

# A ROBUST SATURATED STRATEGY FOR $N$ -PLAYER PRISONER'S DILEMMA

GIULIA GIORDANO\*, DARIO BAUSO<sup>†</sup>, AND FRANCO BLANCHINI<sup>‡</sup>

**Abstract.** We study diffusion of cooperation in an  $n$ -population game in continuous time. At each instant, the game involves  $n$  random individuals, one from each population. The game has the structure of a Prisoner's dilemma, where each player can choose a continuous decision variable associated with the probability of cooperating or defecting. We turn the game into a positive dynamical system. Then, we propose a novel strategy that is the saturation of a polynomial function. The strategy requires to each player exclusively the knowledge of her/his own current average payoff, along with her/his own payoffs in the cooperative and non-cooperative equilibria; no information about other players' payoffs is required. The proposed strategy guarantees local stability of the cooperative equilibrium if the degree  $p$  of the polynomial is greater than or equal to 2. Conversely, the non-cooperative equilibrium becomes unstable, for  $p$  large enough, if and only if a certain Metzler matrix depending on the payoffs has a positive Frobenius eigenvalue. We prove that the  $n$ -dimensional box of all payoffs between the non-cooperative and the cooperative ones is positively invariant. Finally we show that, for  $p$  large, the domain of attraction of the cooperative equilibrium inside this box becomes arbitrarily close to the box itself.

**Key words.** Prisoner's dilemma, Dynamic games, Stability, Equilibria, Invariant sets.

**AMS subject classifications.** 90D-93A-93D

**1. Introduction.** Diffusion of cooperation in society is a core topic at the intersection between engineering and social science. A common paradigmatic model is the repeated Prisoner's dilemma, for which a variety of versions exist. Starting from Smale's definition of good strategies in [44], the evolution of cooperation strategies has been first studied within the framework of asymptotic pseudotrajectories in chain recurrent flows [11]. In a subsequent work, the same game has been discussed in the context of stochastic approximation and differential inclusions [10]. There, as well as in [9], connections with *approachability* in repeated games with vector payoffs [13] have also been discussed in detail (an overview is provided below). A common assumption in [9, 10, 11] is that the decision of each player is based on the knowledge of the whole current average payoff vector. Differently, [32] introduces a simple adaptive rule that, as shown by simulations, yields convergence to the cooperative equilibrium without requiring any knowledge about the mutual dependence of players' payoffs.

Blackwell's Approachability Theorem provides conditions for a set to be approachable. The present paper links to approachability insofar as the players interact repeatedly and adjust their strategies based on the current time-average vector payoff, and convergence of the time-average vector payoff to pre-specified sets is investigated. Approachability has been used to prove convergence in different application domains, such as allocation processes in coalitional games [34], regret minimisation [36, 25], adaptive learning [17, 20, 23, 24], excludability and bounded recall [37], and weak approachability [47]. In the context of cooperative games with transferable utilities,

---

\*Delft Center for Systems and Control, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands. [g.giordano@tudelft.nl](mailto:g.giordano@tudelft.nl)

<sup>†</sup>Jan C. Willems Center for Systems and Control, ENTEG, Fac. Science and Engineering, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands, and Dipartimento dell'Innovazione Industriale e Digitale (DIID), Università di Palermo, V.le delle Scienze, 90128 Palermo, Italy. [dario.bauso@unipa.it](mailto:dario.bauso@unipa.it)

