Optimal Control of Fractional Punishment in Optional Public Goods Game

J. GRAU[1], R. BOTTA[2] and C. E. SCHAERER[3]

ABSTRACT. Punishment is probably the most frequently used mechanism to increase cooperation in Public Goods Games (PGG); however, it is expensive. To address this problem, this paper introduces an optimal control problem that uses fractional punishment to promote cooperation. We present a series of computational experiments illustrating the effects of single and combined terms of the optimization cost function. In the findings, the optimal controller outperforms the use of constant fractional punishment and gives an insight into the period and size of the penalization to be implemented with respect to the defection in the game.

Keywords: optimal control, OPGG, fractional punishment, free-riders.

1 INTRODUCTION

Cooperation is fundamental for success in any human endeavor, but sustaining it is challenging. Over the years, cooperation has been a puzzle studied in various research areas, including natural and social sciences [1, 11, 20]. A key issue in studying the evolution of cooperation is identifying the requirements and mechanisms that can sustain cooperation over time [12, 19, 23, 26]. Two mechanisms have been extensively studied in the literature: a) the application of incentives (such as punishments or rewards) to increase the level of cooperation in the group [18, 21, 23, 30]; and b) the possibility of abstaining from participating, which allows the continuity of cooperation over time [14].

This paper focuses on enhancing cooperation in public goods games (PGG). In this context, punishment serves as a valuable mechanism for improving cooperation, but can be costly. A punishment system requires resources and thus becomes a public good itself [30]. Traditionally,

---

[1]School of Engineering, National University of Asuncion, San Lorenzo, Central, Paraguay – E-mail: jgrau@fiuna.edu.py https://orcid.org/0000-0003-0730-5029

[2]Polytechnic School, National University of Asuncion, San Lorenzo, Central, Paraguay – E-mail: rbotta@pol.una.py https://orcid.org/0000-0001-9778-1461

[3]Polytechnic School, National University of Asuncion, San Lorenzo, Central, Paraguay – E-mail: cschaer@pol.una.py https://orcid.org/0000-0002-0587-7704

punishment is implemented in PGG in two ways: as peer punishment or as pool punishment. These approaches differ based on when they are carried out (before or after the game) and the source of funds for sanctions (cooperators' game benefits or a separate pool set aside to penalize defectors) [5, 24].

Several methods for enforcing sanctions other than peer or pool punishment have been proposed in recent literature. Some include randomly selecting a group of cooperators to share the cost of sanctioning [7], implementing a fixed amount sanction [8], or capping the maximum payment for a defector [32]. In this context, a mechanism called fractional punishment, as presented in [4], reduces the cost of implementing the sanctioning system by punishing only a fraction of defectors in the population while still achieving overall cooperation within the group. This method allows for increased cooperation at a lower cost. However, it does have a drawback in that the proportion of punished free riders remains constant over time without accounting for expected variations based on the results of the applied sanctions.

To minimize the costs of the incentive systems in PGG, some works define an optimal control problem [27, 28, 29]. However, they do not consider free-riders' frequency with respect to the punishment. Consequently, the novelty in this work relies mainly on two points: a) we define an optimal control problem in the fractional punishment sanctioning system to consider the potential variation of the fractional punishment, and b) we introduce an additional term in the optimization objective function to consider the costs associated with sanctioning a single individual.

The results highlight the adaptability of introducing optimal control. Initially, when free-riders are abundant and the cost of punishment is high, a moderate percentage of free-riders are penalized. As the number of free-riders decreases, a more aggressive sanction is used, reducing the cost of punishment. Eventually, the control effort is reduced to a maintenance level, effectively discouraging new sporadic free-riding. This flexible approach significantly reduces the cost of implementing fractional punishment from multiple perspectives making it a versatile option for the final decision-maker.

Finally, this work builds on the idea that the cost of a punishment system is not only a function of how many free-riders are being punished (as discussed when fractional punishment was first introduced) but also of when and how effectively this punishment is being used. In this context, the optimization problem presented are the first attempt to answer such ideas.

The contents of this paper are organized as follows: In §2.1 the OPGG and the fractional punishment model are introduced. The optimal control formulation is defined in §3. In §4, the simulation results for the functional terms relative importance are exhibited. These results are discussed in §5, as well as a comparison between relevant punishment strategies and the optimal solutions found. Finally, the conclusions of the work are presented in §6.

2 CONTROL AND FRACTIONAL PUNISHMENT

In this section, we define the optional public goods game with fractional punishment in which the optimal control will be applied.

2.1 Optional public goods game

The public goods game is an economic experiment that can be used to study the evolution of cooperation in a group. The optional version of the game introduces the possibility of rejecting participation and being independent; therefore, the individual must first decide if he will participate and then whether he will contribute (to be a cooperator) or not (to be a free-rider) to the game. The OPGG system is defined as follows. Let $I = [t_0, t_f]$ denote a time interval for the analyzed period, given the state variable $w(t)$, that is comprised of the frequency of cooperators across time $x(t)$, the frequency of free-riders across time $y(t)$ and the frequency of independents (loners) across time $z(t)$, and for each $t \in I$, the state space is $\mathscr{S}_3 \subset \mathbb{R}^3$ defined as

$$\mathscr{S}_3 := \left\{ [x,y,z]^T \in \mathbb{R}^3 : x,y,z \geq 0 \text{ and } x+y+z = 1 \right\},$$

hence the state variable space is $\mathscr{Y} = \{w(t) \in C(I, \mathscr{S}_3) : \dot{w} \in C(I, \mathbb{R}^3)\}$.

Given an initial condition $w(t_0) = w_0 \in \mathscr{S}_3$ the state equation has the form [4]:

$$
\begin{aligned}
\dot{x} &= x(p_x(w) - \bar{p}(w,v)), \\
\dot{y} &= y(p_y(w,v) - \bar{p}(w,v)), \\
\dot{z} &= z(p_z - \bar{p}(w,v)),
\end{aligned}
\tag{2.1}
$$

where p_i is the expected payoff of the ith strategy, and $\bar{p}$ is its average given by $\bar{p} = x p_x + y p_y + z p_z$; v is the distributed control, that belongs to an admissible space $\mathscr{U} = L^2(I)$ and represents the fractional punishment (see §2.2). It can be shown that the problem (2.1) is well posed (see [4]).

Remark 1. *The simplex $\mathscr{S}_3$ remains invariant under the flow of Equation (2.1) (see [16, 31]); hence, if the state $w(t) \in \mathscr{S}_3$ at $t = t_0$, then $w(t) \in \mathscr{S}_3 \ \forall t$, and it is endowed with the standard norm $\| \cdot \|_2$.*

Remark 2. *In Expression (2.1), the dynamics rely on the expected payoff p_i of the ith strategy with respect to the average population payoff $\bar{p}$ in the sense that, if $p_i > \bar{p}$, then the proportion of the i strategy increases, otherwise it decreases or stays the same.*

To define the payoff in an OPPG, consider that a group of n individuals are invited to participate in the game. The group is therefore composed by n_c cooperators, n_d defectors, and n_l loners ($n = n_c + n_d + n_l$). Assuming that each contribution is equal and normalized ($c = 1$), the payoff of each corresponding strategy is [13, 15]:

$$p_c = r \frac{n_c}{s} - 1, \quad p_d = r \frac{n_c}{s}, \qquad p_l = \sigma, \tag{2.2}$$

where s is the number of players in the game ($s = n_c + n_d$), r is the multiplication factor by which the group contribution is multiplied and σ is the non-participating individual payoff. The parameters r, σ, and n are considered under the following assumptions:

Assumption 1. *The interest rate on the common pool r satisfies $1 < r < n$.*

Condition $1 < r$ means that if all individuals cooperate, they are better off than if all defect. Condition $r < n$ means that each individual is better off defecting by itself than cooperating [13].

