Overpunishing is not necessary to fix cooperation in voluntary public goods games

Fabio Dercole[1,2], Marco De Carli[1], Fabio Della Rossa[1], Alessandro V. Papadopoulos[1]

[1]DEI, Politecnico di Milano

[2]To whom correspondence should be addressed
Address: Via Ponzio 34/5, 20133 Milano, Italy
Ph: +39 02 2399 3484; Fax: +39 02 2399 3412
E-mail: fabio.dercole@polimi.it

Abstract

The fixation of cooperation among unrelated individuals is one of the fundamental problems in biology and social sciences. It is investigated by means of public goods games, the generalization of the prisoner's dilemma to more than two players. In compulsory public goods games, defect is the dominant strategy, while voluntary participation overcomes the social dilemma by allowing a cyclic coexistence of cooperators, defectors, and non-participants. Experimental and theoretical research has shown how the combination of voluntary participation and altruistic punishment—punishing antisocial behaviors at a personal cost—provides a solution to the problem, as long as antisocial punishment—the punishing of cooperators—is not allowed. Altruistic punishment can invade at low participation and pave the way to the fixation of cooperation. Specifically, defectors are overpunished, in the sense that their payoff is reduced by a sanction proportional to the number of punishers in the game. Here we show that qualitatively equivalent results can be achieved with a milder punishing mechanism, where defectors only risk a fixed penalty per round—as in many real situations—and the cost of punishment is shared among the punishers. The payoffs for the four strategies—cooperate, defect, abstain, and cooperate-&-punish—are derived and the corresponding replicator dynamics analyzed in full detail.

Key words: altruistic punishment, finite populations, fixation of cooperation, public goods games, replicator dynamics.

1 Introduction

How cooperation can emerge and be maintained among unrelated individuals in the presence of free-riders—defectors, exploiting the others' effort—is a long-standing puzzle (Dawes, 1980; Hardin, 1968). Public goods games (PGGs; Kagel and Roth, 1997), where participants equally share a public resource irrespectively of their individual contribution, represent the natural tool of investigation, both experimentally and in theoretical models.

Compulsory PGGs, in which individuals are obliged to participate (the natural generalization of the prisoner's dilemma to an arbitrary number of players; Boyd and Richerson, 1988; Hauert and Schuster, 1998), show that "always defect" is the dominant (evolutionarily stable) strategy (Dawes, 1980). When more successful strategies spread (through social learning or natural selection; Schuster and Sigmund, 1983), cooperation will disappear from the population, along with the loss of the public goods.

Several experimental and theoretical studies have identified mechanisms that are able to relax the social dilemma (see Nowak, 2006b, 2012, and refs. therein). When the chance to interact more than once with the same individuals is not vanishing (e.g. in finite populations), the use of memory to reciprocate the others' actions and/or of communications to build and spread reputations can maintain cooperation and, in some cases, allow the invasion of a uniform population of defectors. Similar results have been obtained with memoryless players in heterogeneous environments and/or with particular interaction structures (e.g. assortative grouping) favoring encounters within clusters of cooperators; or by relying on group or kin selection arguments. Most of the proposed mechanisms enhance the evolution of cooperation but fail to explain its fixation—the convergence to "always cooperate" as the dominant strategy.

Voluntary participation has been also shown to overcome the social dilemma, without requiring memory, communications, spatial, and/or interaction structures (Hauert et al., 2002a,b; Semmann et al., 2003). Again, cooperation cannot take over, but rather fluctuates in a rock-paper-scissors manner alternating cooperation, defection, and abstention. Non-participants (called "loners" by Hauert et al., 2002a,b) are supposed to rely on some alternative source for which a group activity is not required.

Cooperation is obviously fostered by the punishment of defectors (Boyd and Richerson, 1992; Fehr and Gächter, 2000). The punishment of antisocial behaviors is however costly and requires the identification of defectors (also required to reciprocate and build reputations). In some cases, there are institutional mechanisms imposing sanctions on defectors, so the cost is covered by the public goods, though experimental results (de Quervain et al., 2004; Egas and Riedl, 2008; Fehr and Fischbacher, 2003; Fehr and Gächter, 2002) have shown that people voluntarily engage in altruistic punishment—paying a personal cost to punish non-contributors.

Altruistic punishment in compulsory PGGs can lead to the fixation of cooperation (Henrich and Boyd, 2001), but the emergence of punishment in a uniform population of defectors remains problematic (and only explained by means of reciprocity and reputation, Sigmund et al., 2001, or group rather than individual selection, Boyd et al., 2010, 2003).

Interestingly, allowing optional participation and altruistic punishment gives a solution to the puzzle, as long as antisocial punishment (Herrmann et al., 2008)—non contributors paying a cost to punish contributors—is not allowed (the apparently illogical motivation and the effects of antisocial punishment will be addressed in the discussion). Fowler (2005) first proposed a voluntary PGG with the four strategies: cooperate, defect, abstain, and cooperate-&-punish. The model shows that punishers can invade and take over any initial antisocial state. It is however based on rather unrealistic assumptions, among which that the proportions of the four strategists in each round of the PGG are the same as in the whole population (or, equivalently, the whole population is involved in each round of the game). The model has been revised by Brandt et al. (2006), who extended the modeling assumption in Hauert et al. (2002a,b)—the PGG is played within a group of N individuals sampled from an infinite population at each round—to the case in which altruistic punishment is allowed. The new model showed a bistable behavior, the evolutionary (replicator) dynamics converging either to the fixation of cooperation (a mix of cooperators and punishers) or to a cooperators-defectors-loners rock-paper-scissors cycle in the absence of punishers. The authors concluded that "the emergence of altruistic punishment was still offering theoretical challenges".

The challenge was resolved by the same group of authors, by relying on the inherent stochasticity of finite populations (Hauert et al., 2007, 2008). They found "surprisingly different" results between the stochastic simulations, where each player (in a finite population) switches, from time to time, to the strategy of a better performing player, and the limiting deterministic (replicator) dynamics describing highly frequent sampling in an infinite population. The stochastic simulations do not show the bistable behavior and converge to the fixation of cooperation even starting from a nearly uniform population of defectors. The apparent contradiction can be explained in terms of the dynamic properties of the two attractors of the bistable dynamics and will be addressed in a dedicated section.

Although criticized in some of the basic assumptions (Boyd and Mathew, 2007; Mathew and Boyd, 2009; Rand and Nowak, 2011), already extended to more institutional and organized forms of punishment (Nakamaru and Dieckmann, 2009; Sasaki et al., 2012; Sigmund et al., 2010), and limited by the hypothesis of perfect knowledge of each player's behavior, the combination of voluntary participation and altruistic punishment remains the sole way to support the emergence and fixation of cooperation among memoryless players in the absence of communications, spatial, and interaction structures. However, all the models presented so far assume a rather strict form of punishment, hereafter referred as "overpunishment," where the payoff of each defector in the group of players involved in a round of the game is reduced by a sanction proportional to the number of punishers in the group. This occurs in experiments (Egas and Riedl, 2008; Fehr and Gächter, 2002) and models (Brandt et al., 2006; Hauert et al., 2007, 2008) of *peer-punishment*, where each punisher imposes a fixed penalty, the fine for having defected, onto each defector; as well as in the model of *pool-punishment* proposed by Sigmund et al. (2010) as a first step toward an institutionalized punishing mechanism, where, however, punishers contribute a fixed amount to the punishing pool. But in many real situations, free-riders only risk a fixed sanction that is independent of the number of punishers in the group—as when free-riding in a public bus—while punishers share the cost of punishment. In this paper, we analyze this latter case, that we name *shared-punishment*.

The new deterministic model is derived and fully analyzed. As a result, the new replicator dynamics are equivalent to the case with peer-punishment (Brandt et al., 2006), so the bistable behavior persists, though the proportion of initial conditions (the initial frequencies of the four strategies) reaching the two alternative regimes—the fixation of cooperation and the rock-paper-scissors cycle—are different. As expected, a milder punishment reduces the proportion of initial conditions leading to the fixation of cooperation. However, the dynamics being equivalent, the stochasticity of finite populations still allows the invasion and fixation of cooperation from any initial state. The conclusion is that overpunishment is not needed, and the same results can be obtained with a gentle punishing scheme.

The paper is organized as follows. First the average payoffs for the four strategies are derived (Sect. 2) and the resulting expressions compared with those with peer-punishment. Then, the corresponding replicator dynamics are analyzed (Sect. 3). We study in particular the equilibrium at which the populations is composed of loners only. This equilibrium is a nonlinear saddle (having all vanishing linear terms in the equations' expansion) lying on the boundary separating the basins of attractions of the two alternative regimes. Due to its nonlinear character, there are parameter settings for the PGG (including the basic setting used by Brandt et al. (2006)) for which the saddle attracts a set of initial states (with nonzero measure) in the four-dimensional simplex. Thus, it behaves as third possible attractor for the evolutionary orbits. By analyzing the orbits' behavior in the vicinity of the loners' equilibrium, and by assuming small and rare random fluctuations, we can correctly estimate the proportions of initial states reaching the two alternative regimes. Finally, the reasons why the fixation of cooperation is to be expected as the sole attractor of the stochastic simulations of a finite population are discussed (Sect. 4). Further discussion and conclusions close the paper (Sect. 5). For the sake of clarity, some of the technical steps are reported in Appendix and can be skipped by the uninterested reader.