<sup>‡</sup>Dipartimento di Matematica, Informatica e Fisica, Università degli Studi di Udine, 33100 Udine, Italy. [blanchini@uniud.it](mailto:blanchini@uniud.it)

the theorem is used to prove that the core is an approachable set under opportune allocation processes. Within the realm of regret minimisation, one wishes to design strategies such that the nonpositive orthant in the space of regrets is an approachable set; there the underlying idea is that a player needs to adapt her/his strategy based on the current regret to make the nonpositive orthant approachable. It can be shown that, when all players have nonpositive regrets, the resulting outcome is an equilibrium for the game. A similar concept can be found in adaptive learning and evolutionary games [43]. We formulate our game in continuous time. The original formulation of approachability is in discrete time and has been adapted to continuous-time repeated games in [25]. There the authors also highlight the connection with Lyapunov theory. The extension to infinite-dimensional space is due to Lehrer [35]. Approachability shares striking similarities with differential game theory and, as such, can be studied using differential calculus and stability theory [39, 45]. Approachability and differential inclusions [2] are studied in [39]. There the authors also highlight that Blackwell's theorem is a generalisation of von Neumann's minimax theorem [48]. We refer the reader to [45] for a set-valued analytical perspective [1, 3]. Remarkably, approachable and discriminating sets can be reframed within the context of set invariance theory [15]. A core concept in approachability is the one of nonanticipative strategy, which resembles the same concept in differential games [7, 18, 45, 42, 46]. Note that classical feedback strategies in differential games are special nonanticipative strategies. A complementary notion is the one of excludability; conditions for a set to be excludable are discussed in [37] under different circumstances: imperfect information, bounded recall, delayed and/or stochastic monitoring. Another concept related to approachability is the one of *attainability* [8, 38]. Attainability can play a role in different application domains such as transportation networks, distribution networks, and production networks.

In this paper, we consider an  $n$ -player repeated Prisoner's dilemma in *mixed strategies*, where each player may choose to cooperate or defect *with probability*. The game describes a scenario where i) the cooperation of a player benefits the rest of the players, ii) defectors benefit from engaging with cooperators, and iii) full cooperation is more profitable to anybody than full non-cooperation. After introducing a new strategy based on partial information, and studying the evolution of the  $n$ -player game, we extend the model to a population game framework and we investigate the diffusion of cooperation in the population. In the case of structured environment, we use a graph to describe the interaction topology and we link the strategy to the network properties; in particular, the connectivity (or degree) of a node (i.e., the number of neighbours of the corresponding player) reflects in the power coefficient of the strategy. In such a scenario, once a random player is chosen, s/he plays repeatedly with  $p$  players: we view this as  $p$  independent identically distributed experiments. The strategy captures the joint probability of defecting in all  $p$  repeated games (or equivalently, one minus the probability of cooperating at least once out of  $p$  plays).

As a main contribution, we propose a novel saturated strategy where each player's decision is based on the exclusive knowledge of her/his own current average payoff and on her/his own payoff in the fully cooperative and fully non-cooperative cases. It is worth stressing that information about the other players is not required, which makes this strategy well suited also for games with incomplete information [4]. We extend preliminary results in [21], limited to the case of 2-player games, to the game with  $n$  players and we revisit the game model as a population game with structured or unstructured interaction environment, providing the following contributions.