Assumption 2. *The payoff σ of the loner strategy satisfies $0 < \sigma < r - 1$.*

This means that a cooperator in a group of cooperators (profiting $(r-1)$ each) is better off than loners that receive σ, but loners are better off than a defector in a group of defectors, where the payoff is equal to 0 [13].

2.2 Fractional punishment model

The fractional punishment is a mechanism to reduce the cost of the sanctioning system by punishing only a subset of the free-riders. Following [4], in this work we consider that only a fraction v $(0 \le v \le 1)$ of the defectors will be punished, modifying the payoff presented in (2.2). In particular, we assume that this set of randomly selected defectors will have their corresponding payoff reduced to 0, while the remaining free-riders will obtain the normal payoff, having an average payoff:

$$p_d = (1-v)\left(\frac{rj}{s}\right) + v0. \tag{2.3}$$

Expected payoffs are used in equations (2.1), since a player does not know in advance the composition of the group he is playing with. To this end, observe that the group composition depends on the frequencies of all strategies in the population.

The corresponding expected payoffs considering the fractional punishment for system (2.1) take the form [4]:

$$
\begin{align}
p_x(x,z) &= \mathbb{E}[p_c] = \sigma z^{n-1} + ra + rxb + (1-r)z^{n-1} - 1, \tag{2.4} \\
p_y(x,z,v) &= \mathbb{E}[p_d] = \sigma z^{n-1} + (1-v)\, rxb, \tag{2.5} \\
p_z(x,z,v) &= \mathbb{E}[p_l] = \sigma. \tag{2.6}
\end{align}
$$

where $a := \frac{1-z^n}{n(1-z)} = \frac{1}{n}\sum_{k=0}^{n-1} z^k$ and $b := \frac{1-a}{1-z} = \frac{1}{n}\sum_{k=0}^{n-2}(n-1-k)z^k$. Then it follows that payoffs in expression (2.1) are polynomials.

3 OPTIMAL CONTROL FORMULATION

In this section, we define a problem of minimization with constraints to find the controller v for the OPGG system (2.1). The cost function for minimization comprises the weighting of four objectives, each with its significance [10, 22].

Our control effort v corresponds with the fraction of defectors that are punished. This control belongs to an admissible space $\mathscr{U} = L^2(t_0, t_f)$, where in our application $v \in [0,1]$. We indicate the dependence of w on v using the notation $w(v)$. Given a target function w^* in $L^2(t_0, t_f)$ and

parameters $\alpha_i \geq 0$, we shall employ the following cost function, which we associate with the state equation (2.1) [9, 17, 22]:

$$
\begin{aligned}
J(w(v)), v) \quad = \quad & \frac{\alpha_1}{2}\|w(t_f, v) - w^*\|_2^2 + \frac{\alpha_2}{2}\int_{t_0}^{t_f}\|w(\tau, v) - w^*\|_2^2 d\tau \\
& + \frac{\alpha_3}{2}\int_{t_0}^{t_f} v^2(\tau)d\tau + \frac{\alpha_4}{2}\int_{t_0}^{t_f}(v(\tau)y(\tau))^2 d\tau.
\end{aligned}
\tag{3.1}
$$

The weights establish the importance of each of the functional terms, *i.e.*, α_1 weights the cost of not being in the desired state w^* at t_f, α_2 weights the accumulated squared error of trajectory w, α_3 weights the value of the controller at all times, and finally α_4 weights the frequency of sanction individuals at all times. In other words, the term associated with α_4 measures the the costs of implementing the control effort given the current frequency of defectors in the population, assuming this two values are proportional (as they usually are).

The optimal control problem for system (2.1) consists of finding a controller $u \in \mathscr{U}$ which minimizes the cost function (3.1) as

$$
J(w, u) := \min_{v \in \mathscr{U}} J(w(v), v) \; s.t. \; (2.1).
\tag{3.2}
$$

Remark 1. *The integrands of the cost-function* (3.1) *belong to* C^2. *Additionally, its associated Hamiltonian function is given by*

$$
H(t, w, v, \lambda) = \frac{\alpha_2}{2}\|w(t, v) - w^*\|_2^2 + \frac{\alpha_3}{2}v^2(t) + \frac{\alpha_4}{2}v^2(t)y^2(t) + \lambda^T f(t, w, v),
$$

is strictly convex with respect to v for w_0 being interior to $\mathscr{S}_3$ and $\alpha_3 > 0$ or $\alpha_4 > 0$; then

$$
H_{vv} = \alpha_3 + \alpha_4 y^2 > 0, \forall t \geq t_0.
$$

This determines necessary and sufficient conditions to obtain a unique optimal solution that minimizes (3.1) *(see* [2, 25]*).*

Next, we discuss the implemented algorithm and the Python's library used for the implementation.

3.1   The GEKKO library and the algorithm

We use a Python package called GEKKO for numerical implementation, specifically designed for optimization and control problems [3]. This package offers three free solvers in its public distribution: APOPT, BPOPT, and IPOPT. This work will use the default solver APOPT unless otherwise indicated.

A pseudo-code of the algorithm implemented in the GEKKO is presented in Algorithm 1.

Algorithm 1 Solving the minimization of the cost function.

1: Declare initial parameters such as number of points, initial and final time, maximum value for v, initial state w_0, weights $[\alpha_1, \alpha_2, \alpha_3, \alpha_4]$, and the values of n, r and σ mentioned in section 2;
2: Initialize the model;
3: Declare all variables and equations;
4: Define the cost function as a minimization objective;
5: Choose a solver and execute it;
6: Display results;

---

4 RESULTS

To appropriately choose the weights in the cost function (3.1), we start by analyzing each element on its own, for example, $[\alpha_1, \alpha_2, \alpha_3, \alpha_4] = [1, 0, 0, 0]$, to understand the single importance of α_1 and so forth. Then, we proceed to examine the most important pairs of combinations between the elements of the cost function (such as α_2 and α_3 or α_2 and α_4). Lastly, we merge the previous findings to obtain a combination that yields better results, which we compare with relevant constant fractional punishment strategies.

In all cases, the parameters mentioned in §2 necessary for defining the problem are as follows: $n = 5$, $r = 3$, and $\sigma = 1$. To simplify the notation, from now on, we will refer to the importance of $\frac{\alpha_1}{2} \|w(t_f, v) - w^*\|_2^2$ as "the importance of α_1". Similar acronyms will be used for the other terms of the functional (3.1).

