}{5}$	$\frac{4b-2c}{5}$	$\frac{2b-c}{2}$
S_7	$-c$	$\frac{b-3c}{4}$	$\frac{b-2c}{3}$	$\frac{b-c}{2}$	$-c$	$\frac{b-2c}{3}$	$\frac{b-2c}{3}$	$\mathbf{\frac{b-c}{2}}$	$-c$	$-c$	$\frac{2b-2c}{3}$	$\frac{2b-2c}{3}$	$\frac{2b-3c}{4}$	$\frac{4b-3c}{5}$	$\frac{2b-c}{2}$	$\frac{2b-c}{2}$
S_8	0	$\frac{b}{2}$	0	$\frac{b}{2}$	$\frac{b}{3}$	b	$\frac{2b}{3}$	b	$\mathbf{0}$	$\frac{3b-c}{5}$	0	$\frac{3b-c}{5}$	$\frac{3b-c}{6}$	$\frac{3b-c}{3}$	$\frac{3b-c}{4}$	$\frac{3b-c}{3}$
S_9	$-\frac{c}{2}$	$\frac{b-2c}{3}$	$\frac{b-c}{3}$	$\frac{b-c}{2}$	$\frac{b-2c}{5}$	$\frac{b-c}{2}$	b	b	$\frac{b-3c}{5}$	$\mathbf{b-c}$	$\frac{b-c}{2}$	$b-c$	$\frac{b-c}{2}$	$\frac{3b-2c}{3}$	$\frac{3b-c}{3}$	$\frac{2b-c}{2}$
S_{10}	0	$\frac{b-c}{3}$	$\frac{b-c}{3}$	$\frac{b-c}{2}$	0	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$\frac{2b-2c}{3}$	0	$\frac{b-c}{2}$	$\mathbf{\frac{b-c}{2}}$	$\frac{2b-2c}{3}$	$\frac{b-c}{2}$	$b-c$	$b-c$	$b-c$
S_{11}	$-\frac{c}{2}$	$\frac{b-2c}{3}$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$-\frac{c}{2}$	$\frac{b-2c}{3}$	$\frac{2b-2c}{3}$	$\frac{2b-2c}{3}$	$\frac{b-3c}{5}$	$b-c$	$\frac{2b-2c}{3}$	$\mathbf{\frac{3b-3c}{4}}$	$\frac{2b-3c}{4}$	$b-c$	$b-c$	$b-c$
S_{12}	$-\frac{c}{2}$	$\frac{b-2c}{4}$	$\frac{b-2c}{4}$	$\frac{b-c}{2}$	$\frac{b-3c}{6}$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$\frac{3b-2c}{4}$	$\frac{b-3c}{6}$	$\frac{b-c}{2}$	$\frac{b-c}{2}$	$\frac{3b-2c}{4}$	$\mathbf{\frac{b-c}{2}}$	$\frac{5b-3c}{6}$	$\frac{5b-3c}{6}$	$\frac{2b-c}{2}$
S_{13}	$-c$	$-c$	$\frac{b-2c}{2}$	$\frac{b-2c}{2}$	$-c$	$-c$	$\frac{3b-4c}{5}$	$\frac{3b-4c}{5}$	$\frac{b-3c}{3}$	$\frac{2b-3c}{3}$	$b-c$	$b-c$	$\frac{3b-5c}{6}$	$\mathbf{\frac{3b-3c}{4}}$	$\frac{3b-2c}{3}$	$\frac{3b-2c}{3}$
S_{14}	$-\frac{2c}{3}$	$-c$	$\frac{2b-4c}{5}$	$\frac{b-2c}{2}$	$-\frac{2c}{3}$	$-c$	$\frac{2b-4c}{5}$	$\frac{b-2c}{2}$	$\frac{b-3c}{4}$	$\frac{b-3c}{3}$	$b-c$	$b-c$	$\frac{3b-5c}{6}$	$\frac{2b-3c}{3}$	$\mathbf{b-c}$	$b-c$
S_{15}	$-c$	$-c$	$\frac{b-2c}{2}$	$\frac{b-2c}{2}$	$-c$	$-c$	$\frac{b-2c}{2}$	$\frac{b-2c}{2}$	$\frac{b-3c}{3}$	$\frac{b-2c}{2}$	$b-c$	$b-c$	$\frac{b-2c}{2}$	$\frac{2b-3c}{3}$	$b-c$	$\mathbf{b-c}$

Table 1: Payoffs for the row player when both players employ deterministic memory-1 strategy in an infinitely repeated donation game with implementation errors[25].The $(p_{\mathrm{CC}}, p_{\mathrm{CD}}, p_{\mathrm{DC}}, p_{\mathrm{DD}})$ form of the memory-1 strategy S_k is Binary(k) with p_{DD} being the least significant bit.

References

[1] R. L. Trivers. The evolution of reciprocal altruism. *The Quarterly review of biology*, 46(1):35–57, 1971.

[2] R. Axelrod. The evolution of cooperation. *Basic Books, New York, NY*, 1984.

[3] M. A. Nowak. Five rules for the evolution of cooperation. *Science*, 314(5805):1560–1563, 2006.

[4] D. S. Wilson and E. O. Wilson. Rethinking the theoretical foundation of sociobiology. *The Quarterly review of biology*, 82(4):327–348, 2007.

[5] L. A. Dugatkin. Cooperation in animals: an evolutionary overview. *Biology and Philosophy*, 17(4):459–476, 2002.

[6] R. L. Riolo, M. D. Cohen, and R. Axelrod. Evolution of cooperation without reciprocity. *Nature*, 414(6862):441–443, 2001.

[7] M. A. Zaggl. Eleven mechanisms for the evolution of cooperation. *Journal of Institutional Economics*, 10(2):197–230, 2014.

[8] A. Traulsen and M. A. Nowak. Evolution of cooperation by multilevel selection. *Proceedings of the National Academy of Sciences*, 103(29):10952–10955, 2006.

[9] M. A. Nowak et al. Evolving cooperation. *Journal of theoretical biology*, 299(0):1–8, 2012.

[10] M. Perc, J. J. Jordan, D. G. Rand, Z. Wang, S. Boccaletti, and A. Szolnoki. Statistical physics of human cooperation. *Physics Reports*, 687:1–51, 2017.

[11] M. Perc and A. Szolnoki. Coevolutionary games—a mini review. *BioSystems*, 99(2):109–125, 2010.

[12] F. Fu, C. Hauert, M. A. Nowak, and L. Wang. Reputation-based partner choice promotes cooperation in social networks. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, 78(2):026117, 2008.

[13] M. Milinski. Tit for tat in sticklebacks and the evolution of cooperation. *Nature*, 325(6103):433–435, 1987.

[14] D. Fundenberg and E. Maskin. Evolution and cooperation in noisy repeated games. *The American Economic Review*, 80(2):274–279, 1990.

[15] M. A. Nowak and K. Sigmund. Tit for tat in heterogeneous populations. *Nature*, 355(6357):250–253, 1992.

[16] M. Van Veelen, J . García, D. G. Rand, and M. A. Nowak. Direct reciprocity in structured populations. *Proceedings of the National Academy of Sciences*, 109(25):9929–9934, 2012.

[17] A. J. Stewart and J. B. Plotkin. From extortion to generosity, evolution in the iterated prisoner's dilemma. *Proceedings of the National Academy of Sciences*, 110(38):15348–15353, 2013.

[18] C. Hilbe, K. Chatterjee, and M. A. Nowak. Partners and rivals in direct reciprocity. *Nature Human Behaviour*, 2(7):469–477, 2018.

[19] X. Chen and F. Fu. Outlearning extortioners: unbending strategies can foster reciprocal fairness and cooperation. *PNAS nexus*, 2(6):pgad176, 2023.

[20] D. Fudenberg and E. Maskin. The folk theorem in repeated games with discounting or with incomplete information. In *A long-run collaboration on long-run games*, pages 209–230. World Scientific, 2009.

[21] C. Hilbe, A. Traulsen, and K. Sigmund. Partners or rivals? strategies for the iterated prisoner's dilemma. *Games and economic behavior*, 92:41–52, 2015.

[22] W. H. Press and F. J. Dyson. Iterated prisoner's dilemma contains strategies that dominate any evolutionary opponent. *Proceedings of the National Academy of Sciences*, 109(26):10409–10413, 2012.

[23] C. Wedekind and M. Milinski. Human cooperation in the simultaneous and the alternating prisoner's dilemma: Pavlov versus generous tit-for-tat. *Proceedings of the National Academy of Sciences*, 93(7):2686–2689, 1996.

[24] M. Nowak and K. Sigmund. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature*, 364(6432):56–58, 1993.

[25] K. Sigmund. *The calculus of selfishness*. Princeton University Press, 2010.

[26] M. Nowak and K. Sigmund. Chaos and the evolution of cooperation. *Proceedings of the National Academy of Sciences*, 90(11):5091–5094, 1993.

[27] M. A. Nowak. *Evolutionary dynamics: exploring the equations of life*. Harvard university press, 2006.

[28] B. Wu, B. Bauer, T. Galla, and A. Traulsen. Fitness-based models and pairwise comparison models of evolutionary games are typically different—even in unstructured populations. *New Journal of Physics*, 17(2):023043, 2015.

[29] B M. Zagorsky, J G. Reiter, K. Chatterjee, and M. A. Nowak. Forgiver triumphs in alternating prisoner's dilemma. *PLoS One*, 8(12):e80814, 2013.

[30] P. D. Taylor and L. B. Jonker. Evolutionary stable strategies and game dynamics. *Mathematical biosciences*, 40(1-2):145–156, 1978.

[31] J. Hofbauer and K. Sigmund. *Evolutionary games and population dynamics*. Cambridge university press, 1998.

[32] J. Hofbauer and W. H. Sandholm. Survival of dominated strategies under evolutionary dynamics. *Theoretical Economics*, 6(3):341–377, 2011.

[33] W. H. Sandholm. *Population games and evolutionary dynamics*. MIT press, 2010.

[34] M. A. Nowak, N. L. Komarova, and P. Niyogi. Evolution of universal grammar. *Science*, 291(5501):114–118, 2001.

[35] L. A. Imhof, D. Fudenberg, and M. A. Nowak. Evolutionary cycles of cooperation and defection. *Proceedings of the National Academy of Sciences*, 102(31):10797–10800, 2005.

[36] N. L. Komarova. Replicator–mutator equation, universality property and population dynamics of learning. *Journal of theoretical biology*, 230(2):227–239, 2004.

[37] M. H. Duong and T. A. Han. On equilibrium properties of the replicator–mutator equation in deterministic and random games. *Dynamic Games and Applications*, 10:641–663, 2020.

[38] D. FP. Toupo and S. H. Strogatz. Nonlinear dynamics of the rock-paper-scissors game with mutations. *Physical Review E*, 91(5):052907, 2015.

[39] M. Mobilia. Oscillatory dynamics in rock–paper–scissors games with mutations. *Journal of Theoretical Biology*, 264(1):1–10, 2010.

[40] D. FP. Toupo, D.G. Rand, and S. H. Strogatz. Limit cycles sparked by mutation in the repeated prisoner's dilemma. *International Journal of Bifurcation and Chaos*, 24(12):1430035, 2014.

[41] M. Nowak and K. Sigmund. Game-dynamical aspects of the prisoner's dilemma. *Applied Mathematics and Computation*, 30(3):191–213, 1989.