- The strategy is a saturated polynomial function parametrised in the degree  $p$ , which represents a *patience* parameter. For  $p = 1$ , we have a linear saturated strategy, meaning immediate reaction of each player to non-cooperation on the part of the other players; for larger  $p$  we have, initially, a patient reaction.
- Patience pays: for  $p \geq 2$ , the cooperative equilibrium becomes locally stable.
- By taking  $p$  large, we can locally destabilise the non-cooperative equilibrium if and only if a Metzler matrix depending on the system payoffs has a positive real dominant (Frobenius) eigenvalue.
- There exists a *distrust region*  $\mathcal{D}$  in the space of payoffs, for which the non-cooperative equilibrium solution is an attractor: all of the trajectories originating in the region converge to non-cooperation. The region is characterised by small initial values of the average payoff for all players (global distrust).
- The hyper-rectangle  $\mathcal{R}_{\alpha,\delta}$  in the space of payoffs, associated with the non-cooperative and the cooperative payoff values, is a positively invariant and attractive set for the system.
- Trajectories originating in the positive orthant outside the distrust region  $\mathcal{D}$  are ultimately bounded inside the hyper-rectangle  $\mathcal{R}_{\alpha,\delta}$ .
- The domain of attraction to the cooperative equilibrium inside  $\mathcal{R}_{\alpha,\delta}$  becomes arbitrarily close to the whole hyper-rectangle as  $p$  becomes sufficiently large: hence the non-cooperative equilibrium, even if locally stable, has a local domain of attraction in  $\mathcal{R}_{\alpha,\delta}$  that is arbitrarily small if  $p$  is sufficiently large.
- Under suitable assumptions on the payoffs, the proposed strategy is a *good strategy*: it is never profitable for any player to switch from this strategy to steady defection.
- In relation to a population game formulation, we can design the strategy to have maximal diffusion of cooperation in the population, which occurs when the damage for the cooperator is less than the benefit for the defector in random matching.
- In a structured environment, convergence to the cooperative equilibrium means that each player cooperates with all of her/his neighbours, thus providing steady-state configuration of the co-evolving network.

The parameter  $p$  bears resemblance to the existing concept of ‘delay’ of a player in responding to defectors/cooperators when using the so-called ‘inertia strategies’ [5]. The parameter  $p$  can also be used to weigh the ‘altruistic nature’ of a player who adopts proportional tit-for-tat strategies, see, e.g., zero determinant memory-one strategies for discrete-time games developed in [26, 27, 28, 29, 30, 31]. It is worth to emphasise the link with risk-seeking behavioural strategies as developed in Kahneman and Tversky’s prospect theory [33]. By increasing  $p$  we assume that the players are increasingly robust against unfavourable events (noncooperation of the other players) and more sensitive to favourable events (cooperation of the other players).

The game is introduced in Section 2, while Section 3 presents some needed preliminaries. Section 4 provides stability results, Section 5 investigates the system dynamics inside  $\mathcal{R}_{\alpha,\delta}$ , and Section 6 studies the conditions under which the strategy is *good*, i.e. no player has interest in abandoning the strategy and defecting. Section 7 provides a population-game perspective. Section 8 discusses a numerical example and proposes simulations that illustrate the evolution of the game with the proposed saturated strategy. Section 9 provides concluding comments and directions for future work.

**2. The  $n$ -Player Game.** In the considered Prisoner’s dilemma with  $n$  players, each player decides either to cooperate or to defect, in order to maximise her/his own

payoff. The average payoff vector  $[x_1 \ x_2 \ \dots \ x_n]^\top$  evolves according to a repeated game, depending on the players' choice. The decision variable associated with player  $k$  is denoted by  $\omega_k \in \{0, 1\}$ , where

- $\omega_k = 1$  means that player  $k$  cooperates;
- $\omega_k = 0$  means that player  $k$  defects.

Then, the payoff function for player  $k$ , resulting from the decision of all the players, is given by

$$\phi^{(k)}(\omega_1, \omega_2, \dots, \omega_n) > 0.$$

Since the game is repeated, the outcome is filtered over time. This corresponds to considering a mixed strategy, namely, a continuous decision variable in the interval

$$u_1^k \in [0, 1],$$

which represents the cooperation frequency for player  $k$ , along with the complementary variable

$$u_0^k \doteq 1 - u_1^k \in [0, 1],$$

which represents the defection frequency. For instance, if  $u_1^k = \frac{1}{2}$ , then player  $k$  cooperates half of the times s/he is involved in a play. In the case of a population, this number reflects the inclination of the individuals towards cooperation.