4.1 Importance of α_1

The value of α_1 is important for minimizing the final state error at time t_f. This situation is observed in Figure 1, which illustrates a simulation where the final state is close to full cooperation. However, when the only requirement is for the final state to be close to full cooperation (vertex $x = 1$), the trajectory can oscillate in the state space and closely approach the vertex $z = 1$. This situation means that in the context of a public good, simply minimizing this error does not guarantee sustainability for the whole period since approaching $z = 1$ could lead to bankruptcy caused by the abandonment of most of the game participants.
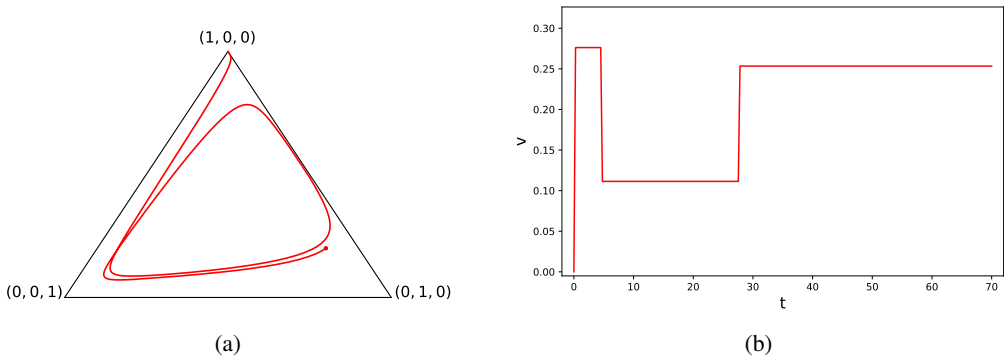
Figure 1: **Importance of** α_1**. Part 1.** Results obtained by simulating from 0 to 70 time units with 250 steps, from initial state $w_0 = [0.2, 0.7, 0.1]^T$ (marked with a dot), using the controller curve shown in b. a) State space and trajectory of the system. b) Control effort v over time.
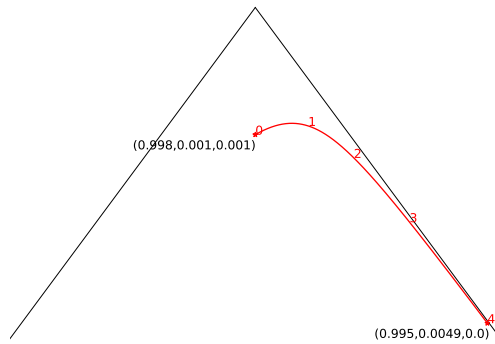

Figure 2: **Importance of** α_1**. Part 2.** Result obtained by simulating from 0 to 4 time units with 600 steps, with no controller ($v = 0$), starting at $w_0 = [0.998, 0.001, 0.001]^T$. A zoomed in simplex border can be seen in black, along with red time stamps in the trajectory traveled. The initial state error amounts to around 0.002449 while the final state error amounts to 0.007008, showing how unnoticeable small time deviations could be at the end of the simulation for this system.

Note that because the system is continuous, it cannot quickly deviate from the state of full co-operation (vertex $x = 1$). Hence, if the goal is to minimize the state error at all times (see §4.2) rather than just minimizing the final state error at time t_f, a boundary layer will emerge. This situation is shown in Figure 2. However, depending on the desired accuracy, this solution can be acceptable. With this consideration, α_1 is excluded from the subsequent analysis by setting $\alpha_1 = 0$.

4.2   Importance of α_2

Considering only the accumulated state error across the entire time period, we aim to reduce the error as quickly as possible without considering the magnitude of the punishment costs. This is observed in Figure 3. The control effort reaches its maximum value, indicating that we are penalizing all free-riders, leading to the fastest reduction in error possible. This trend continues throughout the entire period, ensuring that no further error accumulates once we are close to, or have reached, the desired state w^*.

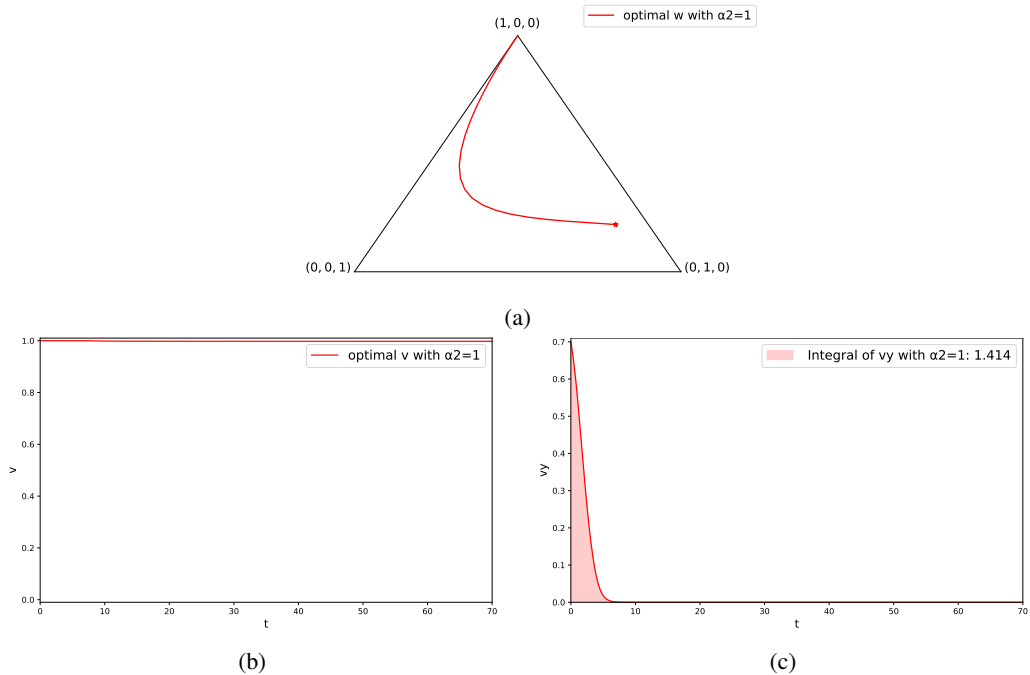Figure 3: **Importance of α_2.** Result obtained by simulating from 0 to 70 time units with 250 steps, only considering α_2 and with starting state $w_0 = [0.2, 0.7, 0.1]^T$. a) State space and trajectory of the system. b) Control effort v over time. c) Proportion of punished individuals yv over time. The red shade represents the area under the curve and the legend shows its value.

In practice, for most real-world systems, it is impossible to sanction all free riders at all times. This impossibility may be due to the high cost of punishing or the difficulty of identifying and addressing all (or most) free riders. However, the obtained numerical result is the minimum estimation for all scenarios of the accumulated errors.

4.3   Importance of α_3 and α_4

When analyzing the importance of α_3, the optimal controller minimizes v to its minimum value (*i.e.*, keeping it at zero the entire period). If this occurs, full cooperation is not achieved, defeating the purpose of the whole controller.

A similar result is obtained when analyzing the importance of α_4. This is because minimizing the frequency of sanctioned individuals implies avoiding sanctioning any of them (*i.e.*, keeping v at zero for the entire period, as with the case of α_3). Figure 4 shows these results[1]. Although the single choice of the terms α_3 or α_4 leads to the no punishment system behavior, they become necessary for combining with a term containing the state error, such as α_2, to achieve a trade-off between minimizing the state error and applying punishment. This will be analyzed in forthcoming sections.