2 The voluntary PGG with peer- and shared-punishment

In the voluntary PGG with altruistic punishment proposed by Brandt et al. (2006) there are four pure strategists, cooperators, defectors, loners, and punishers, whose relative densities (frequencies) in an infinite population are denoted by x, y, z, and w, respectively, $x+y+z+w=1$. Each round of the game is played by a group of $N > 2$ individuals randomly selected. Cooperators, defectors, and punishers participate in the public goods interaction, whereas loners opt for the alternative activity that provides a fixed payoff $p_z = \sigma > 0$. Cooperators and punishers contribute an amount c to the public goods. The total investment is then multiplied by a factor $r > 1$ ($r - 1$ being the investment return) and equally divided among all participants (equivalent results have been obtained for the case of "strictly altruistic" cooperation, where no benefit from the individual contribution returns to the contributor; De Silva et al., 2010).

Defectors therefore benefit from the public goods without contributing, but are identified and sanctioned by the punishers in the group, if there are any. With peer-punishment (Brandt et al., 2006), each punisher imposes a fine $\beta > 0$ onto each defector and, in the presence of defectors, a reduced fine $\alpha\beta$, $0 < \alpha < 1$, onto each cooperator (for not contributing to the punishing; non-punishing contributors are second-order free-riders) at personal costs γ and $\alpha\gamma$, respectively, $0 < \gamma < \beta$. With shared-punishment, we assume that, in the presence of defectors and punishers, each defector (cooperator) is sanctioned a fixed fine β ($\alpha\beta$) at a cost γ ($\alpha\gamma$), and that the total cost of the punishment is equally shared among the punishers.

Denoting by n_x, n_y, n_z, and n_w the numbers of cooperators, defectors, loners, and punishers in the group, $n_x + n_y + n_z + n_w = N$, and by $S = n_x + n_y + n_w$ the number of participants in the public goods, the payoffs for the participants are those reported in Table 1. Note that the joint activity requires, by definition, a group of participants to be remunerable, so that a single participant is forced to behave as a

$S>1$	p_x	p_y	p_w
$n_w=0$	$\frac{r\,c\,n_x}{S}-c$	$\frac{r\,c\,n_x}{S}$	–
$n_w>0$ and $n_y=0$	$\frac{r\,c\,(n_x+n_w)}{S}-c=(r-1)\,c$	–	$\frac{r\,c\,(n_x+n_w)}{S}-c=(r-1)\,c$
$n_w>0$ and $n_y>0$ with peer-punishment	$\frac{r\,c\,(n_x+n_w)}{S}-c-\alpha\beta\,n_w$	$\frac{r\,c\,(n_x+n_w)}{S}-\beta\,n_w$	$\frac{r\,c\,(n_x+n_w)}{S}-c-\alpha\gamma\,n_x-\gamma\,n_y$
$n_w>0$ and $n_y>0$ with shared-punishment	$\frac{r\,c\,(n_x+n_w)}{S}-c-\alpha\beta$	$\frac{r\,c\,(n_x+n_w)}{S}-\beta$	$\frac{r\,c\,(n_x+n_w)}{S}-c-\frac{\alpha\gamma\,n_x+\gamma\,n_y}{n_w}$
$S=1$	σ	σ	σ

Table 1: PGG payoffs.

loner (case $S=1$).

Other constraints to the game are:

(i) $\sigma<(r-1)c$, as loners must get less than contributors when everybody is contributing;

(ii) $r<N$, which guarantees that defection dominates cooperation in the absence of loners and punishers (the gain for a single cooperator switching to defection is $c(1-r/N)$ per round);

We also constrain the numerical values of the game parameters to the following ranges:

(iii) $N\in(2,100]$, $r=(2,50]$, $c=1$, $\sigma=(0,49)$, $\beta=(0,50]$, $\gamma=(0,50)$, $\alpha=(0,1)$.

When unable to prove statements analytically, we test them against the feasible parameter settings on a fine grid over the above ranges.

Note that the contribution c is kept fixed in (iii). In fact, by measuring payoffs (and therefore also σ, β, and γ) in units of c, we can always set $c=1$ (this corresponds to replace σ, β, and γ with $\tilde{\sigma}\,c$, $\tilde{\beta}c$, and $\tilde{\gamma}c$, where $\tilde{\sigma}$, $\tilde{\beta}$, and $\tilde{\gamma}$ are the new scaled parameters, factor c in front of all payoffs, see Table 1, and divide all payoffs by c). However, in the following we prefer to leave c indicated as a parameter of the PGG.

Given the frequencies (x,y,z,w) of the strategies in the population, the average payoff for each strategy is computed (following Hauert et al., 2002a) by summing the payoff values corresponding to all possible compositions for the group of N players, each weighted by the probability of sampling the composition from the population. With peer-punishment, the average payoffs (in the form presented in Hauert et al.,

2008) are

$$
\begin{aligned}
P_x &= z^{N-1}\sigma + rc(x+w)B(z) - cF(z) - w(N-1)G(y)\alpha\beta, && \text{(1a)}\\
P_y &= z^{N-1}\sigma + rc(x+w)B(z) \qquad\qquad - w(N-1)\beta, && \text{(1b)}\\
P_z &= \sigma, && \text{(1c)}\\
P_w &= \underbrace{z^{N-1}\sigma}_{\text{loner activity}} + \underbrace{rc(x+w)B(z) - cF(z)}_{\text{public goods interaction}} - \underbrace{\left[x(N-1)G(y)\alpha\gamma + y(N-1)\gamma\right]}_{\text{punishment}}, && \text{(1d)}
\end{aligned}
$$

with

$$B(z) := \frac{1}{1-z}\left(1 - \frac{1-z^N}{N(1-z)}\right) \tag{2a}$$

determining the average benefit returned by the public goods,

$$F(z) := 1 + (r-1)\,z^{N-1} - \frac{r}{N}\frac{1-z^N}{1-z} \tag{2b}$$

measuring the effective cost of contributing, and

$$G(y) := 1 - (1-y)^{N-2} \tag{2c}$$

taking into account that cooperators are punished only in the presence of defectors in the group.

With shared-punishment, the average payoffs (computed in Appendix A1) only differ in the punishment terms, that result in

$$
\begin{aligned}
P_x &= \cdots - G_x(y,w)\alpha\beta, && \text{(3a)}\\
P_y &= \cdots - G_y(w)\beta, && \text{(3b)}\\
P_w &= \cdots - \left[\frac{x}{w}G_x(y,w)\alpha\gamma + \frac{y}{w}G_y(w)\gamma\right] && \text{(3c)}
\end{aligned}
$$

with

$$
\begin{aligned}
G_x(y,w) &:= 1-(1-y)^{N-1} - (1-w)^{N-1} + (1-y-w)^{N-1} \quad \text{and} && \text{(4a)}\\
G_y(w) &:= 1-(1-w)^{N-1} && \text{(4b)}
\end{aligned}
$$

respectively measuring the probability for cooperators and defectors to be punished.

Comparing the punishment terms in Eqs. (1) and (3) is straightforward. In Eqs. (1a,b) the punishment, when imposed, is proportional to the average number of punishers in the group ($w(N-1)$) and the individual cost of punishing in Eq. (1d) is proportional to the average number of individual to be punished ($x(N-1)$ cooperators and $y(N-1)$ defectors). With shared-punishment, Eq. (3), the punishment, when imposed, is simply $\alpha\beta$ for cooperators and β for defectors, while the w at denominator in Eq. (3d) indicates that the total cost of punishment is equally shared among the punishers.

3 The corresponding replicator dynamics

The replicator equation $\dot{x} = x(P_x - \bar{P})$, being $\bar{P} = xP_x + yP_y + zP_z + wP_w$ the average payoff in the population, and similarly for the other strategies, describes the strategies' evolution through social learning or natural selection (Schuster and Sigmund, 1983) in the four-dimensional simplex $x + y + z + w = 1$.

The (numerically generated) replicator dynamics on the boundary faces of the simplex are portrayed in Fig. 1. Panels a and b respectively show the dynamics with peer- and shared-punishment for the basic parameter setting used by Brandt et al. (2006). Obviously, the punishing scheme makes a difference only when both defectors and punishers are present, i.e. on the faces $z = 0$ and $x = 0$ (the triangles x-y-w and y-z-w; the other two triangles are therefore not reported in panel a). However, the differences are only quantitative, i.e., the dynamics with peer- and shared-punishment are equivalent on the boundary faces.

In the following, we briefly review and integrate the analysis of the four faces separately, as previously discussed in Hauert et al. (2002a,b) with no punishment and in Brandt et al. (2006) and Hauert et al. (2008) with peer-punishment. In particular, we show that different scenarios are possible (both with peer- and shared-punishment for suitable parameter settings) on the faces $z = 0$ (Fig. 1c) and $x = 0$ (Fig. 1d,e; scenarios in panel e are only possible with peer-punishment).