Given the *instantaneous payoff*  $\phi_k$  of player  $k$ , we can define its *average payoff over the time interval*  $(0, \tau]$  as

$$x_k(\tau) \doteq \frac{1}{\tau} \int_0^\tau \phi_k(\theta) d\theta.$$

In principle, the evolution of this variable would lead to an equation that is not time-homogeneous. However, following [45], we can denote the final time as  $\tau = e^t$  and, with a change of variables, obtain a homogeneous equation for  $x_k(e^t)$ :  $\dot{x}_k(t) = -x_k(t) + \phi_k(t)$ . The variable change does not affect the average computation, only the interval is changed (see [45] for a derivation). Moreover, since the strategies we consider are memoryless, they are not affected by time scaling.

In what follows, for brevity, we will denote by  $x_k(t)$  the (average) *payoff of player*  $k$ , obtained by averaging the *instantaneous payoff*  $\phi_k(t)$  over the time interval  $(0, e^t]$ .

The payoff  $x_k(t)$  of player  $k$  evolves over time according to the equation:

$$(2.1) \quad \dot{x}_k = -x_k + \sum_{\omega \in \Omega} \phi^{(k)}(\omega_1, \omega_2, \dots, \omega_n) u_{\omega_1}^1 u_{\omega_2}^2 \dots u_{\omega_n}^n,$$

where the sum is extended to all Boolean choices  $\omega = (\omega_1, \omega_2, \dots, \omega_n) \in \Omega = \{0, 1\}^n$ .

For instance, if all the players cooperate, then we have  $u_1^k = 1$  and  $u_0^k = 0$  for all  $k$ , and the payoffs satisfy the equation [45]

$$\dot{x}_k = -x_k + \phi^{(k)}(1, 1, \dots, 1).$$

In this case, the players converge to the cooperative equilibrium point  $\bar{x}_k = \phi^{(k)}(1, 1, \dots, 1)$  for all  $k$ .

We work under the following standing assumption.

ASSUMPTION 1. *The game payoffs satisfy the following three conditions.*

- (a) *The cooperation of player  $h$  is more profitable than non-cooperation for all others: for all  $k \neq h$  and for any choice of the  $\omega_i$ , with  $i \neq h$ ,*

$$\phi^{(k)}(\omega_1, \omega_2, \dots, \underbrace{1}_{=\omega_h}, \dots, \omega_n) > \phi^{(k)}(\omega_1, \omega_2, \dots, \underbrace{0}_{=\omega_h}, \dots, \omega_n).$$

- (b) *Conversely, defection advantages the traitor: for all  $k$  and for any choice of the  $\omega_i$ , with  $i \neq k$ ,*

$$\phi^{(k)}(\omega_1, \omega_2, \dots, \underbrace{0}_{=\omega_k}, \dots, \omega_n) > \phi^{(k)}(\omega_1, \omega_2, \dots, \underbrace{1}_{=\omega_k}, \dots, \omega_n).$$

- (c) *For all players, full cooperation is more profitable than full non-cooperation:*

$$\phi^{(k)}(1, 1, \dots, 1) > \phi^{(k)}(0, 0, \dots, 0), \quad \text{for all } k. \quad \diamond$$

We consider a decentralised strategy in which each player knows in advance

- her/his own payoff in the fully cooperative case,  $\alpha_k \doteq \phi^{(k)}(1, 1, \dots, 1)$ ;
- her/his own payoff in the fully non-cooperative case,  $\delta_k \doteq \phi^{(k)}(0, 0, \dots, 0)$ .

Note that  $\alpha_k > \delta_k$  for all  $k$ , in view of Assumption 1(c).

We consider a repeated continuous-time game with the following rules.

- At each time instant, each player adopts a memoryless strategy exclusively depending on her/his own payoff  $x_k(t)$ :

$$(2.2) \quad \begin{cases} u_1^k(t) = u_1^k(x_k(t)), \\ u_0^k(t) = u_0^k(x_k(t)) = 1 - u_1^k(x_k(t)). \end{cases}$$

- Each player ignores the actions of the other players.
- Each player ignores the payoff functions of the other players.
- All players act continuously in time, selecting at each time instant the strategy  $u_1^k(t)$ , and  $u_0^k(t) = 1 - u_1^k(t)$ , subject to the constraints  $0 \leq u_1^k(t) \leq 1$  and  $0 \leq u_0^k(t) \leq 1$ .