(a)

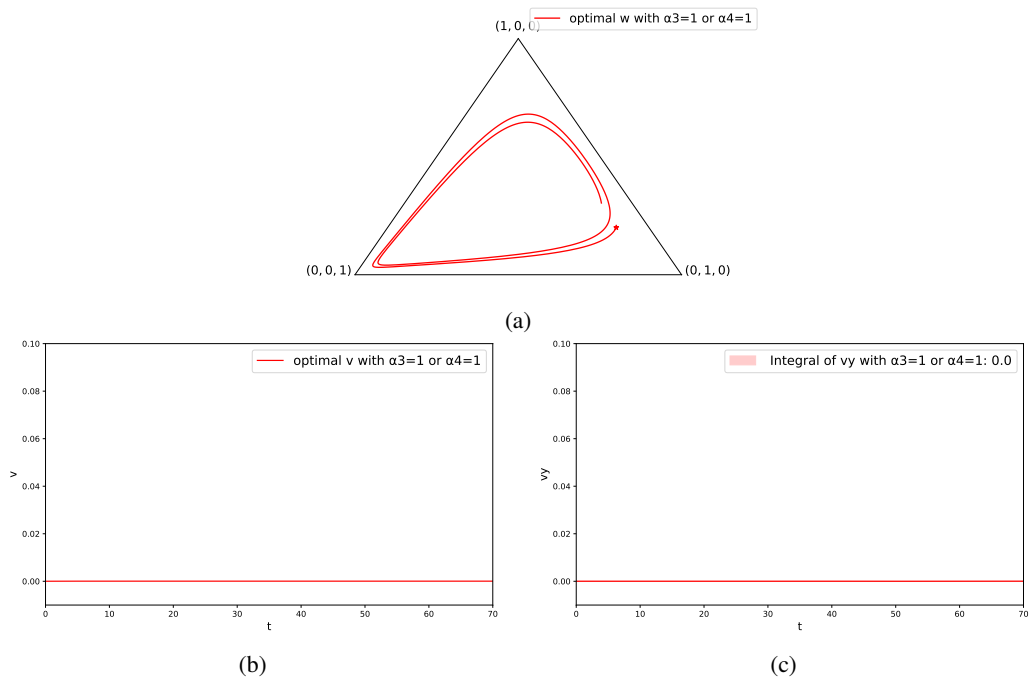(b)   (c)

Figure 4: **Importance of α_3 or α_4.** Result obtained by simulating from 0 to 70 time units with 600 steps, minimizing the control effort or the amount of sanctioned individuals, with starting state $w_0 = [0.2, 0.7, 0.1]^T$. a) State space and trajectory of the system. b) Control effort v over time. c) Proportion of punished individuals yv over time. The red shade represents the area under the curve which in this case is zero.

---

[1]The numerical system trajectory in Figure 4 presents a discrepancy with respect to the closed system trajectory of the free punishment system (see [4]: Figure 1a). This is due to the numerical integrator implemented [3].

4.4   Relative importance of α_3 with respect to α_2

When aiming to reduce costs while enhancing collaboration, it is attractive to explore using a combination of α_2 and α_3 in the cost function (3.1). This approach enables us to balance minimizing the state error trajectory with penalizing only a necessary fraction of free-riders.

To explore the relationship between α_2 and α_3, we initially assume that α_2 is much greater than α_3. Afterwards, we decrease the value of α_2, while ensuring that $\alpha_2 + \alpha_3 = 1$ for all cases. In the corresponding figures, we use the following color sequence: red, blue, green, cyan, and magenta, which aligns with the mentioned procedure. This process is repeated in the subsequent sections.

When α_2 is much more important than α_3 (observe in Figure 5a and b) full cooperation is achieved. This is obtained by applying an important amount of punishment at the beginning and maintaining a small amount of punishment at the end.

As α_2 becomes smaller with respect to α_3 (as seen in Figure 6a and b) a boundary layer emerges in the state trajectory and the controller. Consequently, depending on the pre-specified final error accuracy, the state trajectory can be undesirable for obtaining full cooperation.
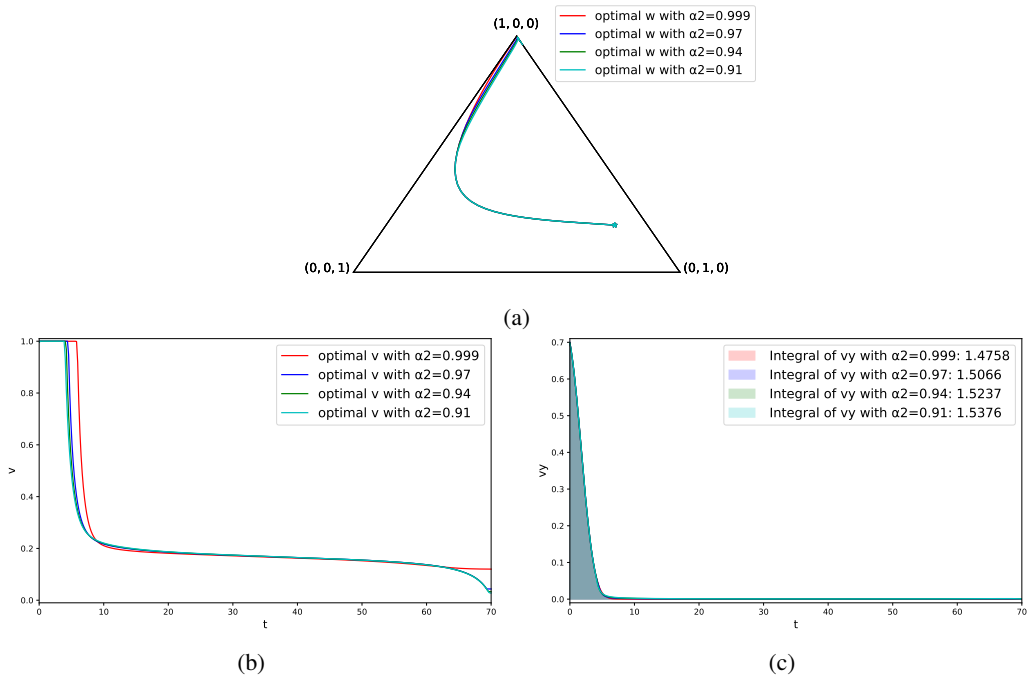