[42] M. Kandori, G. J. Mailath, and R. Rob. Learning, mutation, and long run equilibria in games. *Econometrica: Journal of the Econometric Society*, pages 29–56, 1993.

[43] A. Eriksson and K. Lindgren. Cooperation driven by mutations in multi-person prisoner's dilemma. *Journal of theoretical biology*, 232(3):399–409, 2005.

[44] D. Helbing, A. Szolnoki, M. Perc, and G. Szabó. Defector-accelerated cooperativeness and punishment in public goods games with mutations. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, 81(5):057104, 2010.

[45] M. Spichtig and M. Egas. When and how does mutation-generated variation promote the evolution of cooperation? *Games*, 10(1):4, 2019.

[46] G. Ichinose, Y. Satotani, and H. Sayama. How mutation alters the evolutionary dynamics of cooperation on networks. *New Journal of Physics*, 20(5):053049, 2018.

[47] K. Feng, S. Han, M. Feng, and A. Szolnoki. An evolutionary game with reputation-based imitation-mutation dynamics. *Applied Mathematics and Computation*, 472:128618, 2024.

[48] J. Tkadlec, C. Hilbe, and M. A. Nowak. Mutation enhances cooperation in direct reciprocity. *Proceedings of the National Academy of Sciences*, 120(20):e2221080120, 2023.

[49] D. G. Rand, H. Ohtsuki, and M. A. Nowak. Direct reciprocity with costly punishment: Generous tit-for-tat prevails. *Journal of theoretical biology*, 256(1):45–57, 2009.

[50] K. Panchanathan and R. Boyd. Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature*, 432(7016):499–502, 2004.

[51] H. Brandt and K. Sigmund. The good, the bad and the discriminator—errors in direct and indirect reciprocity. *Journal of theoretical biology*, 239(2):183–194, 2006.

[52] N. E. Glynatsi, M.A. Nowak, and C. Hilbe. Conditional cooperation with longer memory. *arXiv preprint arXiv:2402.02437*, 2024.

[53] M. A. Nowak and K. Sigmund. Evolution of indirect reciprocity. *Nature*, 437(7063):1291–1298, 2005.

[54] M. A. Nowak and K. Sigmund. Evolution of indirect reciprocity by image scoring. *Nature*, 393(6685):573–577, 1998.

[55] A. Szolnoki, G. Szabó, and L. Czakó. Competition of individual and institutional punishments in spatial public goods games. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, 84(4):046106, 2011.

[56] HW Lee, C. Cleveland, and A. Szolnoki. When costly migration helps to improve cooperation. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 32(9), 2022.

[57] M. A. Nowak, A. Sasaki, and D. Taylor, C .and Fudenberg. Emergence of cooperation and evolutionary stability in finite populations. *Nature*, 428(6983):646–650, 2004.

[58] M. A. Nowak and K. Sigmund. The dynamics of indirect reciprocity. *Journal of theoretical Biology*, 194(4):561–574, 1998.

[59] H. Ohtsuki and Y. Iwasa. The leading eight: social norms that can maintain cooperation by indirect reciprocity. *Journal of theoretical biology*, 239(4):435–444, 2006.

[60] F. P. Santos, F. C. Santos, and J. M. Pacheco. Social norm complexity and past reputations in the evolution of cooperation. *Nature*, 555(7695):242–245, 2018.

[61] C. Hilbe, L. Schmid, J. Tkadlec, K. Chatterjee, and M. A. Nowak. Indirect reciprocity with private, noisy, and incomplete information. *Proceedings of the national academy of sciences*, 115(48):12241–12246, 2018.

[62] L. Schmid, K. Chatterjee, C. Hilbe, and M. A. Nowak. A unified framework of direct and indirect reciprocity. *Nature Human Behaviour*, 5(10):1292–1302, 2021.

[63] S. Pal, C. Hilbe, and N. E. Glynatsi. The co-evolution of direct, indirect and generalized reciprocity. *arXiv preprint arXiv:2411.03488*, 2024.

[64] M. A. Nowak and R. M. May. Evolutionary games and spatial chaos. *Nature*, 359(6398):826–829, 1992.

[65] E. Lieberman, C. Hauert, and M. A. Nowak. Evolutionary dynamics on graphs. *Nature*, 433(7023):312–316, 2005.

[66] H. Ohtsuki, C. Hauert, E. Lieberman, and M. A. Nowak. A simple rule for the evolution of cooperation on graphs and social networks. *Nature*, 441(7092):502–505, 2006.

[67] B. Allen, G. Lippner, YT Chen, B. Fotouhi, N. Momeni, ST Yau, and M. A. Nowak. Evolutionary dynamics on any population structure. *Nature*, 544(7649):227–230, 2017.

[68] M. C. Couto and S. Pal. Introspection dynamics in asymmetric multiplayer games. *Dynamic Games and Applications*, 13(4):1256–1285, 2023.

[69] [DATA] Stabilizing Unconditional Cooperation, 2024. Available online at: `https://zenodo.org/records/14246418`.

Supplementary Information

Stabilizing unconditional cooperation

1 The Replicator-Mutator equation I

We consider an infinitely large population where the relative frequency of ALLC, TFT and ALLD are x, y and z respectively (with $x + y + z = 1$). The rate of change of these relative frequencies follow the dynamical equations,

$$\begin{aligned} \frac{\mathrm{d}x}{\mathrm{d}t} &= x(f_x - \bar{f}) - \mu x, \\ \frac{\mathrm{d}y}{\mathrm{d}t} &= y(f_y - \bar{f}) + \mu x, \\ \frac{\mathrm{d}z}{\mathrm{d}t} &= z(f_z - \bar{f}). \end{aligned} \tag{D1}$$

Here f_x, f_y and f_z are the fitness of the three types, given by their average payoff in the well-mixed population. These are,

$$\begin{aligned} f_x &= (b-c)(x+y) - cz \\ f_y &= (b-c)x + \left(\frac{b-c}{2}\right) y \\ f_z &= bx \end{aligned} \tag{1}$$

The term $\bar{f}$ denotes the average fitness of an individual in the population. That is, $\bar{f} = xf_x + yf_y + zf_z$. The term $\mu (\geq 0)$ here is the unconditional mutation rate from ALLC to TFT. In this model, we assume that mutations from ALLC to TFT occur whenever the source, ALLC, is present in the population ($x > 0$). In addition, the rate at which mutated TFT offsprings arise is proportional to x with proportionality constant, μ.

By adding up the dynamical equations in Eq. (D1), we get $\dot{x}+\dot{y}+\dot{z}=0$. This implies that the triangular simplex $\Delta_2 := \{(x,y,z) \in \mathbb{R}^3 \mid x+y+z=1, x,y,z \geq 0\}$ is invariant to (D1). If the dynamics begins in Δ_2 it remains there forever. We only focus on the properties of (D1) in this set. Additionally, subsets $\{(x,y,z) \in \Delta_2 \mid z=0\}$ and $\{(x,y,z) \in \Delta_2 \mid x=0\}$ of Δ_2 are also invariant to Eq. (D1). These subsets form two edges of the triangular simplex in which ALLD and ALLC are respectively absent. Finally, the last edge, $\{(x,0,z) \mid x+z=1\}$ is only invariant if $\mu=0$, otherwise not. The proposition below characterizes all rest points of (D1) in Δ_2.

Proposition 1 (Rest points of (D1) in Δ_2)**.** *For the replicator-mutator equation (D1), the following are the rest points in Δ_2 provided associated conditions (if any) are met.*

1. $\mathbf{x}_{\mathrm{TFT}} := (0,1,0)$,

2. $\mathbf{x}_{\mathrm{ALLD}} := (0,0,1)$,

3. $\mathbf{x}_{\mathrm{ALLC}} := (1,0,0)$ *if and only if* $\mu = 0$,

4. $\mathbf{x}_{\mathrm{E}} := \left(1-\sqrt{\frac{2\mu}{b-c}},\ \sqrt{\frac{2\mu}{b-c}},\ 0\right)$, *if and only if* $0 < \mu < \frac{b-c}{2}$,

5. $\mathbf{x}_{\mathrm{I}} := \left(\frac{(b-c)(c+\mu)^2}{2b(c^2-\mu(b-c))},\ \frac{c+\mu}{b},\ 1-\frac{c+\mu}{b}-\frac{(b-c)(c+\mu)^2}{2b(c^2-\mu(b-c))}\right)$ *if and only if* $0 \leq \mu < \mu_1$.

where, μ_1 is given by,

$$\mu_1 := b\left(1-\sqrt{1+\frac{c^2}{(b-c)^2}}\right)+\frac{c^2}{b-c}. \tag{2}$$

We include proofs of Propositions in the section **Proofs**. The above proposition lists five possible rest points of (D1) in Δ_2. The first three, $\mathbf{x}_{\mathrm{TFT}}$, $\mathbf{x}_{\mathrm{ALLD}}$, and $\mathbf{x}_{\mathrm{ALLC}}$ are three corners of the simplex. In these points, the population only consists of TFT, ALLD or ALLC respectively. The fourth rest point, $\mathbf{x}_{\mathrm{E}}$, when it exists, is strictly in the interior of the edge where ALLD is absent (i.e., both ALLC and TFT are present). Finally, the last rest point, $\mathbf{x}_{\mathrm{I}}$ when it exists, is strictly in the interior of Δ_2 (i.e., all strategies coexist). We note the following degenerate cases: a) when $\mu=0$, $\mathbf{x}_{\mathrm{E}} = \mathbf{x}_{\mathrm{ALLC}}$, b) when $\mu=(b-c)/2$, $\mathbf{x}_{\mathrm{E}} = \mathbf{x}_{\mathrm{TFT}}$. We also note that $\mu_1 < \mu_2 := (b-c)/2$ under the assumption that $b > c > 0$ (see end of Proof of Proposition 1). In the proposition below, we characterize the dynamic stability of these rest points.

Proposition 2 (Asymptotic stability of rest points of (D1))**.** *The following statements describe the asymptotic stability of rest points of (D1) with respect to perturbations in* Δ_2*. Each statement subsumes that the relevant condition for the fixed point's existence is met (from Proposition 1).*

1. *The rest points,* $\mathbf{x}_{\mathrm{ALLD}}$ *and* $\mathbf{x}_{\mathrm{ALLC}}$*, lying at corners of* Δ_2*, are not stable.*
2. *The third corner rest point,* $\mathbf{x}_{\mathrm{TFT}}$*, is stable if and only if* $\mu > (b-c)/2$*.*
3. *The rest point* $\mathbf{x}_{\mathrm{E}}$*, that lies on the boundary where ALLD is absent, is stable if and only if* $\mu > \mu_1$*.*
4. *The rest point,* $\mathbf{x}_{\mathrm{I}}$*, lying strictly in the interior of* Δ_2*, is stable if and only if* $\mu > 0$*.*

Furthermore, if the fixed point $\mathbf{x}_{\mathrm{E}}$ *exists, any solution starting from the set* $\{(x, 1-x, 0) \mid 0 < x \leq 1\}$*, asymptotically converges to* $\mathbf{x}_{\mathrm{E}}$ *(the set includes* $x = 1$ *if* $\mu > 0$*, otherwise not). Additionally, at* $\mu = 0$*, the interior fixed point* $\mathbf{x}_{\mathrm{I}}$ *acts as a center for periodic orbits in* Δ_2*.*

First, the Proposition claims that rest points corresponding to homogenous populations of ALLC and ALLD are not stable. We show that introducing TFT in small amounts to an ALLD population destabilizes it. Similarly, introducing ALLD destabilizes an ALLC population. When $\mu \leq (b-c)/2$, a homogenous TFT population is also not stable since a small introduction of ALLC grows. However, if mutations are too frequent, $\mu > (b-c)/2$, a homogenous TFT population is stable. Second, the Proposition claims that the value of μ determines the stability of the ALLC-TFT coexistence. This coexistence is stable when $\mu_1 < \mu < (b-c)/2$. When $\mu \leq \mu_1$, this coexistence is not stable; a small introduction of ALLD into this mixed population grows. Finally, the Proposition claims that the three types — ALLC, TFT and ALLD — can stably co-exist, if conditions allow for such a coexistence in the first place *and* the mutation rate is positive (that is, $0 < \mu < \mu_1$). At $\mu = 0$, this coexistence acts as a center for periodic orbits. We also show that any population containing ALLC but no ALLD, asymptotically reaches this ALLC-TFT coexistence. This also implies that this coexistence is always stable to perturbations that do not involve ALLD.