3.1 The face $w = 0$

The voluntary PGG with no punishment is analyzed in detail in Hauert et al. (2002a). If $r > 2$ there is a unique interior equilibrium M surrounded by infinitely many (neutrally stable) periodic orbits (cooperators-defectors-loners rock-paper-scissors cycles; see Fig. 1b). If $r \leq 2$ (not shown in the figure) there are no interior equilibria and all orbits come from and converge to the loners' equilibrium $z = 1$ (so-called

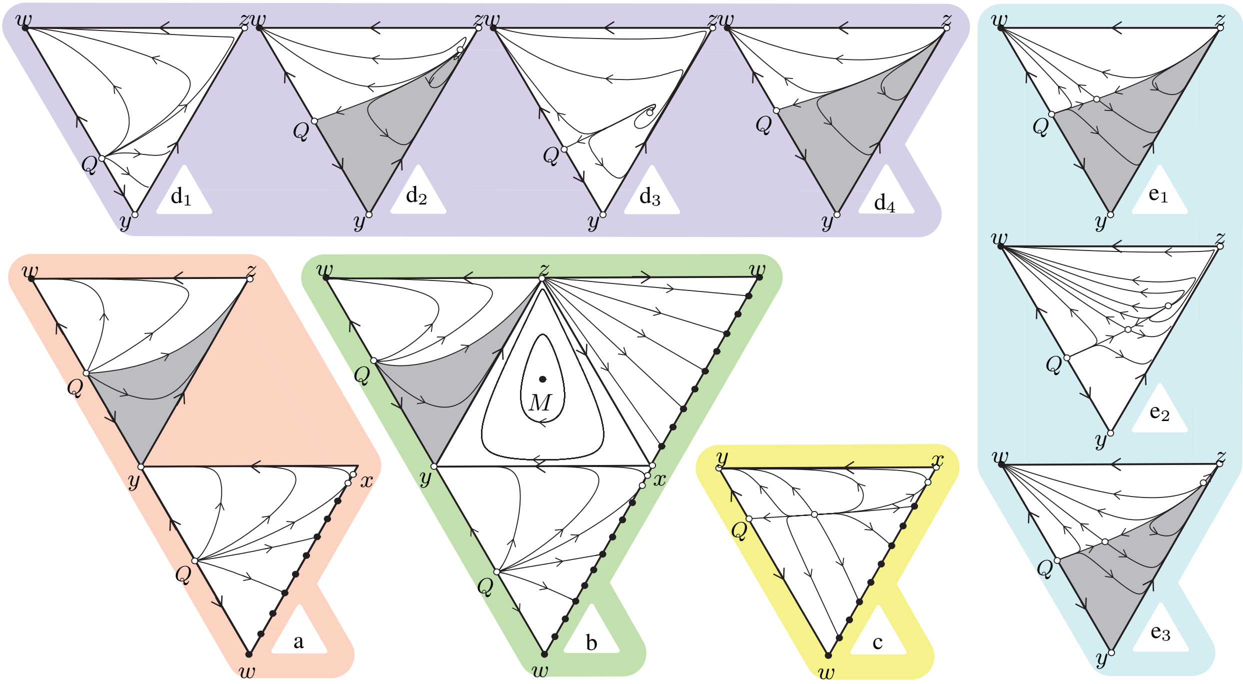

Figure 1: Panels a and b: replicator dynamics on the boundary faces of the simplex with (a) and without (b) overpunishment. Parameter values: $N = 5$, $r = 3$, $c = 1$, $\sigma = 1$, $\beta = 1.2$, $\gamma = 1$, $\alpha = 0.1$ (the basic setting used by Brandt et al. (2006)). Panel c: alternative scenario for the face $z = 0$, obtained with overpunishment. Parameter perturbations from the basic setting: $\gamma = 0.3$, $\alpha = 0.8$. Panels d,e: alternative scenarios for the face $x = 0$, obtained with overpunishment. Parameter perturbations from the basic setting: $N = 4$, $\sigma = 0.1$,

	d$_1$	d$_2$	d$_3$	d$_4$	e$_1$	e$_2$	e$_3$
β, γ	1.2, 0.4	0.6, 0.55	0.6, 0.2	0.68, 0.67	0.8, 0.75	0.8, 0.4	0.8, 0.6

Equivalent dynamics for cases c and d are obtained without overpunishment for the same parameter values.

homoclinic orbits to $z = 1$).

3.2 The face $y = 0$

In the absence of defectors, cooperators and punishers behave the same and have higher payoff than loners. Their frequencies therefore increase by maintaining the initial ratio up to fixation ($x + w = 1$) (as shown in the x-z-w triangle of Fig. 1b). The x-w edge is made of infinitely many (neutrally stable) equilibria.

3.3 The face $z = 0$

The x-w edge is (obviously) stationary also in the absence of loners, but the stability of equilibria is different. An equilibrium $(1 - w, 0, 0, w)$ is (neutrally) stable if defectors cannot invade, i.e., if $P_y < P_x = P_w$. This yields

$$w > \frac{N-r}{N(N-1)}\frac{c}{\beta} := w_{\min} < 1 \qquad \text{if } \beta > \frac{N-r}{N(N-1)}c := \beta_{\min}, \text{ never otherwise} \tag{5a}$$

with peer-punishment and

$$w > 1-\left(1-\frac{N-r}{N}\frac{c}{\beta}\right)^{\frac{1}{N-1}} := w_{\min} < 1 \text{ if } \beta > \frac{N-r}{N}c := \beta_{\min}, \qquad \text{never otherwise} \tag{5b}$$

with shared-punishment. There are stable equilibria close enough ($w > w_{\min}$) to the punishers' vertex $w = 1$ only if the sanction β is larger than the threshold $\beta_{\min}$.

Note that $\beta_{\min}$ is smaller with peer-punishment (recall that $N > 2$) and that $w_{\min}$ in (5a) is smaller than in (5b) (proved in Appendix A2), so that peer-punishment favors the stability of x-w equilibria. Also note that assuming $\beta > c$ (as indeed done by Brandt et al. (2006)) implies $\beta > \beta_{\min}$, so there would always be stable equilibria on the x-w edge. With peer-punishment, however, the punishment for a defector can exceed c even for $\beta < c$, so parameter settings with $\beta < c$ might also be interesting to explore (whereas $\beta > c$ should be required with shared-punishment). Since our interest is in the fixation of cooperation, in the following we assume $\beta > \beta_{\min}$.

On the y-w edge, both strategies cannot be invaded. At $y = 1$, $P_w < P_y$ is guaranteed by the PGG assumptions (see Appendix A6 for the evaluation of P_w at $w = 0$). At $w = 1$, $P_y < P_w$ requires the condition on β in (5). Under this assumption, imposing $P_y = P_w$ at the generic point $(0, 1 - w, 0, w)$, one gets a unique solution corresponding to the equilibrium Q in Fig. 1. With peer-punishment, an analytic expression is available, $y_Q := (\beta - \beta_{\min})/(\beta + \gamma)$, $w_Q := 1 - y_Q = (\beta_{\min} + \gamma)/(\beta + \gamma)$, whereas y_Q can be determined only numerically with shared-punishment (the proof of the uniqueness is reported in Appendix A3). Similarly to what happens on the x-w edge, peer-punishment yields a lower w_Q (numerically verified for the feasible parameter settings in the ranges indicated in note (iii), Sect. 2). Thus, as expected, overpunishment increases the chances for cooperation to go to fixation, in the sense that the fraction of initial conditions on the face $z = 0$ leading to the x-w edge is larger (compare the faces $z = 0$ in Fig. 1a,b).

The stability of equilibrium Q is also studied in Appendix A3. If α is sufficiently small, then $P_x > P_w$ at Q (cooperators are not punished at $\alpha = 0$), so cooperators can invade, hence Q repels also transversally to the y-w edge (Fig. 1a,b). All orbits in the interior of the x-y-w triangle come from point Q and converge either to $y = 1$ or to one of the stable equilibria on the x-w edge. The orbit separating the two cases is tangent to the x-w edge at $w = w_{\min}$. In contrast, $P_x < P_w$ at Q when γ is sufficiently small (punishers punish at no cost at $\gamma = 0$), so cooperators cannot invade and Q is a saddle (Fig. 1c). In this case an interior repellor is also present and one of the orbits emanating from it is tangent to the x-w edge at $w = w_{\min}$ and separates (together with the stable manifold of Q) the initial conditions reaching $y = 1$ from those leading to the x-w edge. The two situations are separated by a so-called transcritical bifurcation (Kuznetsov, 2004; Meijer et al., 2009) for the corresponding replicator equation.

3.4 The face $x = 0$

The interior orbits come from point Q also in the y-z-w triangle in Fig. 1a,b. The only attractor here is the equilibrium $w = 1$. However, it does not attracts all interior initial conditions, as some (shaded in the figure) are attracted by the saddle equilibrium $z = 1$. This peculiar behavior is made possible by the nonhyperbolicity of the saddle (having all vanishing linear terms in the equations' expansion) and is proved in Appendix A4. A particular orbit emanating from Q and converging to $z = 1$ separates the basins of attraction of the two alternative regimes. The loners' equilibrium, though unstable, thus behaves as an attractor (in the Milnor's sense—nonzero measure of the basin of attraction).

For different parameter settings other scenarios are possible. Those reported in panels d_1–d_4 can occur with peer- as well as shared-punishment, whereas scenarios e_1–e_3 are only possible with peer-punishment (numerically verified, see note (iii) in Sect. 2). The transitions among the various scenarios involve:

- the transcritical bifurcation at which equilibrium Q changes stability (see Appendix A3); transitions from panel a to d_2 and d_4–e_1;
- the bifurcations of the nonlinear saddle (see Appendix A4); transitions a–d_1 and d_2–d_4;
- a so-called saddle-node bifurcation (Kuznetsov, 2004; Meijer et al., 2009) at which two internal equilibria (a saddle and a repellor) appear/disappear; transitions a–e_3 and d_1–e_2 (the bifurcation has been characterized numerically and has not been found with shared-punishment).

In all cases, $w = 1$ remains the only proper attractor, with $z = 1$ behaving as such in some of the cases (where its basin of attraction is shaded in Fig. 1).

3.5 The simplex interior

Depending on the initial condition in the simplex interior, the replicator dynamics converge to one of three possible regimes: one of the stable equilibria on the x-w edge, with the consequent fixation of cooperation; one of the stable cycles on the face $w = 0$; the loners' equilibrium $z = 1$ (numerically verified with a mix of simulation and continuation—Dhooge et al., 2002—techniques for the feasible parameter settings in the ranges indicated in note (iii), Sect. 2). The latter possibility went unnoticed in Brandt et al. (2006) and Hauert et al. (2008) and is again due to the nonlinear character of the saddle equilibrium $z = 1$ (having all vanishing linear terms).