The rationale is that, in practical situations of repeated games (such as, for instance, the transmission problem discussed in Section 8.2), an overall cross-surveillance among players is impossible. Each player must decide whether to cooperate based on the evaluation of her/his own payoff only. Variables  $u_1^k(t)$  (respectively  $u_0^k(t)$ ) represent the probability of cooperating (respectively non-cooperating). In a practical context where the decisions are associated with discrete events, this would roughly correspond to the fraction of times when the player cooperates in a small interval  $\Delta t$ .

REMARK 1. *Although  $u_1^k(t)$  and  $u_0^k(t)$  could be regarded as probabilities, we are considering a purely deterministic problem. This is different in nature from the case of stochastic dynamic games (see [41]).*

Define the functions

$$(2.3) \quad \sigma_k(x_k(t)) = \left[ \text{sat}_{[0,1]} \left( \frac{\alpha_k - x_k(t)}{\alpha_k - \delta_k} \right) \right]^p,$$

where  $p$  is a positive integer and  $\text{sat}_{[0,1]}$  is the saturation function:

$$\text{sat}_{[0,1]}(\xi) = \begin{cases} 0, & \text{for } \xi < 0, \\ \xi, & \text{for } 0 \leq \xi \leq 1, \\ 1, & \text{for } \xi > 1. \end{cases}$$

Then, the proposed strategy for each player is:

$$(2.4) \quad \begin{cases} u_1^k(x_k(t)) = 1 - \sigma_k(x_k(t)), \\ u_0^k(x_k(t)) = \sigma_k(x_k(t)). \end{cases}$$

REMARK 2. For each player, the proposed strategy is **exclusively based** on her/his own average payoff ( $x_k(t)$ ) and on her/his own payoff in the fully cooperative ( $\alpha_k$ ) and fully non-cooperative ( $\delta_k$ ) cases; no information about the other players is required. If the current average payoff of a player is lower than expected (which must be due to non-cooperation of some other players), then s/he can decide to stop cooperating, without any actual knowledge of the decision/strategy of the other players. For instance, consider  $m$  vendors who agree on a fair common price for their product (which would result in a certain average profit for each of them, based on an expected demand); if one vendor decides to secretly break the deal and lower the price to attract more customers, then the other vendors see their current cumulative profit decrease with respect to their expectation and, based on this information only, they can decide, in turn, to reduce the price (up to the minimum level) to contrast the negative trend. As we will show, the proposed strategy allows a player to properly react to non-cooperation, even without having any explicit knowledge of the other players and of their strategies. Parameter  $p$  has a fundamental role and represents patience:  $p = 1$  induces a linear reaction to values of  $x_k$  that are smaller than  $\alpha_k$ , while  $p = 2$  induces a weaker quadratic reaction, and so on.  $\diamond$

In this framework, our main goal is to establish which values of the payoffs and of the patience parameter  $p$  lead to cooperation, given the proposed strategy.

**3. Invariance of the Payoff Polytope.** Given  $x = [x_1 \ x_2 \ \dots \ x_n]^\top$ ,  $\omega = (\omega_1, \omega_2, \dots, \omega_n)$  and  $\Phi(\omega) = [\phi^{(1)}(\omega) \ \phi^{(2)}(\omega) \ \dots \ \phi^{(n)}(\omega)]^\top$ , the dynamical system can be written in vector form as:

$$(3.1) \quad \dot{x} = -x + \sum_{\omega \in \Omega} \Phi(\omega) u_{\omega_1}^1(x_1) u_{\omega_2}^2(x_2) \dots u_{\omega_n}^n(x_n),$$

where  $\Omega = \{0, 1\}^n$ . Denoting by

$$(3.2) \quad \mathcal{V} = \text{conv}\{\Phi(\omega), \omega \in \Omega\}$$

the convex hull of all the  $2^n$  points  $\Phi(\omega)$ , with  $\omega \in \Omega$ , the following result holds.