Corollary 1. *The strategy ALLC attains maximal frequency at a stable coexistence when* $\mu = \mu_1$*.*

Proposition 3 (Extinction of ALLD from any mixed population when μ is sufficiently high)**.** *If* $\mu \geq \mu_1$*, any solution* $(x(t), y(t), z(t))$ *of (D1) with* $(x(0), y(0), z(0)) \in \Delta_2$ *such that* $z(0) \in [0, 1)$ *has* $\lim_{t\to\infty} z(t) = 0$*.*

This Proposition claims that the dynamics (D1) eliminates ALLD from any starting population that has ALLD and other strategies (either TFT, ALLC or both), provided the mutation rate, μ is atleast μ_1.

2 The Replicator-Mutator equation II

We now study an alternate Replicator-Mutator dynamic in the infinite population. In this dynamic, the rate at which ALLC mutates to TFT is proportional to both the frequency of ALLC and fitness of ALLC. The types evolve according to the equations

$$\begin{aligned} \frac{\mathrm{d}x}{\mathrm{d}t} &= x(g_x(1-\mu) - \bar{g}), \\ \frac{\mathrm{d}y}{\mathrm{d}t} &= y(g_y - \bar{g}) + \mu x g_x, \\ \frac{\mathrm{d}z}{\mathrm{d}t} &= z(g_z - \bar{g}). \end{aligned} \tag{D2}$$

Here g is the effective fitness of each type, defined as $g_i = f_i + c$. The term f_i is the average payoff of type i, as expressed in Eq. (1), and c, the cost of cooperation, is the shared baseline fitness of the types. In this model $0 \leq \mu \leq 1$. We define $\bar{g} := xg_x + yg_y + zg_z = \bar{f} + c$ as the average fitness of the population. We add c to each type's fitness to ensure non-negative rate of growth for each type in Δ_2. We find that the dynamics (D1) and (D2) qualitatively share the same phase portraits in Δ_2, except that their bifurcation points in μ are different. For a comparison, see Fig. 5 from main-text in which we plot phase potraits of (D1) and Fig. S5. in which we plot the phase potraits of (D2). We summarize the dynamical properties of (D2) in the propositions below[a].

Proposition 4 (Rest points of (D2) in Δ_2)**.** *For the replicator-mutator equation (D2), the following are the rest points in Δ_2 provided associated conditions (if any) are met.*

1. $\mathbf{x}_{\text{TFT}} := (0, 1, 0)$,
2. $\mathbf{x}_{\text{ALLD}} := (0, 0, 1)$,
3. $\mathbf{x}_{\text{ALLC}} := (1, 0, 0)$ *if and only if* $\mu = 0$,
4. $\mathbf{x}_{\text{E},2} := \left(1 - \sqrt{\frac{2b\mu}{b-c}}, \sqrt{\frac{2b\mu}{b-c}}, 0\right)$, *if and only if* $0 < \mu < \frac{b-c}{2b}$,

---

[a] We checked using the symbolic computation software Mathematica that adding any baseline fitness $\alpha > c$ to each type's fitness also yields qualitatively similar phase portraits to dynamic (D1). We avoid that formulation because of lack of clean expressions.

5. $\mathbf{x}_{\mathrm{I},2} := (x_{\mathrm{I},2}, y_{\mathrm{I},2}, 1 - x_{\mathrm{I},2} - y_{\mathrm{I},2})$ *if and only if* $\quad 0 \leq \mu < \mu_1/b := \mu_2$

where, $x_{\mathrm{I},2}$ *and* $y_{\mathrm{I},2}$ *are given by,*

$$x_{\mathrm{I},2} := \begin{cases} \dfrac{c^2 - bc(2-\mu)\mu - (1-\mu)c\sqrt{c^2 - b^2(2-\mu)\mu}}{b(b-c)(2-\mu)\mu} & \mu > 0 \\ \dfrac{b-c}{2b} & \mu = 0 \end{cases} \tag{3}$$

$$y_{\mathrm{I},2} := \frac{bc(2-\mu) - c^2 - c\sqrt{c^2 - b^2(2-\mu)\mu}}{b(b-c)(2-\mu)} \tag{4}$$

Similar to (D1), the new dynamic (D2) has five rest points. The corners corresponding to TFT and ALLD populations are rest points. An ALLC population is a rest point only if $\mu = 0$. There is an interior rest point if $\mu \in [0, \mu_2)$ and there is a rest point strictly in the interior of the ALLC-TFT edge if $\mu \in (0, (b-c)/(2b))$. Moreover, like before $\mathbf{x}_{\mathrm{TFT}} = \mathbf{x}_{\mathrm{E},2}$ at $\mu = (b-c)/(2b)$ and $\mathbf{x}_{\mathrm{ALLC}} = \mathbf{x}_{\mathrm{E}}$ at $\mu = 0$.

Proposition 5 (Asymptotic stability of rest points of (D2))**.** *The following statements describe the asymptotic stability of rest points of (D2) with respect to perturbations in* Δ_2*. Each statement subsumes that the relevant condition for the fixed point's existence is met (from Proposition 4).*

1. *The rest points,* $\mathbf{x}_{\mathrm{ALLD}}$ *and* $\mathbf{x}_{\mathrm{ALLC}}$*, lying at corners of* Δ_2*, are not stable.*
2. *The third corner rest point,* $\mathbf{x}_{\mathrm{TFT}}$*, is stable if and only if* $\mu > (b-c)/(2b)$.
3. *The rest point* $\mathbf{x}_{\mathrm{E},2}$*, that lies on the boundary where ALLD is absent, is stable if and only if* $\mu > \mu_2$.
4. *The rest point,* $\mathbf{x}_{\mathrm{I},2}$*, lying strictly in the interior of* Δ_2*, is stable if and only if* $\mu > 0$.

Furthermore, if the fixed point $\mathbf{x}_{\mathrm{E},2}$ *exists, any solution starting from the set* $\{(x, 1-x, 0) \mid 0 < x \leq 1\}$*, asymptotically converges to* $\mathbf{x}_{\mathrm{E},2}$ *(the set includes* $x = 1$ *if* $\mu > 0$*, otherwise not). Additionally, at* $\mu = 0$*, the interior fixed point* $\mathbf{x}_{\mathrm{I},2}$ *acts as a center for periodic orbits in* Δ_2*.*

The proof of these Propositions on the rest points of (D2) and their stability proceeds identically as the Proofs of Proposition 1 and 2.

Proofs

Proof of Proposition 1. We represent the dynamics (D1) in vector format as

$$\dot{\mathbf{x}} = (\mathrm{G}_x(\mathbf{x}), \mathrm{G}_y(\mathbf{x}), \mathrm{G}_z(\mathbf{x})) \tag{5}$$

where $\mathbf{x} := (x, y, z)$, and $\mathrm{G}_x(\mathbf{x}) = x(f_x - \bar{f}) - \mu x$, $\mathrm{G}_y(\mathbf{x}) = y(f_y - \bar{f}) + \mu x$, and $\mathrm{G}_z(\mathbf{x}) = z(f_z - \bar{f})$. A rest point, $\bar{\mathbf{x}} := (\bar{x}, \bar{y}, \bar{z})$, of the dynamic is defined as a point where $\dot{\mathbf{x}} = \mathbf{0}$, or equivalently,

$$\mathrm{G}_x(\bar{\mathbf{x}}) = \mathrm{G}_y(\bar{\mathbf{x}}) = \mathrm{G}_z(\bar{\mathbf{x}}) = 0. \tag{6}$$

Since we are only interested in rest points that lie in Δ_2, we use $\bar{x} + \bar{y} + \bar{z} = 1$.

Rest points at corners of Δ_2: The corners $(0, 1, 0)$ and $(0, 0, 1)$ always satisfy these conditions and are thus rest points, unconditionally. In addition, if $\mu = 0$, the remaining corner, $(1, 0, 0)$, is also a rest point.

Rest points at the boundary of Δ_2: We now search for rest points in sets corresponding to the interior of the boundary edges of Δ_2 (i.e., edges excluding the corners). These are a) the TFT-ALLD edge: $\{(0, y, 1-y) \in \Delta_2 \mid y \in (0, 1)\}$, b) the ALLC-ALLD edge: $\{(x, 0, 1-x) \in \Delta_2 \mid x \in (0, 1)\}$, and lastly c) the ALLC-TFT edge: $\{(x, 1-x, 0) \in \Delta_2 \mid x \in (0, 1)\}$. At the TFT-ALLD edge, $\dot{x} = 0$, but

$$\frac{\mathrm{d}y}{\mathrm{d}t} = y^2 \left(\frac{b-c}{2}\right)(1-y) > 0. \tag{7}$$

Thus, there is no rest point in this set. At the ALLC-ALLD edge, $\dot{y} = 0$ only if $\mu = 0$. However, at $\mu = 0$

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -x(1-x)c < 0 \tag{8}$$

on this set. So, there is no rest point in this set either. Finally at the ALLC-TFT edge, $\dot{z} = 0$ and

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{x(b-c)}{2}\left((1-x)^2 - \frac{2\mu}{b-c}\right). \tag{9}$$

There is a rest point in this set if $0 < \mu \leq (b-c)/2$ and this rest point is given by

$$\bar{\mathbf{x}} = \mathbf{x}_{\mathrm{E}} := \left(1 - \sqrt{\frac{2\mu}{b-c}}, \sqrt{\frac{2\mu}{b-c}}, 0\right) \tag{10}$$