Similarly to what happens on the face $x = 0$, there are parameter settings for which $z = 1$ is attracting a set of initial conditions with nonzero measure in the simplex interior. This attractive behavior cannot be proved by means of the techniques used in Appendix A4 (effective only when the dynamics are two-dimensional), but is confirmed by the numerical simulations. In fact, there are parameter settings (including the case of Fig. 1a,b) for which the simulations starting from a significant fraction of equally spaced initial conditions in the simplex interior seem to converge to $z = 1$. Although, numerically, there is no way to know whether $z = 1$ is attracting the orbit or behaving like an hyperbolic saddle—there are orbits spending arbitrarily long times close to $z = 1$ in both cases—there are also parameter settings for which none of our simulations converge to $z = 1$.

There can be therefore three attractors for the replicator dynamics in the simplex interior: A, the set of stable equilibria on the x-w edge; B, the interior of the face $w = 0$; C, the loners' equilibrium $z = 1$. Fig. 2 shows the (numeriaclly estimated) basins of attractions of the three attractors for the basic parameter setting used in Fig. 1a,b. The general message is that cooperation goes to fixation if punishers are initially enough, but punishers cannot invade a uniform population of defectors. The required initial frequency of punishers is however vanishing close to the loners' equilibrium and this is the key feature for the fixation of cooperation when allowing small stochastic fluctuations in the frequencies of the four strategies (see Sect. 4).

Stochastic fluctuations indeed play a relevant role close to the loners' equilibrium. The orbits attracted by $z = 1$, and even those passing sufficiently close to it when $z = 1$ is behaving like an hyperbolic saddle, spend so much time there, that the deterministic description of the evolutionary dynamics becomes questionable. Allowing stochastic fluctuations in the four frequencies after a long time spent close to $z = 1$ means that, from time to time on a slower time scale, tiny fractions of the population make a trial in the

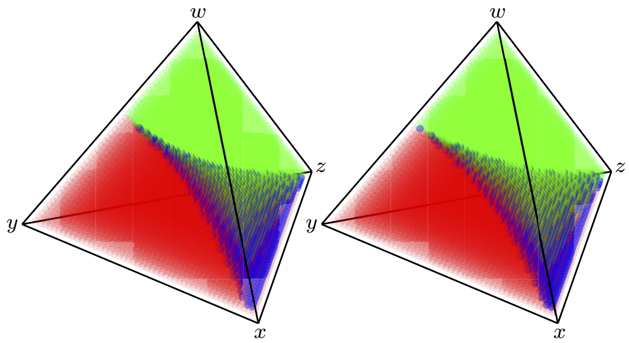

Figure 2: Replicator dynamics in the simplex interior with peer- (left) and shared- (right) punishment: green, red, and blue initial conditions lead to attractor A, B, and C, respectively (see Appendix A5 for the details on the classification algorithm). Parameter values as in Fig. 1a,b.

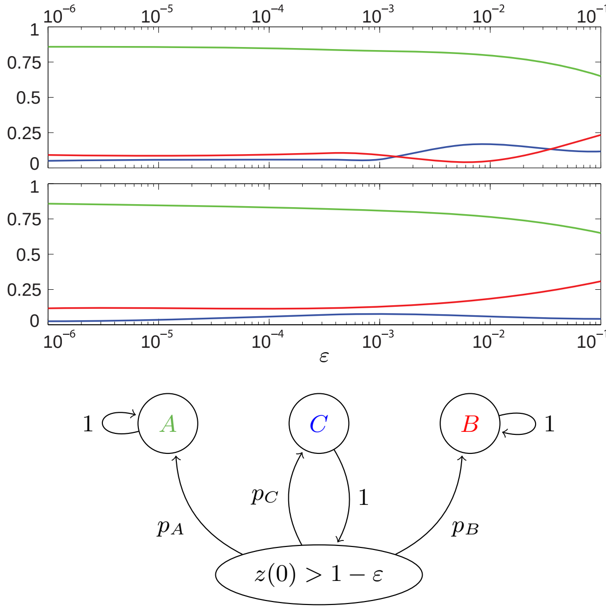

Figure 3: Fractions of initial conditions with $z > 1 - \varepsilon$ leading to attractors A, B, and C with peer- (top) and shared- (bottom) punishment (see Appendix A5 for the details on the classification algorithm). Parameter values as in Fig. 1a,b. The limiting ($\varepsilon \to 0$) fractions p_A, p_B, and $p_C = 1 - p_A - p_B$ can be used as probabilities to reach A, B, and C randomly starting close to $z = 1$.

PGG. Since the three outcomes A, B, and C can all be obtained when starting from an initial condition close to $z = 1$, as shown in Fig. 3, stochastic fluctuations close to $z = 1$ make A and B the only two ultimate attractors for the evolutionary dynamics, according to the Markov process in the figure (nodes A and B are the only absorbing states of the Markov process).

Fig. 4 reports the analysis with respect to the model parameters of the fractions of (equally spaced) initial conditions converging to the three possible attractors A, B, and C (see Appendix A5 for the details on the classification algorithm). Green, red, and blue dashed lines show the fractions associated to attractors

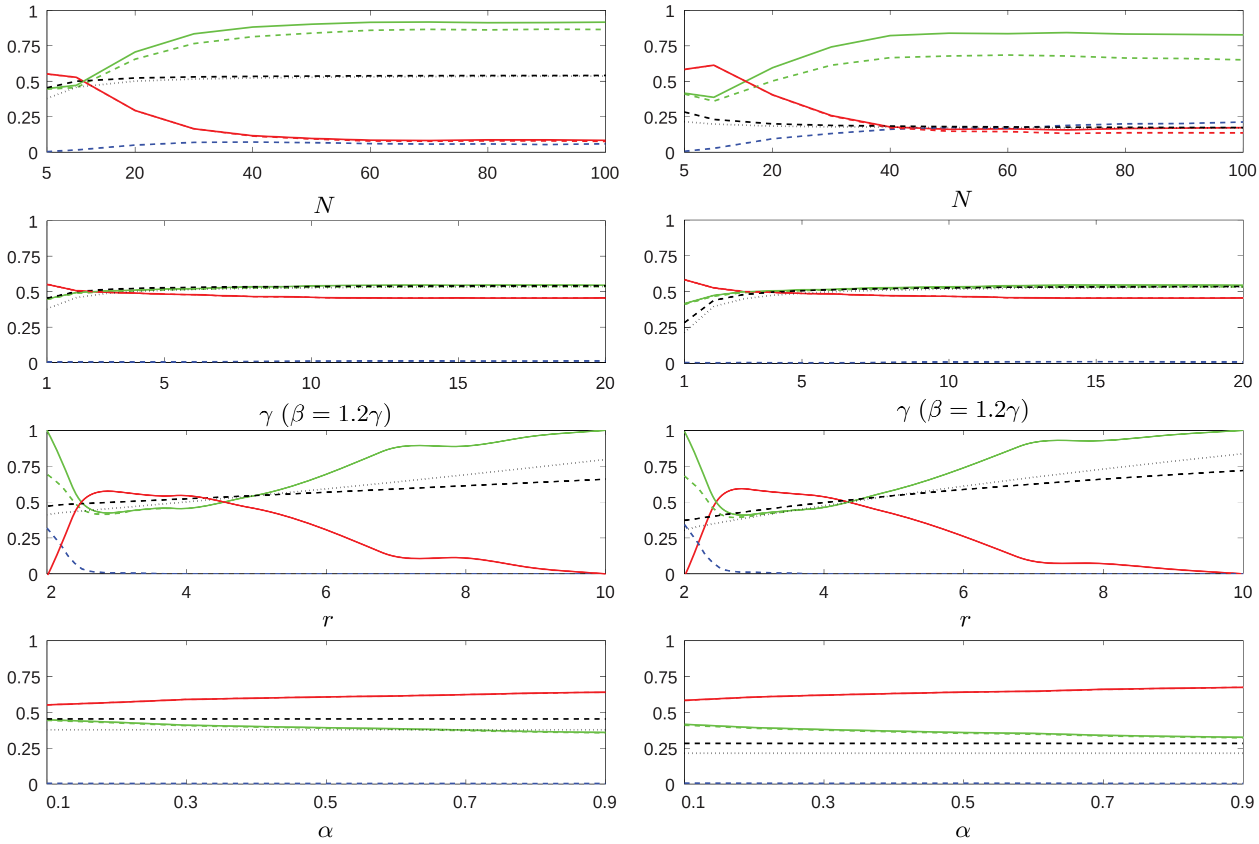

Figure 4: Parametric analysis of the fractions of initial conditions leading to attractors A, B, and C with peer- (left) and shared- (right) punishment. Green, red, and blue dashed lines: fractions leading to attractor A, B, and C, respectively; green and red solid lines: corrected fractions leading to attractor A and B; black dashed: y_Q; black dotted: area of the triangle Q-($w = w_{\min}$)-($w = 1$) within the face $z = 0$. Other parameters as in Fig. 1a,b.

A, B, and C. The solid green and red lines show the corrected fractions associated to attractors A and B when taking stochastic fluctuations close to $z = 1$ into account (the C fraction, multiplied by $p_A/(1-p_C)$ is added to the A fraction, and similarly for the B fraction, to count any possible visit to C before ultimately reaching A or B, see Fig. 3). The black dashed line reports the frequency y_Q, used by Brandt et al. (2006) as a "rule of thumb" for the fraction associated to A. From the dynamics on the face $z = 0$ (Fig. 1a–c), one can indeed see that the larger is y_Q, the larger is the set of initial conditions converging to the x-w edge. The same can be said for the frequency $1 - w_{\min}$ (see (5)) and what matters even more is the area of the triangle Q-($w = w_{\min}$)-($w = 1$) (black dotted line in Fig. 4).