Figure 5: **Relative importance of α_3 with respect to α_2. Part 1.** Result obtained by simulating from 0 to 70 time units with 400 steps, such that α_2 is 0.999, 0.97, 0.94 and 0.91 and with starting state $w_0 = [0.2, 0.7, 0.1]^T$. a) State space and trajectory of the system. b) Control effort v over time. c) Proportion of punished individuals yv over time. The corresponding shade represents the area under the curve and the legend shows its value.
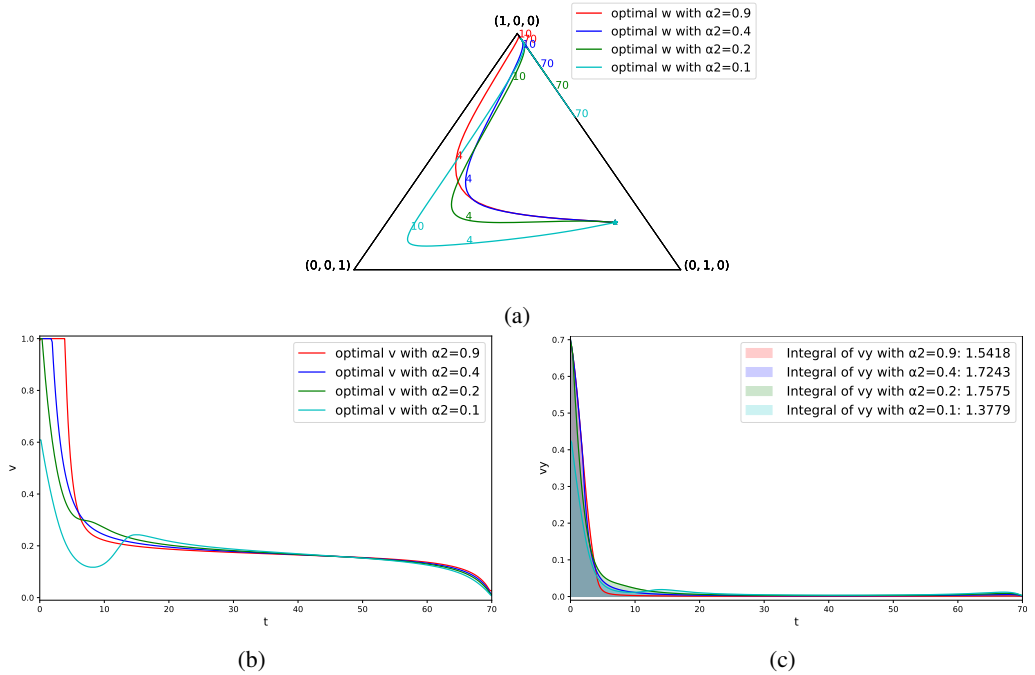
(a)

(b)

(c)

Figure 6: **Relative importance of α_3 with respect to α_2. Part 2.** Result obtained by simulating from 0 to 70 time units with 400 steps, such that α_2 is 0.9, 0.4, 0.2 and 0.1 with starting state $w_0 = [0.2, 0.7, 0.1]^T$. a) State space and trajectory of the system with timestamps in each corresponding color. b) Control effort v over time. c) Proportion of punished individuals yv over time. The corresponding shade represents the area under the curve and the legend shows its value.

4.5 Relative importance of α_4 with respect to α_2

In the real world, another important cost-related factor is the number of punished individuals. To minimize this, we consider the terms associated with the values of α_2 and α_4 in the cost function (3.1). We use a similar methodology to the previous section, maintaining the same color ordering and ensuring that $\alpha_2 + \alpha_4 = 1$ while reducing the value of α_2.

Figure 7 shows that when the value of α_4 is progressively increased, it reduces both the initial impulse of the controller (Figure 7) and the number of punished individuals (Figure 7). This provides a counter-intuitive punishment strategy, since in a phase of high defection, the controller punishes fewer individuals. On the contrary, when the value of α_4 is decreased, it generates higher values for both the control effort and the number of punished individuals (see Figures 8 and c)). Still, after a certain minimum threshold value of α_4, the state trajectory becomes more oscillatory in the simplex (see Figures 7 and 8).

(a)

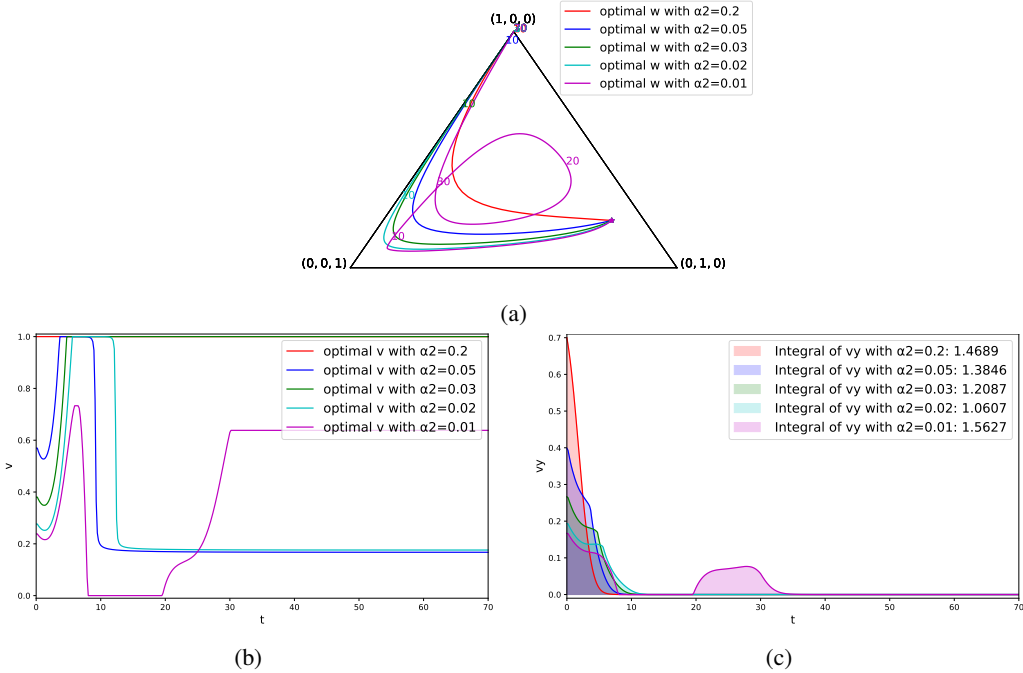(b)   (c)

Figure 7: **Relative importance of α_4 with respect to α_2. Part 1.** Result obtained by simulating from 0 to 70 time units with 400 steps, such that α_2 is 0.2, 0.05, 0.03, 0.02 and 0.01 (with timestamps for this last case) and with starting state $w_0 = [0.2, 0.7, 0.1]^T$. a) State space and trajectory of the system. b) Control effort v over time. c) Proportion of punished individuals yv over time. The corresponding shade represents the area under the curve and the legend shows its value.

The results show that to achieve full cooperation, we can allow for an oscillatory state trajectory by sanctioning a very small number of individuals (represented by the integral of yv). However, in practice, this can lead to a high level of desertion and make the institution susceptible to bankruptcy. In general, even if we do not consider oscillatory trajectories, the results reveal an unexpected effect: we can actually improve cooperation by reducing the number of sanctioned individuals during periods of high free-riding and increasing it during periods of more cooperation.