Rest points strictly in the interior Δ_2: Finally we search for rest points that belong in the subset of Δ_2

where at every point, all strategies are present, $int\Delta_2 := \{\mathbf{x} \in \Delta_2 \mid x, y, z \geq 0\}$. By using $\mathrm{G}_y(\bar{\mathbf{x}}) = \mathrm{G}_z(\bar{\mathbf{x}}) = 0$, we obtain that a fixed point $\bar{\mathbf{x}}$ must respect the following two equations simultaneously,

$$\bar{x}\left(\bar{y}^2(b-c)-\mu\right) = \bar{y}^2(1-\bar{y})\left(\frac{b-c}{2}\right) \tag{11}$$

$$\bar{y}^2(b-c)(1-y)(b\bar{y}-c-\mu)(\bar{y}^2(b-c)-2\mu) = 0 \tag{12}$$

Besides the already evaluated fixed points, the point which satisfies these conditions simultaneously are

$$\mathbf{x}_{\mathrm{I}} := \left(\frac{(b-c)(c+\mu)^2}{2b(c^2-\mu(b-c))},\ \frac{c+\mu}{b},\ 1-\frac{c+\mu}{b}-\frac{(b-c)(c+\mu)^2}{2b(c^2-\mu(b-c))}\right) \tag{13}$$

provided $\mu \neq c^2/(b-c)$. In order for $\mathbf{x}_{\mathrm{I}}$ to be a fixed point in $int\Delta_2$, the elements must satisfy $0 < x_{\mathrm{I}}, y_{\mathrm{I}}, z_{\mathrm{I}} < 1$. These give rise to the following conditions on μ,

$$x_{\mathrm{I}} > 0 \iff \mu < \frac{c^2}{b-c} \tag{14}$$

$$x_{\mathrm{I}} < 1 \iff \mu < -(b+c)+\sqrt{\frac{(b+c)(b^2+3c^2)}{b-c}} =: \mu' \tag{15}$$

$$y_{\mathrm{I}} < 1 \iff \mu < b-c \tag{16}$$

$$z_{\mathrm{I}} > 0 \iff \mu < \frac{c^2}{b-c} - b\left(\sqrt{1+\frac{c^2}{(b-c)^2}}-1\right) =: \mu_1 \tag{17}$$

We do not explicitly include $y_{\mathrm{I}} > 0$ since it is trivial and always holds. Also, since these conditions simultaneously imply $z_{\mathrm{I}} < 1$, we do not write it explicitly. From the above conditions, one derives that $z_{\mathrm{I}} > 0 \implies x_{\mathrm{I}} > 0$ and $y_{\mathrm{I}} < 1$. For the latter, one needs to subtract μ_1 from $b-c$ to check that the difference is always positive. Thus, $\mu < \mu_1 \implies \mu < b-c$. Finally we compare the conditions for $z_{\mathrm{I}} > 0$ and $x_{\mathrm{I}} < 1$. We define a variable $u = 1/(1-c/b)$ that lies in $(1,\infty)$ as $b > c > 0$. The difference between μ' and μ_1 in terms of u can be simplified as

$$\frac{1}{b}\left(\mu'-\mu_1\right) = -(1+u)+\sqrt{(2u-1)\left(1+3\left(\frac{u-1}{u}\right)^2\right)}+\sqrt{(u-1)^2+1} \tag{18}$$

$$\geq -(1+u)+\sqrt{(2u-1)}+\sqrt{(u-1)^2+1} \quad \geq 0 \quad \text{if } u \in (1,\infty) \tag{19}$$

The relation $\mu' \geq \mu_1$ means that $z_{\mathrm{I}} > 0$ implies $x_{\mathrm{I}} < 1$ and so $\mu < \mu_1$ is the only relevant condition for the existence of the interior fixed point $\mathbf{x}_{\mathrm{I}}$.

Finally we show that under the assumption $b > c > 0$, $\mu_1 < \mu_2 := (b-c)/2$. Some algebraic steps verify that the sign of $\mu_2 - \mu_1$ is the same as the sign of $2b\sqrt{b^2 - 2bc + 2c^2} - (b^2 + c^2)$ if one assumes $b > c > 0$. Considering $t = c/b$ this is equivalent to checking the sign of $L(t) := 2\sqrt{1 - 2t + 2t^2} - (1 + t^2)$ for $t \in (0, 1)$. We note that $L(0) = 1$, $L(1) = 0$ and

$$\frac{\mathrm{d}L}{\mathrm{d}t} = -2t - \frac{2(1-2t)}{\sqrt{1-2t+2t^2}} < 2(t-1) < 0 \tag{20}$$

for $t \in (0, 1)$ and therefore $L(t) > 0$ in this interval which implies $\mu_2 > \mu_1$ for $b > c > 0$. ☐

Proof of Proposition 2. Stability of $\mathbf{x}_{\mathrm{ALLD}}$: To show that $\mathbf{x}_{\mathrm{ALLD}}$ is unstable, we take a point $\mathbf{x}_{\mathrm{ALLD}}(\varepsilon) := (0, 1-\varepsilon, \varepsilon)$ in the neighbourhood of $\mathbf{x}_{\mathrm{ALLD}}$. This point lies in the edge connecting $\mathbf{x}_{\mathrm{TFT}}$ and $\mathbf{x}_{\mathrm{ALLD}}$. All solutions starting in this set remain here (i.e., $\dot{x} = 0$). For any $\varepsilon > 0$, it is easy to verify that $\dot{z} < 0$ at $\mathbf{x}_{\mathrm{ALLD}}(\varepsilon)$. Therefore $\mathbf{x}_{\mathrm{ALLD}}$ is not stable.

Stability of $\mathbf{x}_{\mathrm{ALLC}}$: The eigenvalues of the Jacobian of the right hand side G of the dynamic (D1) at $\mathbf{x}_{\mathrm{ALLC}}$ is given by

$$\Lambda_{\mathrm{G}}(\mathbf{x}_{\mathrm{ALLC}}) = (c, -(b-c), -\mu) \tag{21}$$

Since we assume $c > 0$, at $\mu > 0$, there is a single eigenvalue that is positive. The eigenvector corresponding to this eigenvalue is

$$\left(-\frac{c}{c+\mu}, -\frac{\mu}{c+\mu}, 1\right).$$

This eigenvector is parallel to the simplex and thus $\mathbf{x}_{\mathrm{ALLC}}$ is not stable when $\mu > 0$. At $\mu = 0$, when the one of the eigenvalues is 0, we observe that at any point $\mathbf{x}_{\mathrm{ALLC}}(\varepsilon) := (\varepsilon, 0, 1-\varepsilon)$ on the invariant ALLC-ALLD edge, $\dot{x} < 0$. So, $\mathbf{x}_{\mathrm{ALLC}}$ not stable for $\mu = 0$ either, and thus never stable.

Stability of $\mathbf{x}_{\mathrm{TFT}}$: The eigenvalues of the Jacobian of the right hand side G of the dynamic (D1) at $\mathbf{x}_{\mathrm{TFT}}$ is given by

$$\Lambda_{\mathrm{G}}(\mathbf{x}_{\mathrm{TFT}}) = \left(-\frac{b-c}{2}, -\frac{b-c}{2}, \left(\frac{b-c}{2} - \mu\right)\right) \tag{22}$$

Since we assume $b > c > 0$, the first two eigenvalues are strictly negative. The third eigenvalue is negative (or positive) when μ is strictly more (or strictly less) than $(b-c)/2$. The eigenvector corresponding to this eigenvalue, $(-1, 1, 0)$, is parallel to the simplex and thus $\mathbf{x}_{\mathrm{TFT}}$ is stable (or not) when μ is greater (or less) than $(b-c)/2$.

Stability of $\mathbf{x}_{\mathrm{E}}$: The eigenvalues of the Jacobian of the right-hand side G around the rest point $\mathbf{x}_{\mathrm{E}} = (1-y_{\mathrm{E}}, y_{\mathrm{E}}, 0)$ where $y_{\mathrm{E}} = \sqrt{2\mu/(b-c)}$ is given by

$$\Lambda_{\mathrm{G}}(\mathbf{x}_{\mathrm{E}}) = \left(\mu-(b-c), \quad \sqrt{2}\left(\sqrt{\mu}-\sqrt{\frac{b-c}{2}}\right), \quad \frac{-c(c+\mu)+b(c-\sqrt{2\mu(b-c)}+\mu)}{b-c}\right) \tag{23}$$

Since the fixed point $\mathbf{x}_{\mathrm{E}}$ only exists when $0 < \mu < (b-c)/2$, the first two eigenvalues are strictly negative. The third eigenvalue is strictly negative if and only if

$$b(\mu-\sqrt{2\mu(b-c)}+c)-(c+\mu)c < 0 \implies \tag{24}$$

$$\mu > \frac{c^2}{b-c} - b\left(\sqrt{1+\frac{c^2}{(b-c)^2}}-1\right) = \mu_1 \tag{25}$$

$$\tag{26}$$

We confirm that the eigenvector corresponding to this third eigenvalue

$$\mathbf{v}_\lambda = \left(\frac{c\sqrt{2(b-c)}\,(4\mu-c+b)-c\sqrt{\mu}\,(5b+5c-2\mu)}{\left(\sqrt{2(b-c)}-\sqrt{\mu}\right)\left(\sqrt{2\mu(b-c)}(2c-\mu)+b(\mu-c)+c^2-3c\mu\right)}, \right. \tag{27}$$

$$\left. \frac{\mu\left(\sqrt{2\mu(b-c)}-b+c\right)}{\sqrt{2\mu(b-c)}(2c-\mu)+b(\mu-c)+c^2-3c\mu}, \quad 1\right)$$

which satisfies $\mathbf{v}_\lambda \cdot (1,1,1)^\intercal = 0$. Thus the eigenvector is parallel to the simplex and therefore the sign of its corresponding eigenvalue determines its asymptotic stability in the invariant simplex.

Stablity of $\mathbf{x}_{\mathrm{I}}$: Finally for the rest point $\mathbf{x}_{\mathrm{I}}$, the eigenvalues of the Jacobian has the following form

$$\Lambda_{\mathrm{G}}(\mathbf{x}_{\mathrm{I}}) = \Bigg(-\frac{(b-c)(c+\mu)^2}{2\left(-b\mu + c^2 + c\mu\right)}, \tag{28}$$
$$\frac{1}{2b^2\left(-b\mu + c^2 + c\mu\right)^2}\left(\alpha(b,c,\mu)\mu + i\sqrt{\beta(b,c,\mu)}\right),$$
$$\frac{1}{2b^2\left(-b\mu + c^2 + c\mu\right)^2}\left(\alpha(b,c,\mu)\mu - i\sqrt{\beta(b,c,\mu)}\right)\Bigg)$$

where α and β are polynomials on μ given by

$$\alpha(b,c,\mu) = b^2(b-c)(c+\mu)\left(b\mu - c^2 - c\mu\right) \tag{29}$$
$$\beta(b,c,\mu) = b^2(b-c)^4(c+\mu)^2\left(\mu - \frac{c^2}{b-c}\right)^2\left((c+\mu)^2(b-c)\left(\frac{c^2}{b-c} - \mu\right) + b^2\mu\left(\mu - \frac{2c^2}{b-c}\right)\right) \tag{30}$$

From Proposition 1, we note that this interior fixed point only exists when $\mu < \mu_1 < c^2/(b-c)$. One can confirm that for these values $\alpha < 0$ while β may change signs from positive to negative depending on μ. The eigenvalues $\alpha \pm i\sqrt{\beta}$ do not vanish to zero regardless. At $\mu = 0$, there is one negative real eigenvalue and two purely imaginary complex conjugate eigenvalues. The eigenvectors corresponding to these eigenvalues at $\mu = 0$ are given respectively by

$$v_{\lambda,1} = \left(1, \frac{2b}{b-c}, 1\right) \tag{31}$$
$$v_{\lambda,2} = \left(-\frac{b+ic}{b-ic}, \quad \frac{2ic}{b-ic}, \quad 1\right) \tag{32}$$
$$v_{\lambda,3} = \left(\frac{b+ic}{b-ic}, \quad -\frac{2ic}{b-ic}, \quad 1\right) \tag{33}$$

where $v_{\lambda,1} \cdot (1,1,1)^{\intercal} > 0$ but the eigenvectors corresponding to purely imaginary eigenvalues satisfy $v \cdot (1,1,1)^{\intercal} = 0$. Thus at $\mu = 0$, the conjugate imaginary eigenvalues correspond to directions on the simplex. The fixed point is therefore a center to periodic orbits that lie on the simplex. The analysis of eigenvalues for $\mu > 0$ is cumbersome, so we restore to studying the determinant and trace of the Jacobian which have the following simpler forms,

$$\det(J_{\mathrm{G}}(\mathbf{x}_{\mathrm{I}})) = \frac{(b-c)^2(c+\mu)^4(2b^2\mu - (c+\mu)^2(b-c))}{8b^2(c^2-\mu(b-c))^2}, \tag{34}$$

$$\mathrm{tr}(J_{\mathrm{G}}(\mathbf{x}_{\mathrm{I}})) = -\frac{(b-c)(c^2+4\mu c+3\mu^2)}{2(c^2-\mu(b-c))}. \tag{35}$$

One can check that this determinant and trace are strictly negative simultaneously when the numerator of the determinant is strictly negative which occurs when $\mu < \mu_1$ (the trace is negative when this condition holds as $\mu_1 < c^2/(b-c)$). As none of the eigenvalues are zero for $\mu \in (0, \mu_1)$, the negative signs of the determinant and the trace implies all eigenvalues have negative real parts, and so $\mathbf{x}_{\mathrm{I}}$ is asymptotically stable.