Fig. 4 confirms that the frequencies y_Q and the area of the triangle Q-($w = w_{\min}$)-($w = 1$) both correlate with the fraction of initial states leading to the fixation of cooperation. It also confirms that peer-punishment always yields a larger fraction associated to attractor A.

4 The fixation of cooperation

We now discuss why it is licit to expect the invasion and the eventual dominance of altruistic punishers, when considering the stochastic effects of a finite population. In the stochastic simulations (Hauert et al., 2007, 2008), each player updates his/her own strategy by imitating a player who is selected, from time to time, with a probability proportional to the player's performance in the PGG (in another, basically equivalent, scheme used by Sigmund et al., 2010, a focal and a target player are randomly chosen and the focal player adopts the strategy of the target player with a probability that increases with the difference in the expected payoff; see e.g. Nowak, 2006a for further discussion on social learning). In addition, each player has a small probability to blindly change to another strategy at random ("mutation"). This process yields the deterministic replicator dynamics in the limit of frequent sampling in the absence of mutations in a large population. Roughly speaking, one can think of the deterministic orbits as the expected path followed by the stochastic process in the limit of vanishing variance.

Stochastic fluctuations within or transversal to an attractor of the deterministic dynamics may have remarkably different effects. Transversal fluctuations, e.g. moving the system's state from a stable equilibrium on the x-w edge slightly to the simplex interior, are reabsorbed due to the attractiveness, while those within the attractor, e.g. moving the state to a nearby stable equilibrium on the x-w edge, are not and cause a neutral drift in the system's state. The presence of infinitely many (neutrally) stable equilibria and cycles, as those on the x-w edge and on the face $w = 0$ (Fig. 1a–c), is therefore a source of neutral drift that cannot be neglected even at a very small fluctuations' variance.

As a consequence, the punishers' equilibrium $w = 1$ becomes the global attractor of the evolutionary dynamics. In fact, once the deterministic orbit reaches attractor B, the drift among the infinity of cycles guarantees that, sooner or later, the system's state passes close to the loners' equilibrium $z = 1$. There, punishers can invade and lead the state to attractor A. Again, the drift among the infinity of equilibria must be considered, so that the state eventually reaches either $w = 1$ or $w = w_{\min}$, the other extreme of the segment of stable equilibria. At $w = 1$, only a mutation can change the system's state, but the further imitation dynamics bring back the states on the x-w edge, at (or close to) $w = 1$. At $w = w_{\min}$ a transient can be triggered by the invasion of defectors, possibly converging again to attractor B, but eventually punisher will come to dominate again. And even starting from a uniform population of defectors, the same kind of dynamics is triggered by the invasion of loners.

Note that the system's state can spend a significant time at (or close to) any of the simplex vertexes. There, the population is nearly uniform and, if mutations are rare, the evolutionary dynamics are slow. However, while all vertexes but $w = 1$ are unstable, so the dynamics point away from them, the punishers' equilibrium is stable, though not attracting, so the system's state remains in the vicinity. Only the neutral drift along the x-w edge can then move the state significantly away from $w = 1$, but, being random, it often brings the state back to $w = 1$.

In conclusion, as confirmed by the stochastic simulations in Hauert et al. (2007) and Hauert et al. (2008), after the initial transient the system's state remains at (or close to) the punishers' equilibrium for most of the time, and only cooperators invading by neutral drift can break the punishers' dominance. Most of the time, the drift is reabsorbed, but occasionally (when the drift reaches $w = w_{\min}$) it triggers a transient with large oscillations in the strategies' frequencies (temporary passing close to attractor B), that eventually reestablishes another phase of punishers' dominance.

5 Discussion and conclusions

In this study we relaxed the peer-punishment scheme adopted in the voluntary PGG modeled by Brandt et al. (2006) (and in most of the related experiments; de Quervain et al., 2004; Egas and Riedl, 2008; Fehr and Fischbacher, 2003; Fehr and Gächter, 2002). Instead of each punisher in the group of interacting players imposing a sanction onto each defector, each defector is sanctioned a fixed fine and the total cost of punishing is shared among the punishers. We named such a punishing scheme shared-punishment and shown it works just as well, qualitatively, as peer-punishment. In particular, when allowing for the stochastic effects of a finite population (Hauert et al., 2007), both peer- and shared-punishment support the emergence and fixation of cooperation.

Sanctioning defectors (and non-punishing contributors as second-order free-riders) in proportion to the number of punishers among the PGG participants "overpunishes" antisocial behaviors with respect to many real situations. However, such overpunishing has been used in all experimental and theoretical studies, so far, including the first attempt to model the emergence of institutional forms of sanctioning (Sigmund et al., 2010) and the recent investigations of antisocial punishment (non-contributors punishing contributors; see Rand and Nowak, 2011, and ref. therein). While the idea of sharing the cost of punishing is already adopted in the pool-punishment of Sigmund et al. (2010) (in this case it is the cost to the punisher, and not the sanction, that is fixed), our shared-punishment is the first punishing scheme where free-riders only risk a fixed sanction per PGG round, independently of the number of encountered punishers. It can be interpreted as a rudimentary forms of sanctioning institution, where the punishing pool is formed a posteriori among the punishers involved in the public goods interaction.

Shared-punishment is a form of altruistic punishment and, as such, it can be challenged by antisocial punishment (Herrmann et al., 2008). As recently shown by Rand and Nowak (2011), allowing a generalized form (all-to-all) of (over-) punishment, thus including antisocial punishment, annihilates the positive effect of punishment in promoting the evolution of cooperation. In particular, punishing the others' strategy self-protects players from others' invasion (just as altruistic punishers are protected from the invasion of defectors), so that drifting from a uniform population of self-protecting cooperators (resp. defectors, loners) to a uniform population of the same strategy of non-punishers restores the rock-paper-scissors oscillations of the voluntary PGG with no punishment (Hauert et al., 2002a,b). Although rarely observed in western human societies (if not in the form of retaliation; Herrmann et al., 2008), antisocial punishment questions any punishing mechanism aimed at fostering cooperation and should be therefore considered and possibly controlled. This opens further important directions for future research.

We explored our shared-punishment scheme for a wide range of parameters (see note (iii) in Sect. 2). The fractions of initial conditions (the positive initial frequencies of the four strategies cooperate, defect, abstain, and cooperate-&-punish) converging to each of the two ultimate attractors of the deterministic evolutionary dynamics—the fixation of cooperation at a mix of punishing and non-punishing cooperators (A) and the rock-paper-scissors oscillations with no punishers (B)—are reported Fig. 4. As qualitatively discussed in Brandt et al. (2006), the fixation of cooperation (the green fraction) is enhanced by enlarging the group of interacting players (parameter N), by exacerbating the sanction (and proportionally its cost; parameters β and γ), and by increasing the reward (parameter r).

We also considered the effect of punishing cooperators as second-order free-riders (parameter α), an issue that recently received some attention. As discussed in (De Silva et al., 2010; Fowler, 2005; Hauert et al., 2007, 2008), the role of punishing second-order free-riders is surprisingly marginal in voluntary PGG, and this is confirmed by our analysis which shows it even (slightly) opposes the fixation of cooperation. However, second-order punishment gains back importance in the light of the pool-punishment model of Sigmund et al. (2010). There, peer- and pool-punishers compete with non-punishing cooperators, defectors, and loners. Pool-punishers a priori contribute to a punishing pool, a public goods later used to punish first- and second-order free-riders. This facilitate the sanctioning of second-order free-riders—the players not contributing to the punishing pool who later cooperate but do not punish defectors—and, as a result, pool-punishers invade and come to dominate only in the presence of second-order punishment. The natural emergence of institutional forms of sanctioning, not imposed by a higher authority, seems to require second-order punishment.

We conclude by going back to the idea of a fixed sanction. We see as interesting the extension in this direction of the current experimental setups and models of peer- and pool- punishing, possibly including antisocial or all-to-all punishment, with the aim of confirming that avoiding overpunishing yields qualitatively similar results.

Acknowledgments

We thank K. Sigmund for a discussion of an early draft of this manuscript and the two anonymous reviewers for the useful suggestions. This work was supported by the Italian Ministry for University and Research (under contract FIRB RBFR08TIA4).

References

Andronov, A. A., Leontovich, E. A., Gordon, I. J., Maier, A. G., 1973. Theory of Bifurcations of Dynamical Systems on a Plane. Wiley, New York, (translation from Russian).

Berezovskaya, F. S., Novozhilov, A. S., Karev, G. P., 2007. Population models with singular equilibrium. Math. Biosci. 208, 270–299.

Boyd, R., Gintis, H., Bowles, S., 2010. Coordinated punishment of defectors sustains cooperation and can proliferate when rare. Science 328, 617–620.

Boyd, R., Gintis, H., Bowles, S., Richerson, P. J., 2003. The evolution of altruistic punishment. Proc. Natl. Acad. Sci. 100, 3531–3535.

Boyd, R., Mathew, S., 2007. A narrow road to cooperation. Science 316, 1858–1859.

Boyd, R., Richerson, P. J., 1988. The evolution of reciprocity in sizeable groups. J. Theor. Biol. 132, 337–356.

Boyd, R., Richerson, P. J., 1992. Punishment allows the evolution of cooperation (or anything else) in sizable groups. Ethol. Sociobiol. 13, 171–195.

Brandt, H., Hauert, C., Sigmund, K., 2006. Punishing and abstaining for public goods. Proc. Natl. Acad. Sci. 103, 495–497.