THEOREM 3.1. *The polytope (3.2) is robustly positively invariant and attractive for system (3.1) for any choice of the payoff functions in vector  $\Phi(\omega)$  and for any possible choice of the functions  $u_1^k(t) \in [0, 1]$  and  $u_0^k(t) = 1 - u_1^k(t)$ .  $\square$*

*Proof.* For any choice of  $u_{\omega_k}^k(t) \in [0, 1]$ , it can be shown by induction that

$$(3.3) \quad v \doteq \sum_{\omega \in \Omega} \Phi(\omega) u_{\omega_1}^1 u_{\omega_2}^2 \dots u_{\omega_n}^n \in \mathcal{V}.$$

In fact, for  $n = 2$ ,

$$\begin{aligned} v &= \sum_{\omega \in \Omega} \Phi(\omega_1, \omega_2) u_{\omega_1}^1 u_{\omega_2}^2 \\ &= \Phi(1, 1) u_1^1 u_1^2 + \Phi(1, 0) u_1^1 (1 - u_1^2) + \Phi(0, 1) (1 - u_1^1) u_1^2 + \Phi(0, 1) (1 - u_1^1) (1 - u_1^2), \end{aligned}$$

where  $u_1^1 u_1^2$ ,  $u_1^1(1 - u_1^2)$ ,  $(1 - u_1^1)u_1^2$ ,  $(1 - u_1^1)(1 - u_1^2)$  are nonnegative and sum up to one, hence  $v$  is in the convex hull  $\mathcal{V}$ . Assume that the assertion is true for a given  $n$ . If we add a new variable  $(n + 1)$ , we can write  $v$  as

$$v = u_1^{n+1} \sum_{\omega \in \Omega} \Phi(\omega, 1) u_{\omega_1}^1 u_{\omega_2}^2 \dots u_{\omega_n}^n + (1 - u_1^{n+1}) \sum_{\omega \in \Omega} \Phi(\omega, 0) u_{\omega_1}^1 u_{\omega_2}^2 \dots u_{\omega_n}^n.$$

Since the two sums are in the convex hull, their convex combination belongs to the convex hull as well. Therefore, the inclusion (3.3) holds for any  $n$ .

Then, we can write the system as

$$\dot{x} = -x + v, \quad v \in \mathcal{V},$$

where  $\mathcal{V}$  is any convex and compact set (in our case, a polytope). We show that  $\mathcal{V}$  is positively invariant and attractive for this system. Indeed, consider the support functional  $\psi$  of  $\mathcal{V}$ , defined in accordance with [40] as  $\psi(z) = \max_{v \in \mathcal{V}} z^\top v$ . Then  $\mathcal{V} = \{v \in \mathbb{R}^n : z^\top v \leq \psi(z), \forall z \in \mathbb{R}^n\}$ .

Consider the function  $\zeta = z^\top x$ , with  $z \neq 0$  arbitrary. Since, for all  $v \in \mathcal{V}$ , it is  $z^\top v \leq \psi(z)$  for any  $z$ , then  $\dot{\zeta} = -z^\top x + z^\top v \leq -z^\top x + \psi(z) = -\zeta + \psi(z)$ . Necessarily we have  $\limsup_{t \rightarrow \infty} \zeta \leq \psi(z)$  for all  $z$ , hence any trajectory  $x(t)$  converges to the set  $\mathcal{V}$ , and this proves that  $\mathcal{V}$  is attractive. Also, in view of the differential inequality  $\dot{\zeta} \leq -\zeta + \psi(z)$  derived above, the actual trajectory is bounded as  $\zeta(t) \leq \bar{\zeta}(t)$ , where  $\bar{\zeta}(t)$  is the solution of  $\dot{\bar{\zeta}} = -\bar{\zeta} + \psi(z)$ , which monotonically converges to  $\psi(z)$ . Therefore, if  $z^\top x(t_0) \leq \psi(z)$  for any  $z$  (hence,  $x(t_0) \in \mathcal{V}$ ), then  $z^\top x(t) \leq \psi(z)$  for any  $z$  (hence,  $x(t) \in \mathcal{V}$ ) for all  $t \geq t_0$ . This proves the positive invariance of the set  $\mathcal{V}$ .  $\square$