Finally, on the invariant ALLC-TFT edge ($z = 0$), when $0 < \mu < (b-c)/2$, one observes from Eq. (9) that $\dot{x} > 0$ if $0 < x < \mathbf{x}_{\mathrm{E}}$ whereas $\dot{x} < 0$ when $\mathbf{x}_{\mathrm{E}} < x < 1$. $\square$

Proof of Corollary 1. ALLC can exist in a stable coexistence either at $\mathbf{x}_{\mathrm{E}}$ or $\mathbf{x}_{\mathrm{I}}$. Since we assume $b > c > 0$, one can show that $0 < \mu_1 < (b-c)/2$. Thus, for $\mu \in [0, \mu_1)$, the only stable coexistence in which ALLC exists is $\mathbf{x}_{\mathrm{I}}$ and for $\mu \in [\mu_1, (b-c)/2)$, the only stable coexistence in which it exists is $\mathbf{x}_{\mathrm{E}}$. If one observes the expressions for x_{I} and x_{E} (the frequencies of ALLC in $\mathbf{x}_{\mathrm{I}}$ and x_{E} respectively), one notices that they are respectively increasing and decreasing in μ in the respective domains of existence of those rest points. Thus, ALLC attains maximal frequency at a stable coexistence when $\mu = \mu_1$ $\square$

Proof of Proposition 3. We can check that the simplex Δ_2, which is a a plane in $\mathbb{R}^3$, is invariant to the differentiable dynamics (D1). So, we can apply the Poincare-Bendixson theorem[1] on an open and planar subset of $\mathbb{R}^3$ that contains this simplex.

As Δ_2 is invariant to (D1), any solution starting from within is bounded. Therefore, any orbit passing through a point in this simplex has a compact, non-empty ω-limit set. Also, by Proposition 1 and 2, there are only finitely many fixed points, all of which are isolated. So, by Poincare-Bendixson theorem, any orbit passing through a point within this simplex has its ω-limit set as a fixed point or a periodic orbit.

Now we rule out a periodic orbit in Δ_2 for $\mu \geq \mu_1$. The boundary of Δ_2, $bd\Delta_2 := \{(x,y,z) \in \Delta_2 \mid (x=0) \vee (y=0) \vee (z=0)\}$, does not contain a periodic orbit as $(0,1,0)$ is a fixed point. Now, let's assume there is a periodic orbit in the interior of Δ_2. If so, the periodic orbit, which is restricted to a plane, must enclose atleast one fixed point[2]. However, from Proposition 1 and 2, we know that there is no fixed point in the interior of Δ_2 when $\mu \geq \mu_1 > 0$. By contradiction, there is no periodic orbit in the interior of Δ_2 under this condition. Thus, for $\mu \geq \mu_1$, any solution begining in the simplex must converge to a fixed (rest) point.

We examine the limits of solutions that begin from $(x(0), y(0), z(0))$. First we study the case where $z(0) = 0$. In this case the dynamic is restricted in the set $\{\mathbf{x} \in \Delta_2 \mid z = 0\}$ as $\dot{z} = 0$. So, $\lim_{t\to\infty} z(t)$ is trivially 0. Second we study the case where $x(0) = 0$. Here, as $\dot{x} = 0$, the dynamics is restricted to $\{\mathbf{x} \in \Delta_2 \mid x = 0\}$. Additionally, if $z(0) \in [0, 1)$, $\dot{z} \leq 0$ (equality at $z(0) = 0$). For these initial points, $\lim_{t\to\infty} z(t) = 0$.

Now, we study the case where $y(0) = 0$. Unlike the previous two cases, the associated edge of the simplex is not invariant when $\mu > 0$. That is, in addition, if $x(0) \in (0, 1)$, $\dot{y} > 0$ and the solution enters the interior of the simplex: $int\Delta_2 := \Delta_2 \setminus bd\Delta_2$. As previously argued using the Poincare-Bendixson theorem, any solution starting from $int\Delta_2$ must converge to a fixed point when $\mu \geq \mu_1$ (which is the case of interest). We study $\dot{y}$ in a point $(\varepsilon_1, \varepsilon_2, 1 - \varepsilon_1 - \varepsilon_2)$ that is arbitrarily close to the $z = 1$ corner but inside Δ_2. That is, $0 < \varepsilon_1, \varepsilon_2 \ll 1$. We find, upon neglecting third order terms on ε (and assuming $0 < c < b < \infty$),

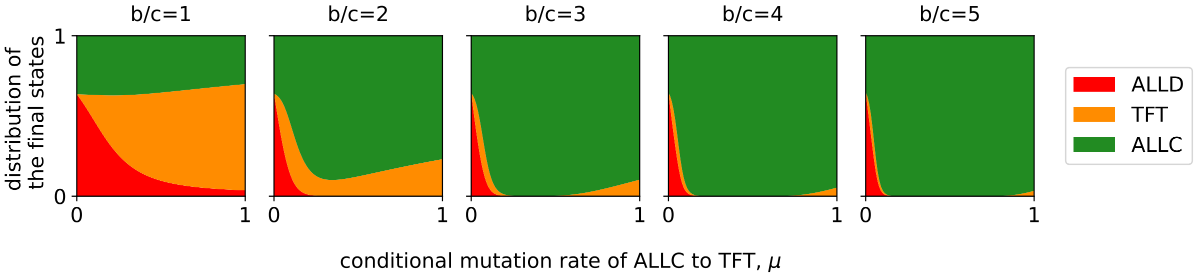
$$\frac{\mathrm{d}y}{\mathrm{d}t} = \big(2\varepsilon_1\mu + \varepsilon_2^2(b - c)\big) > 0 \tag{36}$$

Therefore, no solution from the interior converges to the $z = 1$ corner at which $y = 0$. These solutions can only converge to one of the remaining rest points, all of which satisfy $z = 0$ (QED). One may study $\dot{z}$ at a point $(1 - \varepsilon_1 - \varepsilon_2, \varepsilon_1, \varepsilon_2)$ that is arbitrarily close to $x = 1$ to find that upon neglecting third order terms of ε,

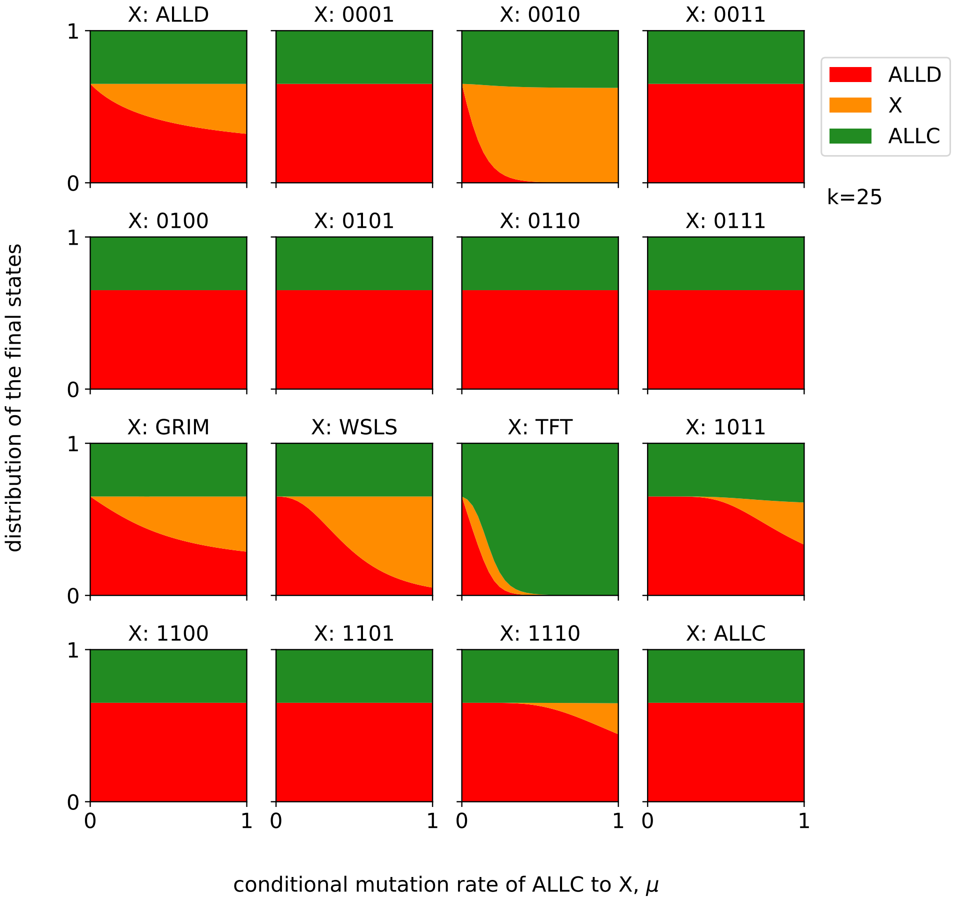
$$\frac{\mathrm{d}z}{\mathrm{d}t} = \varepsilon_2\left(c(1 - \varepsilon_2) - b\varepsilon_1\right) > 0 \tag{37}$$

Solutions also do not converge to the $x = 1$, the ALLC corner either. For $\mu > \mu_1$, all interior solutions converge to $\mathbf{x}_{\mathrm{E}}$ or $\mathbf{x}_{\mathrm{TFT}}$. $\square$