Dawes, R. M., 1980. Social dilemmas. Annu. Rev. Psychol. 31, 169–193.

de Quervain, D. J.-F., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., Fehr, E., 2004. The neural basis of altruistic punishment. Science 305, 1254–1258.

De Silva, H., Hauert, C., Traulsen, A., Sigmund, K., 2010. Freedom, enforcement, and the social dilemma of strong altruism. J. Evol. Econ. 20, 203–217.

Dhooge, A., Govaerts, W., Kuznetsov, Yu. A., 2002. MATCONT: A MATLAB package for numerical bifurcation analysis of ODEs. ACM T. Math. Software 29, 141–164.

Egas, M., Riedl, A., 2008. The economics of altruistic punishment and the maintenance of cooperation. Proc. Roy. Soc. Lond. B 275, 871–878.

Fehr, E., Fischbacher, U., 2003. The nature of human altruism. Nature 425, 785–791.

Fehr, E., Gächter, S., 2000. Cooperation and punishment in public goods experiments. Am. Econ. Rev. 90, 980–994.

Fehr, E., Gächter, S., 2002. Altruistic punishment in humans. Nature 415, 137–140.

Fowler, J. H., 2005. Altruistic punishment and the origin of cooperation. Proc. Natl. Acad. Sci. 102, 7047–7049.

Hardin, G., 1968. The tragedy of the commons. Science 162, 1243–1248.

Hauert, C., De Monte, S., Hofbauer, J., Sigmund, K., 2002a. Replicator dynamics for optional public good games. J. Theor. Biol. 218, 187–194.

Hauert, C., De Monte, S., Hofbauer, J., Sigmund, K., 2002b. Volunteering as Red Queen mechanism for cooperation in public goods games. Science 296, 1129–1132.

Hauert, C., Schuster, P., 1998. Extending the iterated prisoners dilemma without synchrony. J. Theor. Biol. 192, 155–166.

Hauert, C., Traulsen, A., Brandt, H., Nowak, M. A., Sigmund, K., 2007. Via freedom to coercion: The emergence of costly punishment. Science 316, 1905–1907.

Hauert, C., Traulsen, A., Brandt, H., Nowak, M. A., Sigmund, K., 2008. Public goods with punishment and abstaining in finite and infinite populations. Biol. Theory 3, 114–122.

Henrich, J., Boyd, R., 2001. Why people punish defectors: Weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas. J. Theor. Biol. 208, 79–89.

Herrmann, B., Thoni, C., Gächter, S., 2008. Antisocial punishment across societies. Science 319, 1362–1367.

Kagel, J., Roth, A. (Eds.), 1997. The Handbook of Experimental Economics. Princeton University Press, Princeton, NJ.

Kuznetsov, Yu. A., 2004. Elements of Applied Bifurcation Theory, 3rd Edition. Springer-Verlag, Berlin.

Mathew, S., Boyd, R., 2009. When does optional participation allow the evolution of cooperation? Proc. Roy. Soc. Lond. B 276, 1167–1174.

Meijer, H. G. E., Dercole, F., Oldeman, B. E., 2009. Numerical bifurcation analysis. In: Meyers, R. A. (Ed.), Encyclopedia of Complexity and System Science. Springer-Verlag, Berlin, pp. 6329–6352.

Nakamaru, M., Dieckmann, U., 2009. Runaway selection for cooperation and strict-and-severe punishment. J. Theor. Biol. 257, 1–8.

Nowak, M. A., 2006a. Evolutionary Dynamics: Exploring the Equations of Life. Harvard University Press, Cambridge, MA.

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

Nowak, M. A., 2012. Evolving cooperation. J. Theor. Biol. 299, 1–8.

Rand, D. G., Nowak, M. A., 2011. The evolution of antisocial punishment in optional public goods games. Nat. Commun. 2, 434.

Sasaki, T., Brännström, Å., Dieckmanna, U., Sigmund, K., 2012. The take-it-or-leave-it option allows small penalties to overcome social dilemmas. Proc. Natl. Acad. Sci. 109, 1165–1169.

Schuster, P., Sigmund, K., 1983. Replicator dynamics. J. Theor. Biol. 100, 533–538.

Semmann, D., Krambeck, H. J., Milinski, M., 2003. Volunteering leads to rock-paper-scissors dynamics in a public goods game. Nature 425.

Sigmund, K., De Silva, H., Traulsen, A., Hauert, C., 2010. Social learning promotes institutions for governing the commons. Nature 466, 861–863.

Sigmund, K., Hauert, C., Nowak, M. A., 2001. Reward and punishment. Proc. Natl. Acad. Sci. 98, 10757–10762.

group composition	symbol	probability
N players with S participant	$P(N,S)$	$\binom{N}{S}(1-z)^S z^{N-S}$
N players with S participant and n_x cooperators	$P(N,S,n_x)$	$P(N,S)\binom{S}{n_x}\left(\frac{x}{1-z}\right)^{n_x}\left(\frac{y+w}{1-z}\right)^{S-n_x}$
N players with S participant, n_x cooperators, and n_y defectors (and $n_w = S - n_x - n_y$ punishers)	$P(N,S,n_x,n_y)$	$P(N,S,n_x)\binom{S-n_x}{n_y}\left(\frac{y}{1-x-z}\right)^{n_y}\left(\frac{w}{1-x-z}\right)^{S-n_x-n_y}$

Table A1: Sampling probabilities. Sampling from an infinite population follows a binomial distribution. Multiple requirements on the group composition are resolved by means of conditional probabilities: $P(A\cap B\cap C) = P(A)\,P(B|A)\,P(C|A\cap B)$.

Appendix

A1 Computation of the average payoffs

In this appendix we compute the average payoffs P_x, P_y, and P_z for cooperators, defectors, and punishers (3a,b,d) in the case with shared-punishment. We formulate each average as the sum of the possible values—the obtained payoffs according to Table 1—weighted by the probabilities of occurrence of the corresponding compositions for the group of N players, see Table A1.

$$
\begin{aligned}
P_x = {} & \underbrace{\sigma P(N-1,0)}_{\text{no other participants}} \\
& + \underbrace{\sum_{S=2}^{N}\sum_{n_x=1}^{S}\left(\frac{r\,c\,n_x}{S}-c\right)P(N-1,S-1,n_x-1,S-n_x)}_{\text{no punishers}} \\
& + \underbrace{\sum_{S=2}^{N}\sum_{n_x=1}^{S-1}(r-1)c\,P(N-1,S-1,n_x-1,0)}_{\text{with punishers and no defectors}} \\
& + \underbrace{\sum_{S=3}^{N}\sum_{n_x=1}^{S-2}\sum_{n_w=1}^{S-n_x-1}\left(\frac{r\,c\,(n_x+n_w)}{S}-c-\alpha\beta\right)P(N-1,S-1,n_x-1,S-n_x-n_w)}_{\text{with punishers and defectors}}
\end{aligned}
\tag{A1a}
$$

$$
\begin{aligned}
P_y \; = \; & \underbrace{\sigma P(N-1,0)}_{\text{no other participants}} \\
& + \underbrace{\sum_{S=2}^{N} \sum_{n_y=1}^{S} \frac{r\,c\,(S-n_y)}{S} P(N-1,S-1,S-n_y,n_y-1)}_{\text{no punishers}} \\
& + \underbrace{\sum_{S=2}^{N} \sum_{n_y=1}^{S-1} \sum_{n_w=1}^{S-n_y} \left(\frac{r\,c\,(S-n_y)}{S} - \beta \right) P(N-1,S-1,S-n_y-n_w,n_y-1)}_{\text{with punishers}} \qquad \text{(A1b)}
\end{aligned}
$$

$$
\begin{aligned}
P_w \; = \; & \underbrace{\sigma P(N-1,0)}_{\text{no other participants}} \\
& + \underbrace{\sum_{S=2}^{N} \sum_{n_w=1}^{S} (r-1)c\,P(N-1,S-1,S-n_w,0)}_{\text{no defectors}} \\
& + \underbrace{\sum_{S=2}^{N} \sum_{n_w=1}^{S-1} \sum_{n_y=1}^{S-n_w} \left(\frac{r\,c(S-n_y)}{S} - c - \frac{\alpha\gamma\,(S-n_y-n_w)+\gamma\,n_y}{n_w} \right) P(N-1,S-1,S-n_y-n_w,n_y)}_{\text{with defectors}} \qquad \text{(A1c)}
\end{aligned}
$$

The expressions in (3) have then been obtained by resolving the sums with the Newton binomial and have been handled and checked with computer algebra.

A2 The threshold $w_{\min}$

In this appendix we prove that $w_{\min}$ is smaller with peer- (5a) than with shared- (5b) punishment, i.e., that

$$
\frac{N-r}{N(N-1)} \frac{c}{\beta} < 1 - \left(1 - \frac{N-r}{N} \frac{c}{\beta} \right)^{\frac{1}{N-1}} \qquad \text{(A2a)}
$$

for any feasible parameter setting ($N > 2$, $r < N$, and $\beta > \beta_{\min}$). Letting

$$
a = \frac{N-r}{N} \frac{c}{\beta} \qquad \text{and} \qquad b = \frac{1}{N-1},
$$

(A2a) can be rewritten as

$$ab < 1 - (1-a)^b\,, \tag{A2b}$$

with $a \in (0,1)$ (due to $\beta > \beta_{\min}$, see (5b)) and $b \in (0, \frac{1}{2}\,)$ treated as a real number.