4.6   Combining the effects of α_2, α_3 and α_4

By leveraging the obtained relations between α_2 and α_3, and α_2 and α_4, we design a controller that ensures full cooperation while also considering the costs of imposing sanctions during periods of high defection. This approach minimizes the control effort to the optimal level required to sustain cooperation.
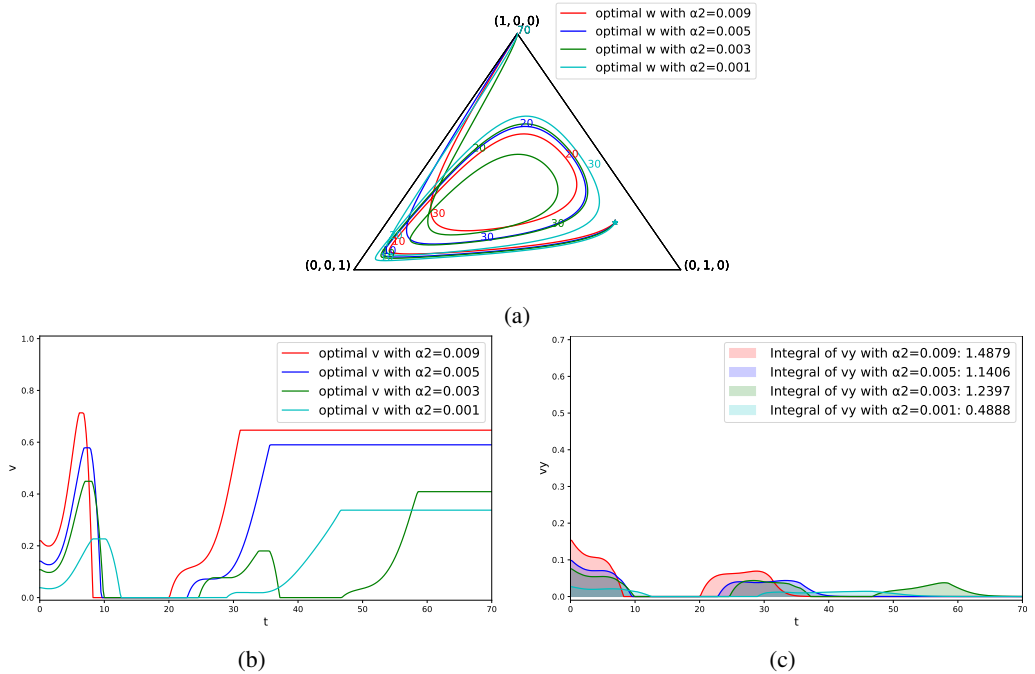
Figure 8: **Relative importance of α_4 with respect to α_2. Part 2.** Result obtained by simulating from 0 to 70 time units with 400 steps, such that α_2 is 0.009, 0.005, 0.003 and 0.001 and with starting state $w_0 = [0.2, 0.7, 0.1]^T$. a) State space and trajectory of the system. b) Control effort v over time. c) Proportion of punished individuals yv over time. The corresponding shade represents the area under the curve and the legend shows its value.

Similarly to previous cases, we begin with a low value of α_2 while maintaining $\alpha_3 = 0.0001$ and $\alpha_2 + \alpha_3 + \alpha_4 = 1$. Afterward, we increase the relative importance of the term associated with α_2 while decreasing the importance of the amount of sanctioned individuals (α_4).

Perhaps the most remarkable result is shown in Figure 9. Observe the controller's effort is smaller during high defection periods while being more aggressive when defection has gone down. Finally, it maintains a constant value at the end to discourage defection. The regulation of weights α_2 and α_4 manifest almost as a time shift of the plateaus in this graph.

Something notable to mention is that when time increases, the value of v for the test scenarios tends toward the value for which the vertex $x = 1$ becomes an attractor. In this case, it is around $v = 1/6$ [6]. Yet, it is not confirmed if this result is just a coincidence given the chosen parameters or if it unveils a special relation of the model.
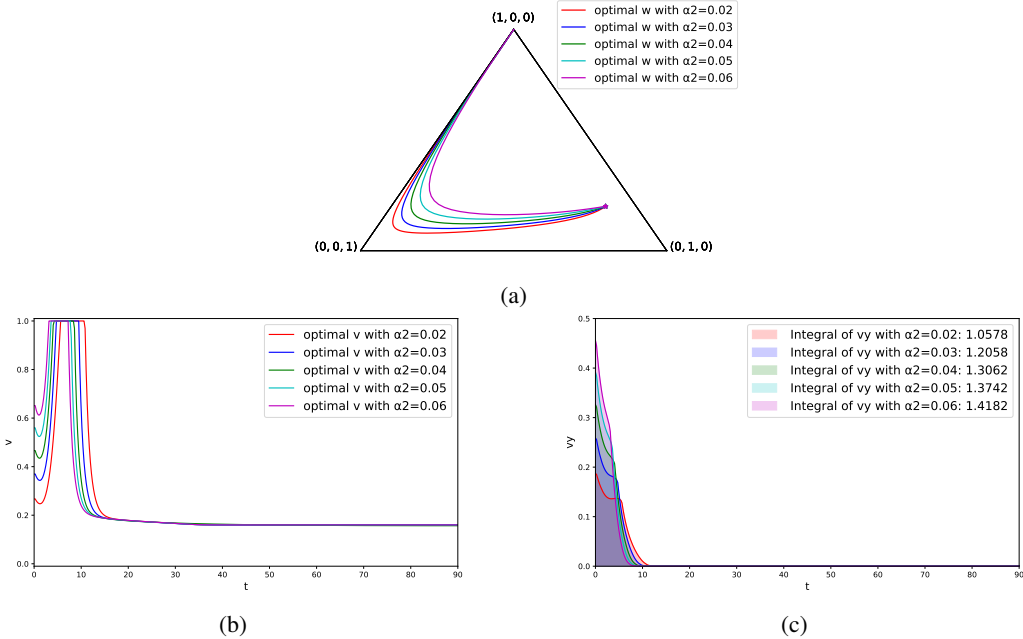
(a)

(b)

(c)

Figure 9: **Combining results. Part 1.** Result obtained by simulating from 0 to 90 time units with 400 steps, such that α_2 is 0.02, 0.03, 0.04, 0.05 and 0.06 and with starting state $w_0 = [0.2, 0.7, 0.1]^T$. a) State space and trajectory of the system. b) Control effort v over time. c) Proportion of punished individuals yv over time. The corresponding shade represents the area under the curve and the legend shows its value.

4.7   A comparison between punishment strategies

In this section, we use a particular scenario (and its associated cost function) to compare finding an optimal solution with using a simple constant value of fractional punishment (or eventually punishing all free-riders at all times). To this end, consider the following cost function:

$$J_1 = \int_0^{20} \left[\frac{0.04}{2} \|w(\tau, v) - w^*\|_2^2 + \frac{0.001}{2} v^2(\tau) + \frac{0.959}{2} y^2(\tau) v^2(\tau) \right] d\tau. \qquad (4.1)$$

For initial conditions $w_0 = [0.2, 0.7, 0.1]^T$, each pair of constant fractional punishment and trajectory $(v, w(v))$, has an associated cost given by $J_1(t_0 := 0, t_f := 20, v, w)$. Figure 10 shows a curve of values of the cost function (4.1) for several values of v.