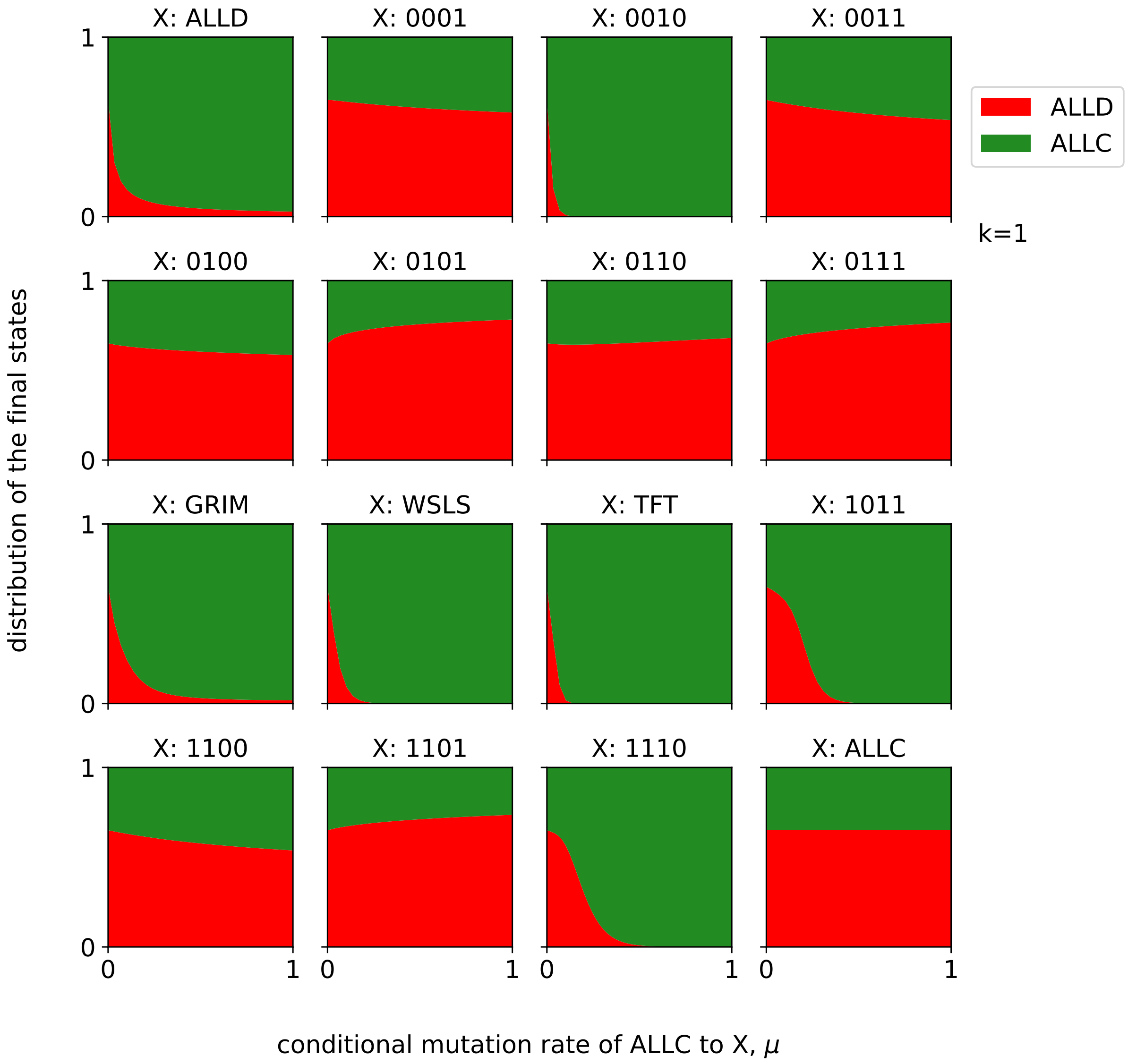
Supplementary Figures



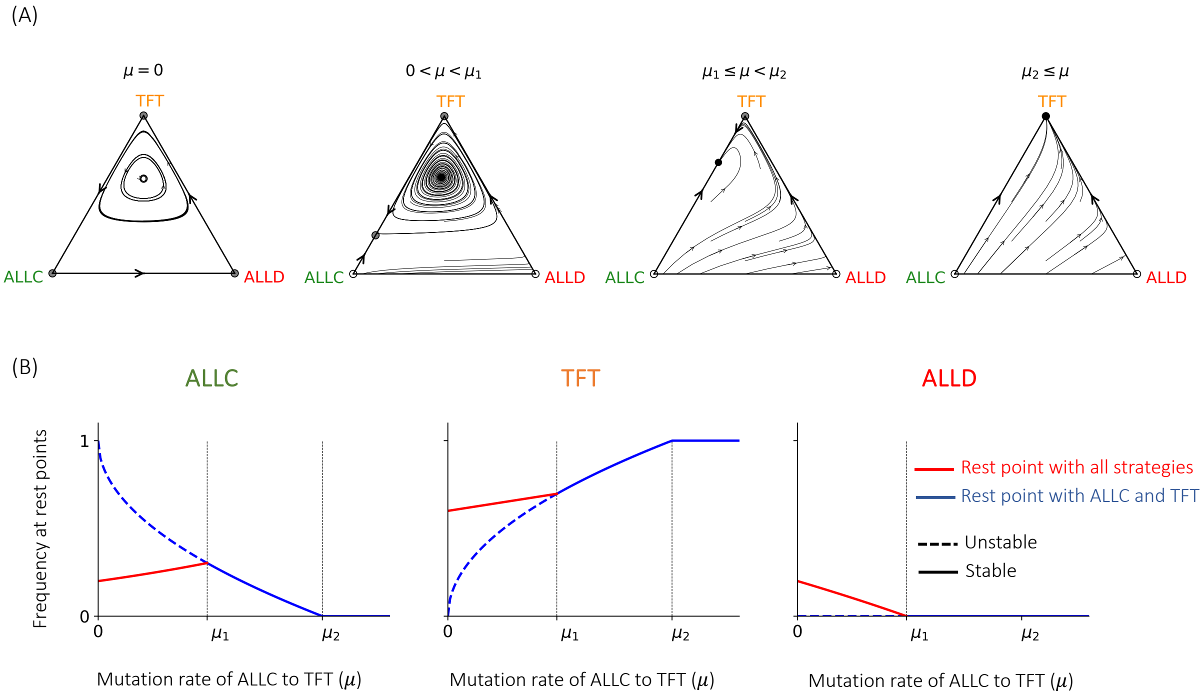
Supplementary Figure S1: Rescue outcomes in the finite population when tit-for-tat is the rescue strategy, as we vary the benefit-to-cost ratio. We plot the distribution of the final states versus μ (exactly as Figure 2 in main text) when tit-for-tat is the rescue strategy. We vary the benefit of cooperation b as we keep c fixed at 1. The population size is 100.



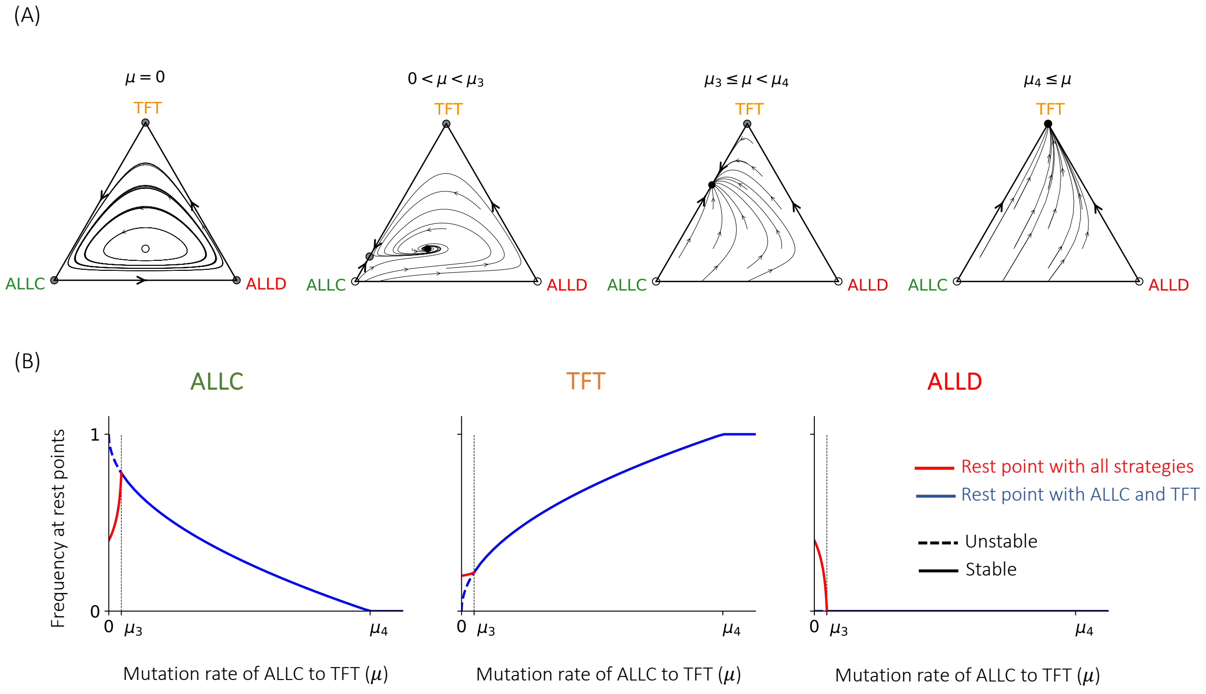
Supplementary Figure S2: Even when rescue is delayed ($k = 25$)**, tit-for-tat provides the best recue, among all deterministic memory-1 strategies**. We perform the exact computations as Figure 2 in main-text, with the exception that $k = 25$ instead of $k = 1$.