Let us rewrite inequality (A2b) as

$$b > f_a(b) := \frac{\log(1-ab)}{\log(1-a)} \tag{A2c}$$

and consider $f_a(b)$ as a a-parametric family of functions of $b \in [0,1]$. Since $f_a(b)$ is nonnegative and concave in $b \in [0,1]$ for any $a \in (0,1)$,

$$\frac{\mathrm{d}^2 f_a}{\mathrm{d}b^2} = -\frac{a^2}{(1-ab)^2\log(1-a)} > 0,$$

and coincides with the lefthand side of inequality (A2c) at $b=0$ and $b=1$, then (A2c) certainly holds for any pair (a,b) with $a \in (0,1)$ and $b \in (0, \frac{1}{2}\,)$.

A3 The equilibrium point Q

A3.1 Uniqueness with shared-punishment

The equilibrium Q on the y-w edge of the simplex is determined by solving $P_y = P_w$ for y_Q at point $(0, y_Q, 0, 1-y_Q)$. This gives the same solutions $y_Q \in (0,1)$ than $Q(y_Q) := (P_w(y_Q) - P_y(y_Q))\,(1-y_Q) = 0$.

With shared-punishment, we have

$$Q(y_Q) = (\beta+\gamma)y_Q^N - \beta y_Q^{N-1} + \left(c\left(1-\frac{r}{N}\right) - (\beta+\gamma)\right)y_Q + \beta - c\left(1-\frac{r}{N}\right).$$

Note that $Q(0) = \beta - \beta_{\min} > 0$ under (5b), $Q(1) = 0$, and $Q'(1) = (N-1)\gamma + \beta_{\min} > 0$ by the PGG assumptions. Thus, there is certainly a solution $y_Q \in (0,1)$. The uniqueness follows from the fact that the curvature $Q''(y) = (N-1)y^{N-3}[N(\beta+\gamma)y - (N-2)\beta]$ cannot change sign twice for $y \in (0,1)$.

A3.2 Stability on the face $z = 0$, transversally to the y-w edge

With peer-punishment, $P_x > P_w$ at Q yields

$$\alpha < \frac{y_Q}{1 - y_Q - (1 - y_Q)^{N-1}} \frac{\gamma}{\beta}.$$

With shared-punishment, it gives

$$\alpha < \frac{y_Q}{1 - y_Q - \dfrac{(1 - y_Q)^N}{1 - y_Q^{N-1}}} \frac{\gamma}{\beta}.$$

Note that the righthand side is positive in both cases (being $y_Q \in (0, 1)$) and possibly larger than one when γ is sufficiently close to β. Thus, equilibrium Q is a repellor (unstable also transversally to the y-w edge) if γ and α are sufficiently large and small, respectively. Vice versa, Q is a saddle ($P_x < P_w$ at Q) if γ is sufficiently small compared to α.

A3.3 Stability on the face $x = 0$, transversally to the y-w edge

With peer-punishment, at Q we have

$$P_w = -(N-1)^2 \left(\beta_{\min} + \frac{\beta - c}{N-1} \right) \frac{\beta_{\min} + \gamma}{\beta + \gamma},$$

with $\beta_{\min} := (r-1)c/(N(N-1))$ given in (5b). Thus, if $\beta > c$, then $P_w < 0$ at Q and $P_z > P_w$ for any $\sigma > 0$, so that equilibrium Q is a repellor (unstable also transversally to the y-w edge). Note that if $\beta = \beta_{\min}$, then $P_w = (r-1)c > 0$ at Q (Q collides with the vertex $w = 1$), so for some $\beta_{\min} < \beta < c$ equilibrium Q is a saddle for sufficiently small σ.

With shared-punishment, P_w at Q can be positive also for $\beta > c$ (e.g. $P_w \simeq 0.19$ for the setting of Fig. 1a,b), so that equilibrium Q can be a saddle if σ is sufficiently small.

A4 Analysis of the loners' equilibrium

In this appendix we analyze the replicator dynamics in the interior of the face $x = 0$ in the vicinity of the loners' equilibrium $z = 1$. In particular, we show that $z = 1$ is a nonhyperbolic saddle, with vanishing Jacobian, able to attract a set of initial conditions with nonzero measure in the face $x = 0$.

The replicator equation on the face $x = 0$ is two-dimensional and we use variables y and w, being $z = 1 - y - w$ and $(y, w) = (0, 0)$ the loners' equilibrium $z = 1$. Expanding in (y, w) powers, the equation reads

$$\dot{y} = Y_2(y,w) + Y_3(y,w) + \cdots\,, \quad \dot{y} = W_2(y,w) + W_3(y,w) + \cdots\,, \tag{A3}$$

where there are no linear terms,

$$\begin{aligned} Y_2(y,w) &:= -(N-1)\sigma y^2 + (N-1)\left(\tfrac{1}{2}\,rc - \beta - \sigma\right) yw \quad \text{and} \\ W_2(y,w) &:= (N-1)\left(\tfrac{1}{2}\,(r-2)c - \gamma - \sigma\right) yw + (N-1)\left((r-1)c - \sigma\right) w^2 \end{aligned}$$

collect quadratic terms with peer- as well as shared-punishment, and $Y_k(y,w)$ and $W_k(y,w)$, $k > 2$, collect the higher order terms (up to order $(N+1)$, with differences between peer- and shared-punishment), and are such that $Y_k(0,w) = W_k(y,0) = 0$, being the (y,w)-axes (representing the y-z and z-w edges) invariant.

We use the blow-up transformation $u = w/y$ (Andronov et al., 1973; Berezovskaya et al., 2007) and study the system in the coordinates (y,u) for small $|y| \neq 0$, where the equilibrium $(y,w) = (0,0)$ has been stretched into the u-axis. With the time-scaling $\tau = yt$ (that inverts the direction of time for $y < 0$), the new equations are

$$\dot{y} = yY_2(1,u) + y^2 Y_3(1,u) + y^3 Y_4(1,u) + \cdots\,, \quad \dot{u} = U_3(1,u) + yU_4(1,u) + y^2 U_5(1,u) + \cdots\,, \tag{A4}$$

with $U(y,w) := y\,W(y,w) - w\,Y(y,w)$.

It is well known that all orbits of system (A3) asymptotically reaching $(y,w) = (0,0)$, forward or backward in time, do that along characteristic directions $w = \bar{u}y$ corresponding to the equilibria $(0,\bar{u})$ of the blown-up dynamics (A4) plus, in our case, the invariant direction $y = 0$. From (A4) it follows that $\bar{u}$ must be a root of the third-order polynomial $U_3(1,u) = W_2(1,u) - uY_2(1,u)$. The associated eigenvalues are $\lambda_u(\bar{u}) := \mathrm{d}U_3(1,u)/\mathrm{d}u|_{u=\bar{u}}$ along the eigenvector $y = 0$ and $\lambda_y(\bar{u}) := Y_2(1,\bar{u})$ along $v_y(\bar{u}) := (Y_2(1,\bar{u}) - \lambda_u(\bar{u}), W_3(1,\bar{u}) - \bar{u}Y_3(1,\bar{u}))$. If such equilibria are all hyperbolic ($\lambda_u(\bar{u})\lambda_y(\bar{u}) \neq 0$), then the dynamics (A4) for small $|y| \neq 0$ are dominated by the linear terms and this induces a partition of a sufficiently small neighborhood of $(y,w) = (0,0)$ into sectors (called Brouwer sectors) with qualitatively known dynamics. We now apply these arguments to system (A3).

The polynomial $U_3(1,u)$ has degree two ($Y_2(1,u)$ is linear in u and $W_2(1,u)$ is quadratic), with roots

root of $U_3(1,u)$	λ_u	λ_y
$\bar{u}_1 = 0$	$-\frac{1}{2}(N-1)(2\gamma-(r-2)c)$	$-(N-1)\sigma$
$\bar{u}_2 = \dfrac{2\gamma-(r-2)c}{2\beta+(r-2)c}$	$\frac{1}{2}(N-1)(2\gamma-(r-2)c)$	$-\frac{1}{2}(N-1)\dfrac{c^2r(r-2)-2c((r-2)\beta+r\gamma)+4(\beta\gamma+\beta\sigma+\gamma\sigma)}{2\beta+(r-2)c}$

Table A2: Roots of $U_3(1,u)$.

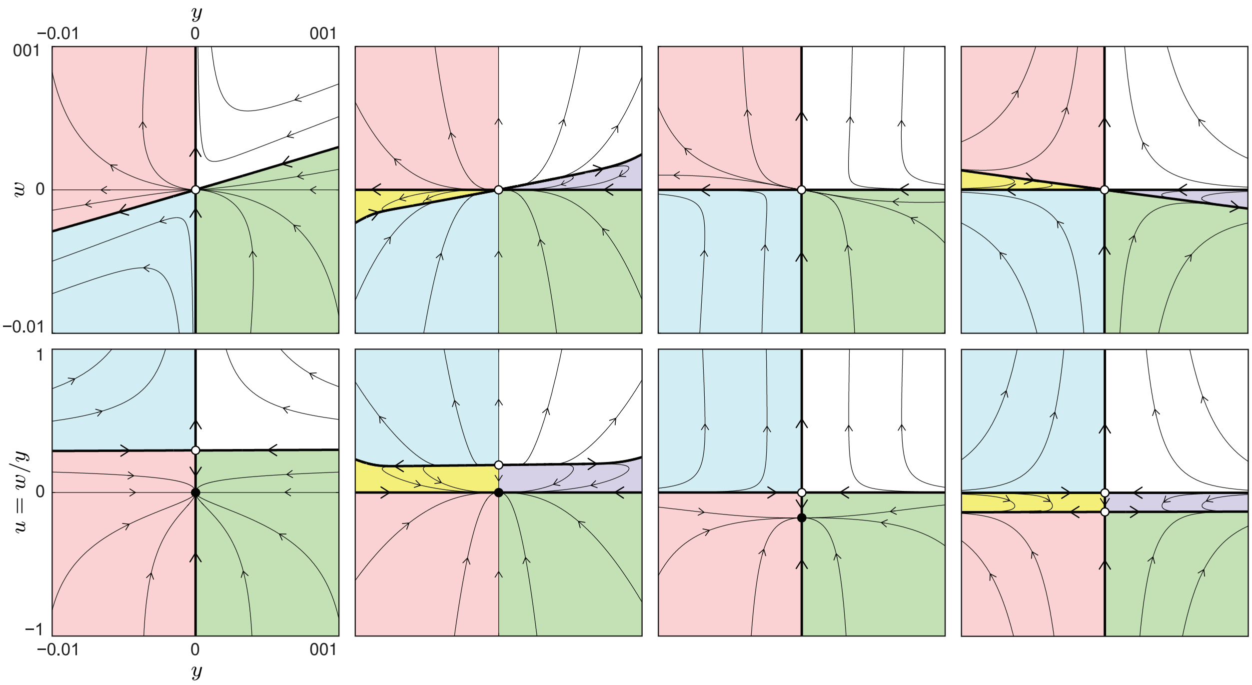

Figure A1: Original (top) and blown-up (bottom) dynamics. Parameter values: a, as in Fig. 1a,b; b, as in Fig. 1e$_1$; c, as in Fig. 1d$_1$ but $\gamma = 0.2$; d, as in c but $\beta = 3$.