Figure 11 shows two relevant cases tested against the optimal trajectory obtained using Algorithm 1. It is observed that the optimal solution has a lower value for the cost function and a lower amount of punished individuals (given by the integral of yv). Table 1 shows the computational time and the cost function value for some representative values of v and the optimal controller $v(t)$. Notice that obtaining the optimal solution requires less time than obtaining a solution by simulating the 101 cases seen in the tests for Figure 10, even when using smaller time steps to find the optimal solution.
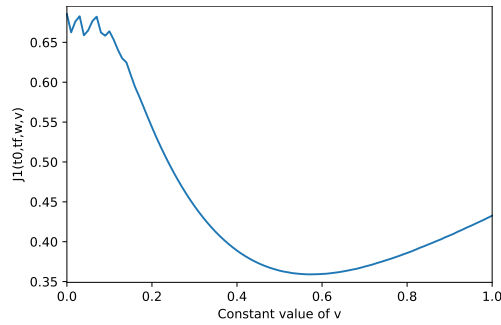
Figure 10: **A comparison between punishment strategies. The cost function when a constant punishment is applied.** Values of cost function (4.1) obtained by taking 101 constant and equally spaced values of v ranging from 0 to 1. Each scenario was tested from initial state $w_0 = [0.2, 0.7, 0.1]^T$ using 20 time units and 400 time steps. From the tests, there is a clear implication that there is a minimum value of J_1 at $v \approx 0.57$, and small deviations from this value do not result in big changes of J_1.
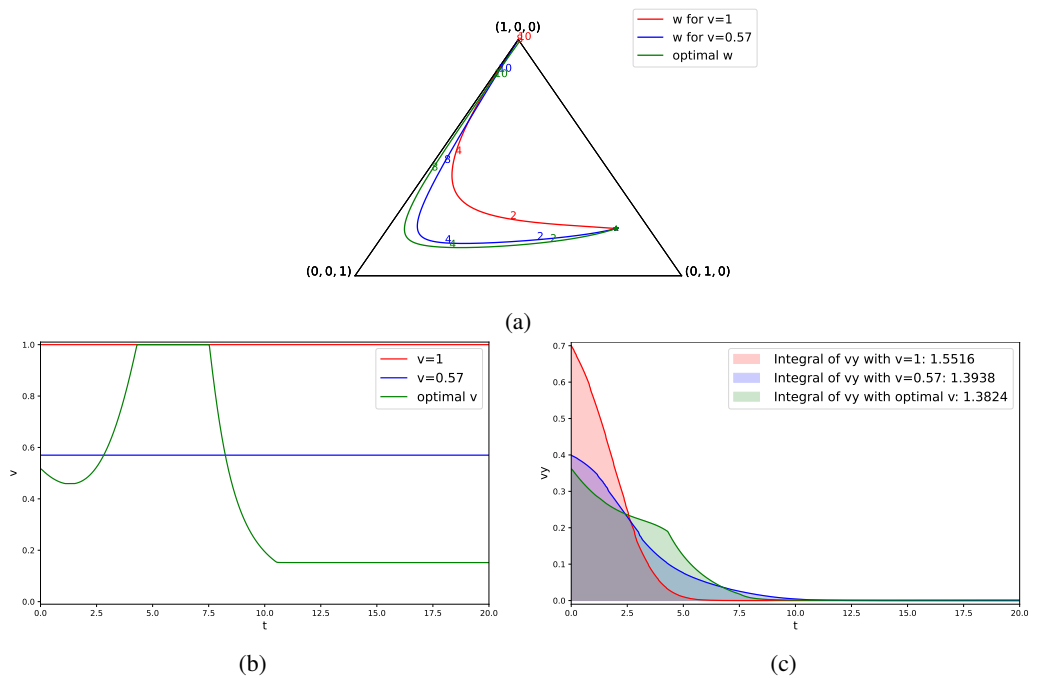

Figure 11: **A comparison between punishment strategies. Constant fractional punishments against the optimal one.** Result obtained by simulating from 0 to 20 time units with 1200 steps, starting state $w_0 = [0.2, 0.7, 0.1]^T$ and using $v = 1$ (in red), $v = 0.57$ (in blue) and an optimal v (in green) as punishment strategies. a) State space and trajectory of the system with some color coded time stamps. b) Control effort v over time. c) Proportion of punished individuals yv over time. The corresponding shade represents the area under the curve and the legend shows its value.

Table 1: Cost comparison between control strategies.

Strategy	Computation time (using the IPOPT solver)	Computed value of J_1
Constant $v = 1$	≤ 15s	0.4400079
Constant $v = 0.57$	≤ 1500s	0.3638641
Optimal v	≤ 30s	0.3503187

Notes. This table presents some key features of the comparison shown in Figure 11. For this case in particular, the IPOPT solver was used, since even though the APOPT solver has been reliable for most of this work. The IPOPT solver is much faster, even when using many more time steps. For the computation time, for row $v = 1$, it just reflects the approximated simulation time; for row $v = 0.57$, it mainly reflects the process of finding the minimum for J_1 among the 101 values of v mentioned in Figure 10 and simulating it again with more time steps.

5   DISCUSSION

Trying to account for the most common costs associated with running a public good, three of them naturally arise: the difference between the desired state of cooperation and the actual state of the system (state error), the fraction of free-riders that will be sanctioned and the total amount of people that will be sanctioned. To solve this problem, a weighted cost function that takes into account all this factors is designed, using weights α_2, α_3 and α_4 for the relevant costs mentioned before.

When analyzing each part of the cost function individually, it is observed from the experiments that if any long-term improvement of cooperation is desired, α_2 must be considered. Nevertheless, at least one of α_3 or α_4 should be used to decrease the sanctioning efforts by reducing the fraction of sanctioned individuals (see §4.1 to §4.5).

As shown in §4.6 a combination of α_2, α_3, and α_4 can be used, such that cooperation improves (because of α_2), fractional punishment is low when there are too many free-riders to sanction (because of α_4). It is also low when the maintenance of cooperation does not deem it necessary to sanction so many people (because of α_3). This is considered a good illustration of the capabilities of the optimal solution obtained for this minimization problem.

Finally in §4.7 an optimal solution for a specific problem generated from cost function (4.1) is compared with two representative alternative sanctioning strategies: a) sanctioning all free-riders and b) sanctioning a constant fraction of free-riders close to the minimizing value of the cost function.

In summary, this research demonstrates that the optimal solution is better than the best constant fractional punishment in terms of the cost function value and the frequency of punished individuals. It is also found in a considerably shorter time, and the sanctions are applied when they are most effective. These findings clearly show the advantage of the optimal solution over the alternative of finding the best constant fractional punishment.

6   CONCLUSION

The central focus of this paper has been the optimal control of the free-rider problem in an optional public good game using fractional punishment as the controller. Several objective minimization functions were tested based on the error concerning a specific equilibrium point (full cooperation $(x, y, z) = (1, 0, 0)$), the controller's effort, and the frequency of punished free-riders.

An important term introduced to minimize the cost of punishing is associated with the effort to implement the control given the current frequency of free riders in the population. This term relates to the real per-unit costs of punishing in a public good or service, where the costs depend on the number of punished free-riders.

Numerical experiments evidence that the optimal solutions found result in a lower value of the cost function compared to punishing all or a constant fraction of free-riders at all times.

REFERENCES

[1] R. Axelrod & W. Hamilton. The evolution of cooperation. *Science*, **211**(27) (1981), 1390 – 1396. URL `http://wuos.org/content/211/4489/1390.short`.

[2] J. Baumaister & A. Leitão. "Introdução à Teoria de Controle e Programação Dinâmica". IMPA, 1st. edition ed. (2014), 399 p.

[3] L. Beal, D. Hill, R. Martin & J. Hedengren. GEKKO Optimization Suite. *Processes*, **6**(8) (2018), 106. doi:10.3390/pr6080106.

[4] R. Botta, G. Blanco & C.E. Schaerer. Fractional Punishment of Free Riders to Improve Cooperation in Optional Public Good Games. *Games*, **12**(1) (2021). doi:10.3390/g12010017. URL `https://www.mdpi.com/2073-4336/12/1/17`.

[5] R. Botta, G. Blanco & C.E. Schaerer. Discipline and punishment in panoptical public goods games. *Sci. Rep.*, **14**(7903) (2024). doi:https://doi.org/10.1038/s41598-024-57842-0.