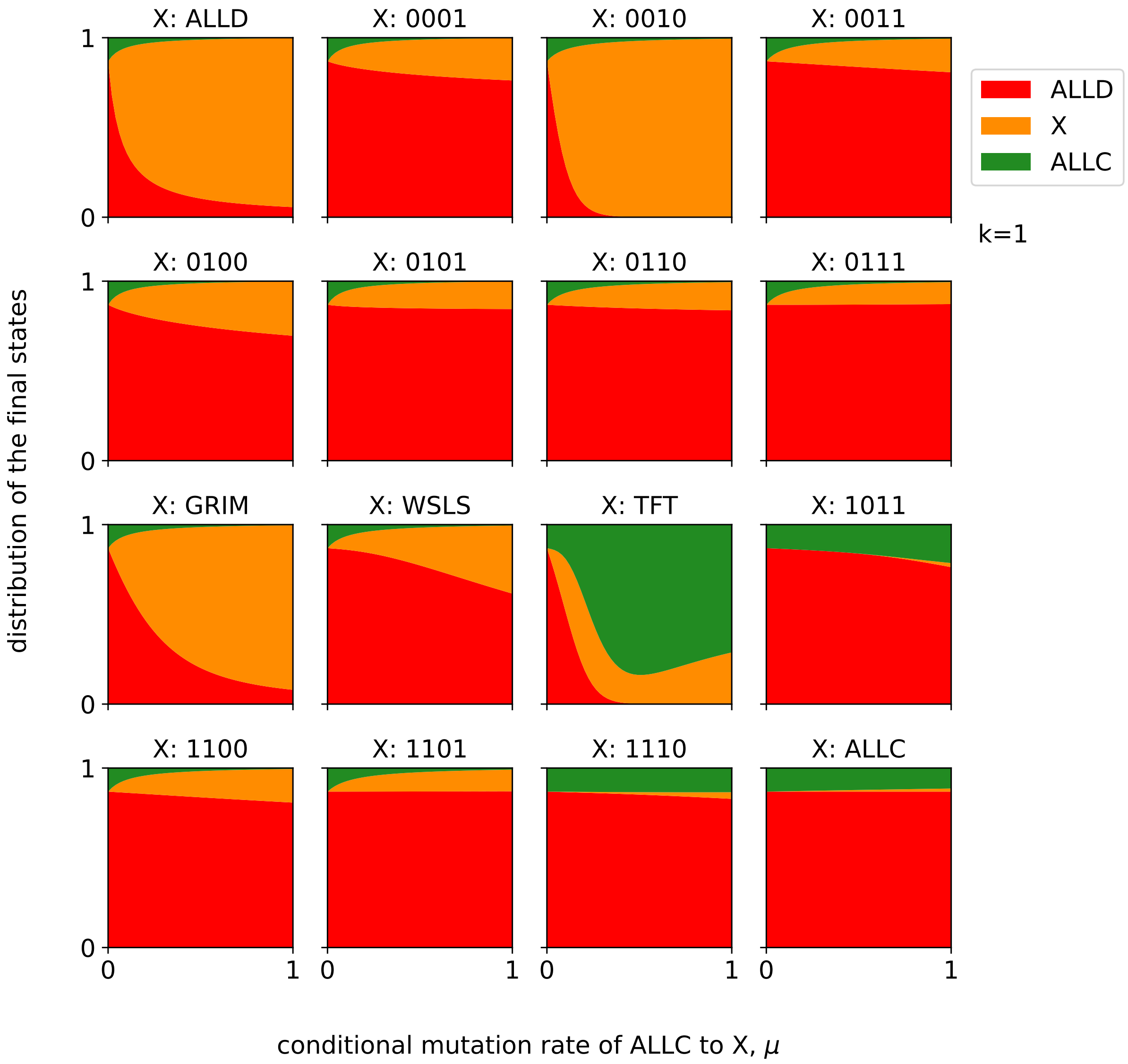
Supplementary Figure S3: Results from a variant of the finite population model. Here strategy X back-mutates to ALLC after ALLD is extinct. We consider an additional step to the baseline version of the finite population model. Here, after ALLD goes extinct, the strategy X mutates to ALLC with probability $\mu_{\text{back}}(>0)$. The process now has two final states: one in which ALLC takes over and the other in which ALLD takes over. Outcomes for ALLD are unaffected as this additional step does not interfere in ALLD's extinction (or fixation). For any value of $\mu_{\text{back}} > 0$, the probability that the process absorbs to the full-ALLC state can be computed by simply adding ρ_{ALLC} and ρ_{X} from the baseline process. Here we use the same parameters as Fig.2 from the main text. For small values of μ, the strategy S_2 provides comparable rescue to S_{10}, TFT.



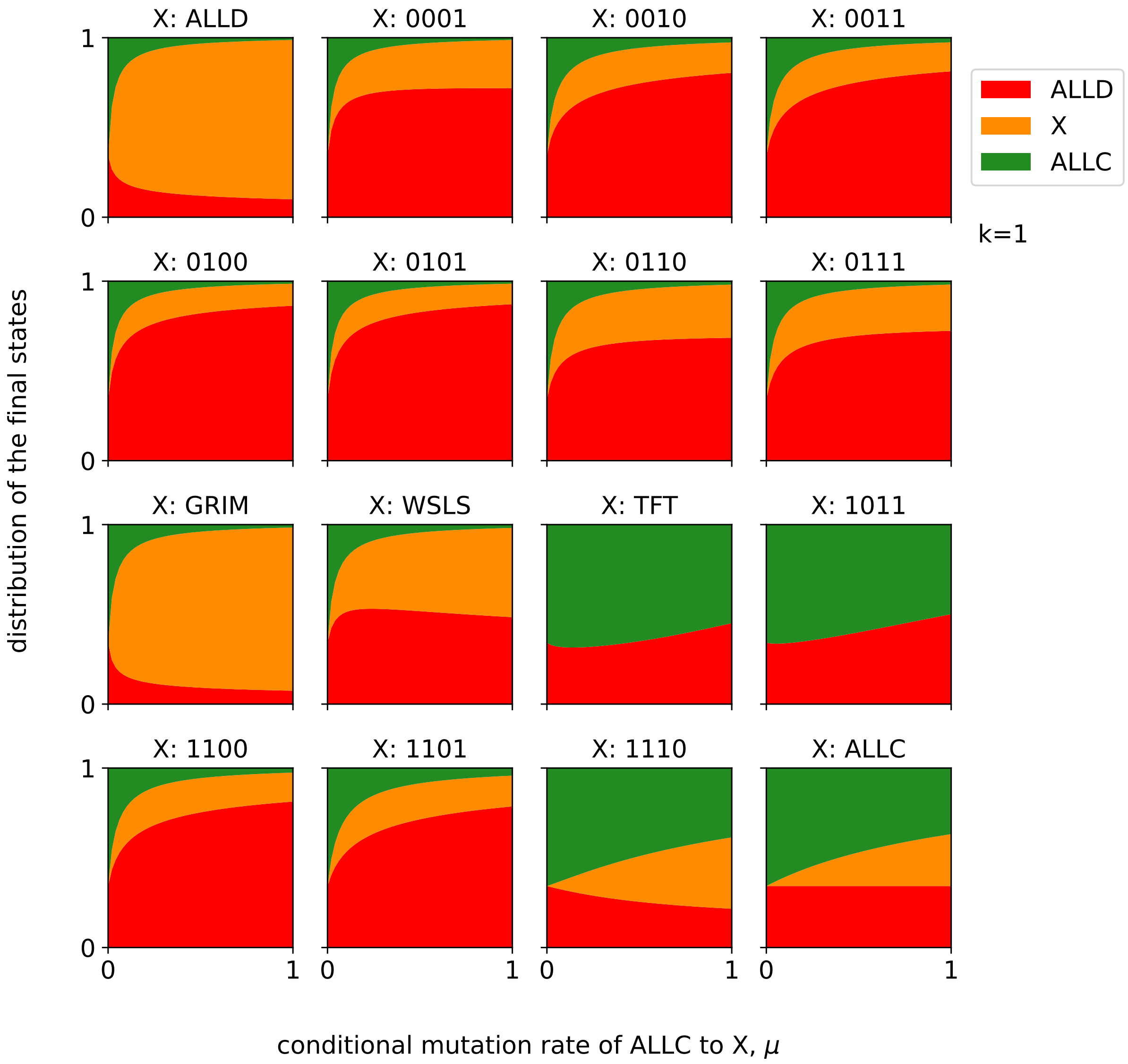
Supplementary Figure S4: Analysis of the Replicator-Mutator Equation I demonstrates an optimal $\mu\ (= \mu_1)$ **at which ALLC is stable in maximal frequency.** Here we recreate Fig. 5 from main text with the parameters $b = 5$ and $c = 3$. Upon comparing this Figure with Fig. 4 from the main text, we find that the maximal frequency of ALLC in a stable equilibrium is lower when the cost of cooperation is higher. The optimal mutation rate, $\mu = \mu_1$ is also higher compared to the $b = 5, c = 1$ case.



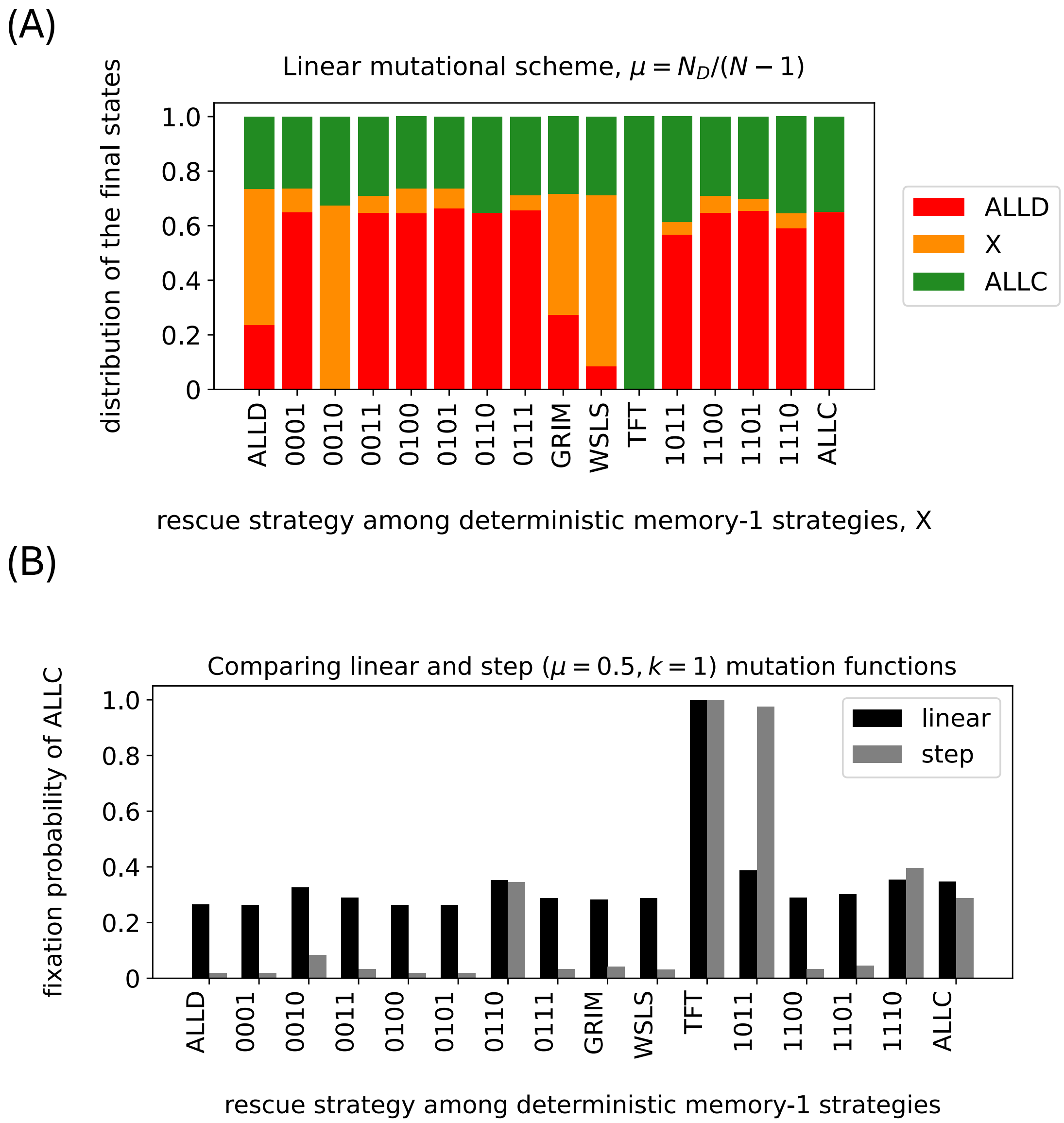
Supplementary Figure S5: Analysis of Replicator-Mutator Equation II demonstrates an optimal μ $(= \mu_3)$ at which ALLC is stable in maximal frequency. Here we recreate Fig. 5 from main text, but considering the dynamic (D2) with the parameters $b = 5$ and $c = 1$. Here $\mu_4 := (b - c)/(2b)$. We observe that the phase-portraits from this dynamic are qualitatively similar to that of dynamic of Replicator-Mutator I, (D1) as shown in Fig. 5 from main-text or Fig. S4 here.