Theorem 3.1 entails the following result, which will be crucial in assessing whether the strategy is good (cf. Definition 6.1 in Section 6).

**COROLLARY 3.2.** *If  $n - r$  players defect (namely,  $u_{\omega_k}^k = 0$  for  $k = r + 1, \dots, n$ ), then the trajectory  $x(t)$  of system (3.1) converges to the convex hull of the  $2^r$  points  $\Phi(\omega_1, \omega_2, \dots, \omega_r, 0, \dots, 0)$ , with  $\omega_i \in \{0, 1\}$  for all  $i = 1, \dots, r$ .  $\square$*

The next corollary ensures that the overall system is positive, namely that the positive orthant is a positively invariant set for the system, regardless of the chosen strategy and payoff functions.

**COROLLARY 3.3.** *For any choice of the payoff functions in vector  $\Phi(\omega)$  and of the strategy functions  $u_1^k(t) \in [0, 1]$  and  $u_0^k(t) = 1 - u_1^k(t)$ , system (3.1) is positive.  $\square$*

#### 4. Stability Analysis of Cooperative and Non-cooperative Equilibria.

In this section, the results in [21] are generalised to the  $n$ -dimensional case. First we show that, when the proposed strategy is adopted with  $p \geq 2$ , the cooperative point is a robustly stable equilibrium.

**THEOREM 4.1.** *For any choice of the payoff functions, the cooperative point  $\bar{x} = \Phi(1, \dots, 1)$ , namely  $\bar{x}_k = \alpha_k$  for all  $k$ , is a locally asymptotically stable equilibrium for system (3.1) with strategy (2.4), provided that  $p \geq 2$ .  $\square$*

*Proof.* If  $\bar{x}_k = \alpha_k$ , then  $\sigma_k(\bar{x}_k) = 0$  and  $u_1^k = 1$  for all  $k$ . Therefore, only one term in the sum survives and the system equation becomes

$$\dot{x} = -\bar{x} + \Phi(1, 1, \dots, 1) = 0,$$

hence  $\bar{x}$  is an equilibrium point. To prove its stability, since the function  $\sigma_k(x_k)$  with  $p \geq 2$  is continuously differentiable at  $x_k = \alpha_k$ , we can apply a linearisation criterion.

We have seen that  $\sigma_k(\alpha_k) = 0$ . By differentiating  $\sigma_k(x_k)$  with respect to  $x_k$ , we have

$$\frac{d\sigma_k(x_k)}{dx_k} = \begin{cases} \frac{-p(\alpha_k - x_k)^{p-1}}{(\alpha_k - \delta_k)^p}, & \text{for } x_k \leq \alpha_k, \\ 0, & \text{for } x_k > \alpha_k, \end{cases}$$

which is also zero at  $x_k = \alpha_k$ . Therefore, the derivative of all the terms

$$u_{\omega_1}^1, u_{\omega_2}^2, \dots, u_{\omega_n}^n$$

is zero at  $x_k = \alpha_k$  and the Jacobian at the cooperative equilibrium is  $J = -I$ , where  $I$  is the identity matrix. This implies local asymptotic stability.  $\square$

Also the non-cooperative point is an equilibrium, whose instability can be ensured under suitable assumptions, since a player moving away from the non-cooperative equilibrium increases the payoffs of all other players.