[6] R. Botta, C.E. Schaerer & G. Blanco. Cooperation and punishment in community managed water supply system. In "Conference of Computationar Interdisciplinary Science, 2019, Atlanta. Anais eletrônicos. Campinas, Galoá, 2019." (2020).

[7] X. Chen, A. Szolnoki & M. Perc. Probabilistic sharing solves the problem of costly punishment. *New Journal of Physics*, **16** (2014). doi:10.1088/1367-2630/16/8/083016.

[8] F. Dercole, M. De Carli, F. Della Rossa & A.V. Papadopoulos. Overpunishing is not necessary to fix cooperation in voluntary public goods games. *Journal of theoretical biology*, **326** (2013), 70–81. doi:10.1016/j.jtbi.2012.11.034. URL `http://dx.doi.org/10.1016/j.jtbi.2012.11.034`.

[9] X. Du, M. Sarkis, C.E. Schaerer & D.B. Szyld. Inexact and truncated Parareal-in-time Krylov subspace methods for parabolic optimal control problems. *Electron. Trans. Numer. Anal.*, **40** (2013), 36–57.

[10] J. Grau, R. Botta & C.E. Schaerer. Optimal Control of Fractional Punishment in Optional Public Goods Game. In "Proceeding Series of the Brazilian Society of Computational and Applied Mathematics", volume 9. SBMAC - Sociedade de Matemática Aplicada e Computacional, São Carlos/SP (2022), p. 1–2.

[11] W.D. Hamilton. The genetical evolution of social behaviour. I. *Journal of theoretical biology*, **7**(1) (1964), 1–16. URL http://www.ncbi.nlm.nih.gov/pubmed/5875341.

[12] G. Hardin. The Tragedy of the Commons. *Science*, **162** (1968), 1243–1248. URL http://www.tandfonline.com/doi/abs/10.1080/19390450903037302.

[13] C. Hauert, S. De Monte, J. Hofbauer & K. Sigmund. Replicator dynamics for optional public good games. *Journal of Theoretical Biology*, **218**(2) (2002), 187–194. doi:10.1006/jtbi.2002.3067.

[14] C. Hauert, S. De Monte, J. Hofbauer & K. Sigmund. Volunteering as Red Queen mechanism for cooperation in public goods games. *Science (New York, N.Y.)*, **296**(5570) (2002), 1129–32. doi:10.1126/science.1070582. URL http://www.ncbi.nlm.nih.gov/pubmed/12004134.

[15] C. Hauert, N. Haiden & K. Sigmund. The dynamics of public goods. *Discrete & Continuous Dynamical Systems - B*, **4**(3) (2004), 575–587. doi:10.3934/dcdsb.2004.4.575. URL http://aimsciences.org//article/doi/10.3934/dcdsb.2004.4.575.

[16] J. Hofbauer & K. Sigmund. "Evolutionary Games and Population Dynamics". Cambridge University Press, Cambridge (1998), 323 p. doi:10.1017/CBO9781139173179. URL http://ebooks.cambridge.org/ref/id/CBO9781139173179.

[17] T.P. Mathew, M. Sarkis & C.E. Schaerer. Analysis of Block Pararel Preconditioners for Parabolic Optimal Control Problems. *SIAM Journal on Scientific Computing*, **32**(3) (2010), 1180–1200. doi:10.1137/080717481.

[18] M. Nowak & R. May. Evolutionary games and spatial chaos. *Nature*, **359** (1992). URL http://tinyurl.com/8sahrcf.

[19] M.A. Nowak. Five Rules for the Evolution of Cooperation. *Science*, **314**(December) (2006).

[20] E. Ostrom. "Governing the commons: The evolution of institutions for collective action". Cambridge university press (1990).

[21] T. Sasaki, A. Brännström, U. Dieckmann & K. Sigmund. The take-it-or-leave-it option allows small penalties to overcome social dilemmas. *Proceedings of the National Academy of Sciences of the United States of America*, **109**(4) (2012), 1165–9. doi:10.1073/pnas.1115219109. URL http://tinyurl.com/99dfe2n.

[22] C.E. Schaerer, T. Mathew & M. Sarkis. Block Iterative Algorithms for the Solution of Parabolic Optimal Control Problems. In M. Daydé, J.M.L.M. Palma, Á.L.G.A. Coutinho, E. Pacitti & J.C. Lopes (editors), "High Performance Computing for Computational Science - VECPAR 2006". Springer Berlin Heidelberg, Berlin, Heidelberg (2007), p. 452–465.

[23] K. Sigmund. Punish or perish? Retaliation and collaboration among humans. *Trends in ecology & evolution*, **22**(11) (2007), 593–600. doi:10.1016/j.tree.2007.06.012. URL `http://www.ncbi.nlm.nih.gov/pubmed/17963994`.

[24] K. Sigmund, C. Hauert, A. Traulsen & H. Silva. Social Control and the Social Contract: The Emergence of Sanctioning Systems for Collective Action. *Dynamic Games and Applications*, **1**(1) (2010), 149–171. doi:10.1007/s13235-010-0001-4. URL `http://www.springerlink.com/index/10.1007/s13235-010-0001-4`.

[25] J.L. Speyer & J.D. H. "Primer on Optimal Control Theory". Advances in Design and Control. SIAM (2010).

[26] R.L. Trivers. The Evolution of Reciprocal Altruism Published by : The University of Chicago Press Stable URL :. *The Quarterly Review of Biology*, **46**(1) (1971), 35–57. URL `http://www.jstor.org/stable/2822435`.

[27] S. Wang, X. Chen & A. Szolnoki. Exploring optimal institutional incentives for public cooperation. *Communications in Nonlinear Science and Numerical Simulation*, **79** (2019), 104914. doi:https://doi.org/10.1016/j.cnsns.2019.104914. URL `https://www.sciencedirect.com/science/article/pii/S1007570419302357`.

[28] S. Wang, X. Chen, Z. Xiao & A. Szolnoki. Decentralized incentives for general well-being in networked public goods game. *Applied Mathematics and Computation*, **431** (2022), 127308. doi:https://doi.org/10.1016/j.amc.2022.127308. URL `https://www.sciencedirect.com/science/article/pii/S0096300322003824`.

[29] S. Wang, X. Chen, Z. Xiao, A. Szolnoki & V.V. Vasconcelos. Optimization of institutional incentives for cooperation in structured populations. *Journal of the Royal Society Interface*, **20**(199) (2023), 20220653.

[30] T. Yamagishi. The provision of a sanctioning system as a public good. *Journal of Personality and Social Psychology*, **51**(1) (1986), 110–116. doi:10.1037/0022-3514.51.1.110. URL `http://doi.apa.org/getdoi.cfm?doi=10.1037/0022-3514.51.1.110`.

[31] E.C. Zeeman. Population dynamics from game theory. In "Global theory of dynamical systems, Lecture Notes in Mathematics", volume 819. Springer, Berlin, Heidelberg (1980), p. 471–497. doi:10.1007/BFb0087009. URL `https://www.lms.ac.uk/sites/lms.ac.uk/files/1980`.

[32] J. Zhang, Y. Zhu, Z. Chen & M. Cao. Evolutionary Game Dynamics Driven by Setting a Ceiling in Payoffs of Defectors. In "Proceedings of the 36th Chinese Control Conference July 26-28, 2017, Dalian, China" (2017), p. 11289–11295.