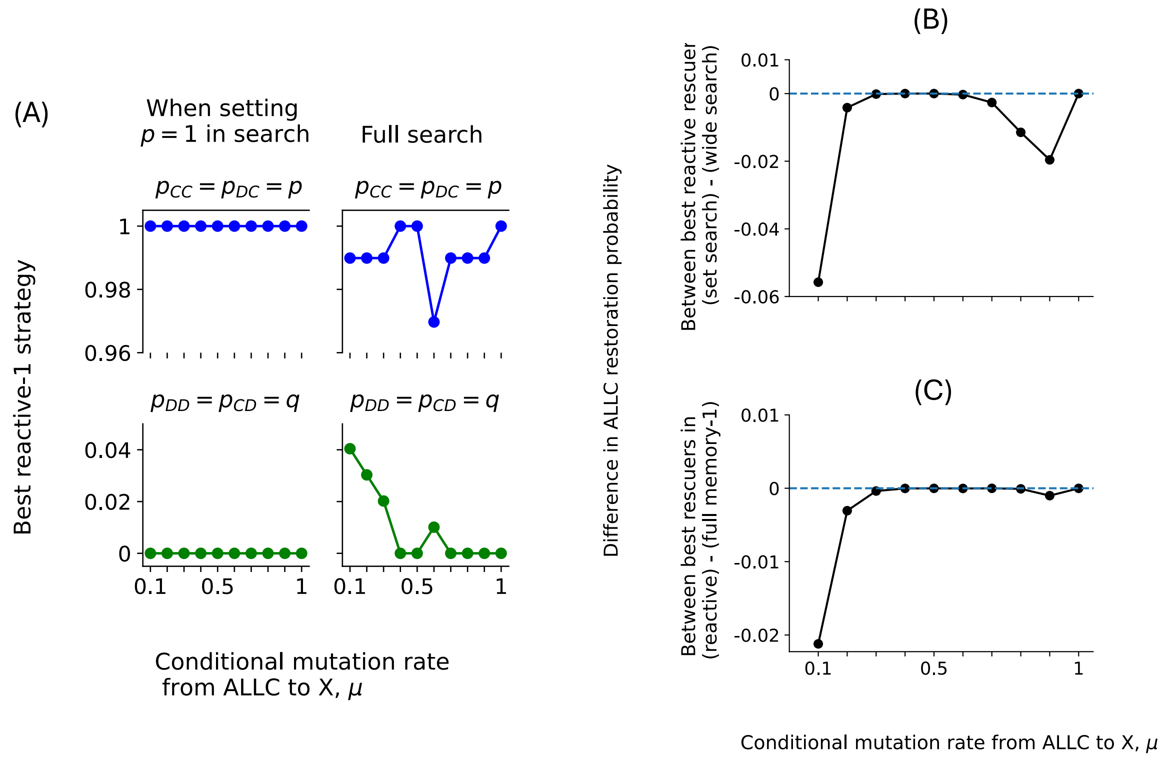
Supplementary Figure S6: Performance of the 16 deterministic memory-1 strategies as rescue strategies when the underlying game is the Prisoner's dilemma with $(R, S, T, P) = (3, 0, 5, 1)$**.** Here we recreate Fig. 2 from main text but with the underlying game as a Prisoner's dilemma that is not a donation game. The payoff for the first player at the outcomes CC, CD, DC and DD are R, S, T and P respectively. We use a population size of 100 and mutation threshold $k = 1$.



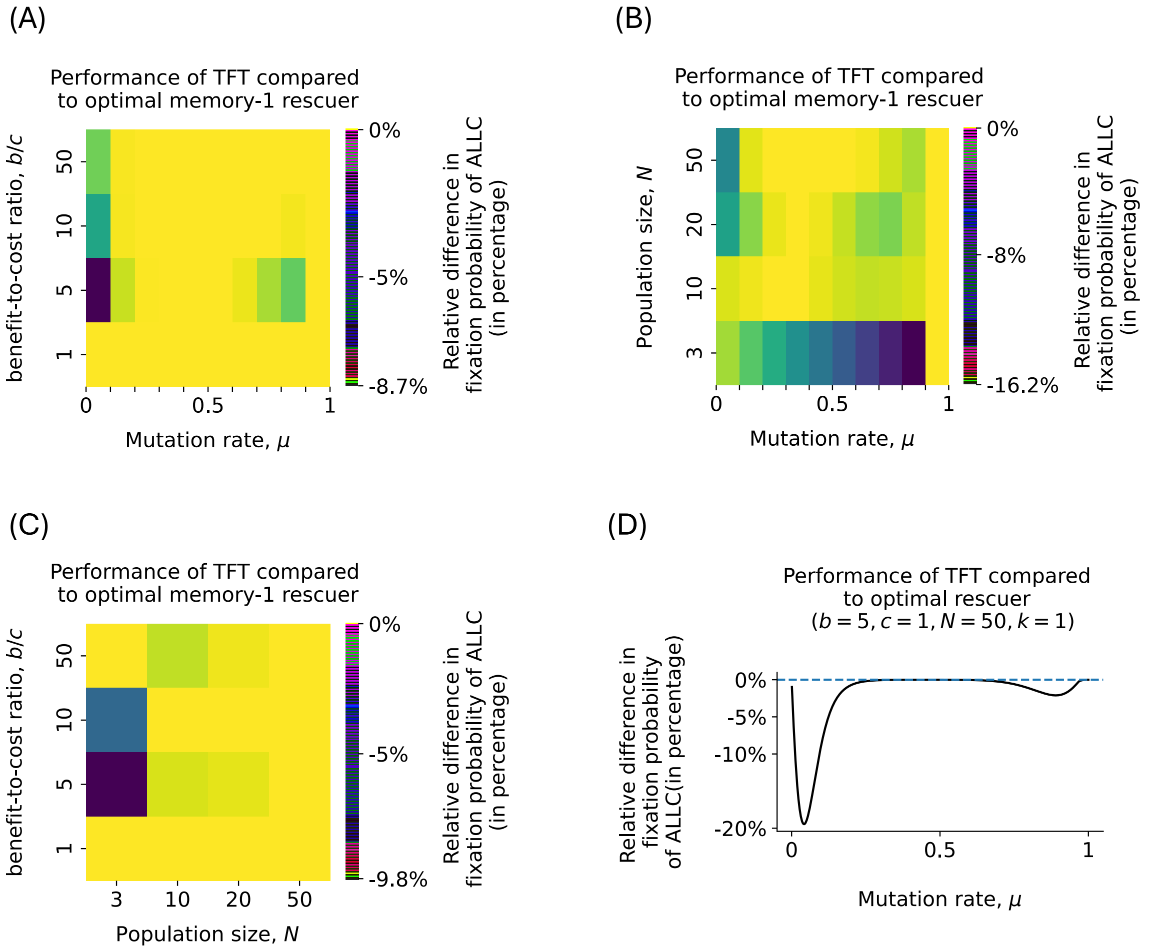
Supplementary Figure S7: Performance of the 16 deterministic memory-1 strategies as rescue strategies when the underlying game is the coordination game. We take $(R, S, T, P) = (5, -20, 4.9, 4)$ **as the stage game payoffs.** Here we recreate Fig. 2 from main text but with the underlying game as a coordination game. The payoff for the first player at the outcomes CC, CD, DC and DD are R, S, T and P respectively. We use a population size of 100 and mutation threshold $k = 1$. We observe that, as a rescuer, Tit-for-Tat offers only a marginal improvement in the restoration chances of ALLC — and in some cases, it even performs worse. In contrast, strategies like GRIM and even ALLD prove effective at eliminating the invading ALLD. This suggests that under a back-mutation mechanism, as shown in Supplementary Figure S3, these strategies would be highly successful in restoring the ground state.



Supplementary Figure S8: The performance of the 16 deterministic memory-1 strategies as rescue strategies when the mutational scheme is not a step function but a linear function on the number of ALLD in the population. We consider a linear mutation scheme where the probability of mutation from ALLC to X increases linearly to the number of ALLD in the population.We choose a linear function that has $\mu = 0$ when no ALLD is present and $\mu = 1$ when $N - 1$ out of N individuals are ALLD. The underlying game is a donation game with $b = 5$ and $c = 1$. We take a population of size 100. In panel A we show the probability of absorption into the final states for each deterministic memory-1 strategy as the rescue strategy. In panel B we compare the performance of the linear mutational function with the step function. We plot the probability that the final state is ALLC for each deterministic memory-1 rescue strategy, using black and grey bars to represent the results under a linear and a step mutation function, respectively. For the step function we choose $\mu = 0.5$ and a trigger threshold $k = 1$.



Supplementary Figure S9: An example where reactive rescue strategies that limit cooperation with ALLC outperform fully cooperative ones. Here we search for the best rescue strategy in the space of reactive strategies, similar to Fig 3A in the main text. We choose a population of size $N = 50$, a mutation threshold, $k = 1$. The benfit and cost of the underlying donation game are $b = 5$ and $c = 1$. (A) We make two independent searches for the best rescue strategy. In the first search, the best rescue strategy is numerically searched among reactive strategies that satisfy $p = 1$ (i.e., those that fully cooperate with ALLC in a direct competition). In the second search, the best rescue strategy is numerically searched in the entire space of reactive strategies $[0, 1]^2$. The results of the latter are identical to those in Fig 3A of main text. We observe that when we restrict the search of best rescue strategy to reactive strategies with $p = 1$, each time the best rescue strategy turns out to be tit-for-tat, i.e., has $q = 0$. However, when the search is performed in the entire space, the best rescue strategy does not necessarily have $p = 1$ (or $q = 0$). (B) We compare the performances of the best rescue strategies across these two independent searches. We find that the best strategies from the second search always perform at least as well as the best strategies from the first search. (C) Finally, we compare the performance of the best reactive rescue strategy and the best memory-1 rescue strategy that we obtain from the respective searchers. We find that the best memory-1 rescue strategies, which fully cooperate with ALLC and fully defect with ALLD perform atleast as good (sometimes better) than the best reactive rescue strategy.



Supplementary Figure S10: Comparing the rescue performance of tit-for-tat with the performance of the optimal rescue strategy. In panels A, B and C, we compare the perfromance of tit-for-tat as a rescue strategy to the performance of the optimal memory-1 rescue strategy, which we find from our search process. We do this comparison for a range of parameter values of the finite population model. In panel A, we vary the benefit-to-cost ratio of the donation game, b/c and the conditional mutation probability μ from ALLC to X as we hold the population size fixed to $N = 50$. In panel B, we vary the population size and the conditional mutation probability, as we fix the benefit-to-cost ratio to 5. Finally in panel C, we vary the benefit-to-cost ratio and population size as we fix the mutation rate to 0.5. To compare, we report the relative difference in the fixation probability of ALLC (in percentage), when TFT is the rescuer versus when the optimal rescue strategy is the rescuer. To compute the relative difference, we divide by the fixation probability of ALLC, when the optimal rescue strategy is the rescuer. In panel D, we similarly compare the performance of Tit-for-Tat with that of the optimal rescue strategy, identified without restriction to a specific strategy space. This analysis mirrors the approach used in Figure 4 of the main text. The optimal rescue strategy fully cooperates with ALLC, fully defects against ALLD, and achieves an arbitrary payoff r when playing against itself. Parameters for panel D: $N = 50$, $b/c = 5$ and mutation threshold, $k = 1$.

References

[1] Gerald Teschl. *Ordinary differential equations and dynamical systems*, volume 140. American Mathematical Society, 2024.

[2] Steven H Strogatz. Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering. 2018.