and associated eigenvalues reported in Table A2 (note that $\bar{u}_1 = 0$, with eigenvector $v_y(\bar{u}_1) = (1,0)$, corresponds to the invariance of the y-axis).

If

$$r < 2\left(1+\frac{\gamma}{c}\right) \tag{A5}$$

(as in Fig. 1a,b), then equilibrium $(0,\bar{u}_1)$ is stable and $(0,\bar{u}_2)$ is unstable ($\lambda_u(\bar{u}_2) = -\lambda_u(\bar{u}_1) > 0$) with positive $\bar{u}_2$ (see Fig. A1a,b). If also $\lambda_y(\bar{u}_2) < 0$, i.e., if

$$c^2r(r-2)-2c((r-2)\beta+r\gamma)+4(\beta\gamma+\beta\sigma+\gamma\sigma) > 0, \tag{A6}$$

(see Table A2), then equilibrium $(0,\bar{u}_2)$ is a saddle (Fig. A1a) and the neighborhood of $(y,w) = (0,0)$ is partitioned into four sectors (the sector separatrices are the thick orbits in the figure). In two sectors system

(A3) behaves like a node (so-called parabolic sectors), stable for $y > 0$ and unstable for $y < 0$; in the other two sectors the system behaves like a saddle (hyperbolic sectors).

If the lefthand side of (A6) is negative, then the equilibrium $(0, \bar{u}_2)$ is a repellor (Fig. A1b). Part of the orbits emanating from it converge to $(0, \bar{u}_1)$ and correspond, in the original coordinates, to homoclinic orbits to $(y, w) = (0, 0)$. The neighborhood of $(y, w) = (0, 0)$ is again partitioned into four sectors, two of parabolic type (one attracting and one repelling) and two of so-called elliptic type (those with the homoclinic orbits).

The two situations are separated by a so-called pitchfork bifurcation (Kuznetsov, 2004; Meijer et al., 2009) for system (A4). Either an internal repellor is present when $(0, \bar{u}_2)$ is a saddle and there are no internal equilibria when $(0, \bar{u}_2)$ is a repellor (so-called sub-critical pitchfork, see Fig. 1d$_2$,d$_4$, respectively), or an internal saddle is present when $(0, \bar{u}_2)$ is a repellor and there are no internal equilibria when $(0, \bar{u}_2)$ is a saddle (super-critical pitchfork, see Fig. 1e$_1$, a). The internal equilibrium (and a symmetric one for $y < 0$ that is irrelevant for our purposes) collides with $z = 1$ at the bifurcation and determines the sector transition from case a to case b in Fig. A1 (d$_2 \rightarrow$ d$_4$ in Fig. 1) and vice versa (e$_1 \rightarrow$ a in Fig. 1).

When condition (A5) is reversed, then equilibrium $(0, \bar{u}_1)$ is a saddle ($\lambda_u(\bar{u}_1) > 0$), whereas $(0, \bar{u}_2)$, with negative $\bar{u}_2$, is stable under (A6) and a saddle if (A6) is reversed (see Fig. A1c,d). In both cases, however, the positive quadrant of the (y, w) plane is an hyperbolic sector, so $(y, w) = (0, 0)$ behaves like a saddle (see, e.g., Fig. 1d$_1$)).

In conclusion, we have proved that with peer- as well as shared-punishment (the only differences in the analysis being in the eigenvector $v_y(\bar{u}_2)$ and in the criticality of the pitchfork bifurcation that involve the third and higher order terms in $Y_k(1, \bar{u}_2)$ and $W_k(1, \bar{u}_2)$) there are feasible parameter settings (including the basic one used in Brandt et al. (2006)) for which a set of initial conditions with nonzero measure in the face $x = 0$ (shaded in Fig. A1a,b) is attracted by the loners' equilibrium $z = 1$.

A5 The classification algorithm

Initial conditions in the simplex interior are selected at the vertexes of regular tetrahedra with edge $1/n$ ($n > 2$) filling the simplex, without considering the vertexes on the simplex boundary. This results in

$$\sum_{i=1}^{n-3} \sum_{j=1}^{n-i-2} \sum_{k=1}^{n-i-j-1} 1 = \tfrac{1}{6}\,(n-1)(n-2)(n-3)$$

equally spaced initial conditions, in the sense that each of the strategies' frequencies $x(0)$, $y(0)$, $z(0)$, $w(0)$ is quantized into multiples $\varepsilon = \frac{1}{n}$ (there are $(n-3)$ ways of choosing $x(0) \in \{\frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-3}{n}\}$, as all frequencies must be positive; given $x(0) = i/n$ there are $(n-i-2)$ ways of choosing $y(0)$; given $y(0) = j/n$ there are $(n-i-j-1)$ ways of choosing $z(0)$; $w(0) = 1 - x(0) - y(0) - z(0)$). In the computation we have used $n = 100$ (156849 initial conditions).

The orbit of the replicator dynamics is followed until one of the conditions below is matched:

A, $x(t_1)+w(t_1) > 1-\varepsilon$ for some $t_1 > 0$ (at which $w(t_1) > w_{\min}$, see (5)) and $z(t) < 1-\varepsilon$ for $t \in [0, t_1]$;

B, $w(t) < \varepsilon$ for t in an interval $[t_1, t_2]$ in which $\dot{z}(t)$ changes sign n_{loop} times with $z(t) < 1-\varepsilon$ for $t \in [0, t_2]$ (we have used $n_{\text{loop}} = 10$);

C, $z(t) > 1-\varepsilon$ for some $t_1 > 0$.

In cases A and B the initial condition is associated to the corresponding attractor. In case C the orbit remains in the vicinity of the loners' equilibrium $z = 1$ for so long that, as described in Sect. 3.5, the deterministic description of the evolutionary dynamics becomes inappropriate. The simulation is therefore interrupted and the initial condition is temporarily associated to a "dummy" state "near C" (corresponding to the state $z(0) > 1-\varepsilon$ in the Markov chain of Fig. 3).

To complete the classification, we need to compute the probabilities p_A and p_B ($p_C = 1-p_A-p_B$) of the Markov chain. For this, a coarser set of initial conditions is used by filling the tetrahedron closest to $z = 1$ with smaller tetrahedra with edge ε/m ($m > 2$) and by considering the $onesixth(m-1)(m-2)(m-3)$ non-boundary vertexes. In the computation we have used $m = 20$ (969 initial conditions). Now the orbit of the replicator dynamics is followed until one of the conditions below is matched:

A, $x(t_1)+w(t_1) > 1-\varepsilon$ for some $t_1 > 0$ (at which $w(t_1) > w_{\min}$, see (5));

B, $w(t) < \varepsilon$ for t in an interval $[t_1, t_2]$ in which $\dot{z}(t)$ changes sign n_{loop} times with $z(t) > \varepsilon$ at the zeros of $\dot{z}(t)$;

C, the maximum time T is reached (we have used $T = 10^9$ and checked that $z(T) > 1-\varepsilon/m$).

The probabilities p_A and p_B are estimated as the fraction of initial conditions yielding cases A and B, respectively.

The classification is completed by counting as associated to attractors A, B, and C, respectively the fractions p_A, p_B, and p_C of the initial conditions in the "dummy" state.

A6 Computational issues

Functions $B(z)$ and $F(z)$ in (2) are not defined at $z = 1$. However, they can easily be defined there by continuity. They are actually smooth functions in $[0, 1]$, as they have the following polynomial expressions:

$$B(z) = \frac{1}{N} \sum_{k=1}^{N-1} k\, z^{N-1-k}, \quad F(z) = r\,(1-z)\,B(z) - (r-1)(1-z^{N-1}).$$

Similarly, the w at the denominator in (3c) makes the punishers' payoff not defined at $w = 0$. However, the following polynomial relations hold:

$$\begin{aligned} \frac{1}{w} G_y(w) &= \sum_{k=1}^{N-1} \binom{N-1}{k} w^{k-1} (1-w)^{N-1-k}, \\ \frac{1}{w} G_x(y, w) &= \frac{1}{w} G_y(w) - \sum_{k=1}^{N-1} \binom{N-1}{k} w^{k-1} (1-y-w)^{N-1-k}. \end{aligned}$$

The above polynomial expressions are numerically more accurate to be computed and have been used in the implementation of the average payoffs (1) and (3).