Before stating this result, we need to show the invariance of the hyper-rectangle

$$\mathcal{R}_{\delta, \alpha} = \{x \in \mathbb{R}^n : \delta_i \leq x_i \leq \alpha_i, \quad i = 1, \dots, n\}.$$

**PROPOSITION 4.2.** *The hyper-rectangle  $\mathcal{R}_{\delta, \alpha}$  is positively invariant for system (3.1) with strategy (2.4).*  $\square$

*Proof.* We show that  $\dot{x}_i \geq 0$  when  $x_i = \delta_i$ , while  $\dot{x}_i \leq 0$  when  $x_i = \alpha_i$ . This implies positive invariance of  $\mathcal{R}_{\delta, \alpha}$  [15, 16].

Let us consider the game from the standpoint of player 1. Since the cooperation of other players is profitable by assumption, the payoff of player 1 in the fully non-cooperative case is  $\delta_1 = \phi^{(1)}(0, 0, \dots, 0) \leq \phi^{(1)}(0, \omega_2, \dots, \omega_n)$ , while in the fully cooperative case it is  $\alpha_1 = \phi^{(1)}(1, 1, \dots, 1) \geq \phi^{(1)}(1, \omega_2, \dots, \omega_n)$ . If  $x_1 = \delta_1$ , then player 1 is not cooperating, namely  $\omega_1 = 0$ , hence  $u_1^1 = 0$  and  $u_0^1 = 1$ . Then

$$\begin{aligned} \dot{x}_1 &= -\delta_1 + \sum_{\omega_1=0} \phi^{(1)}(0, \omega_2, \dots, \omega_n) u_0^1 u_{\omega_2}^2 \dots u_{\omega_n}^n \\ &\geq -\delta_1 + \sum_{\omega_1=0} \phi^{(1)}(0, 0, \dots, 0) [1] u_{\omega_2}^2 \dots u_{\omega_n}^n \\ &= -\delta_1 + \phi^{(1)}(0, 0, \dots, 0) = -\delta_1 + \delta_1 = 0. \end{aligned}$$

If  $x_1 = \alpha_1$ , player 1 is cooperating, namely  $\omega_1 = 1$ , hence  $u_1^1 = 1$  and  $u_0^1 = 0$ . Then

$$\begin{aligned} \dot{x}_1 &= -\alpha_1 + \sum_{\omega_1=1} \phi^{(1)}(1, \omega_2, \dots, \omega_n) u_1^1 u_{\omega_2}^2 \dots u_{\omega_n}^n \\ &\leq -\alpha_1 + \sum_{\omega_1=1} \phi^{(1)}(1, 1, \dots, 1) [1] u_{\omega_2}^2 \dots u_{\omega_n}^n \\ &= -\alpha_1 + \phi^{(1)}(1, 1, \dots, 1) = -\alpha_1 + \alpha_1 = 0. \end{aligned}$$

The same argument can be repeated for all players and this completes the proof.  $\square$

We are now ready to introduce a result on the attractiveness of  $\mathcal{R}_{\delta, \alpha}$ .

**PROPOSITION 4.3.** *The hyper-rectangle  $\mathcal{R}_{\delta, \alpha}$  is attractive for system (3.1) with strategy (2.4), hence there cannot be equilibria in  $\mathbb{R}^n \setminus \mathcal{R}_{\delta, \alpha}$ .*  $\square$

*Proof.* Along the same lines as in the proof of Proposition 4.2, assume that the payoff for player 1 is  $\bar{x}_1 < \delta_1$ . Hence, strategy (2.4) will lead the player to defect,

since  $u_1^1 = 0$  and  $u_0^1 = 1$ , thus  $\omega_1 = 0$ . Then, for  $x_1 = \bar{x}_1$ ,

$$\begin{aligned} \dot{x}_1 &= -\bar{x}_1 + \sum_{\omega_1=0} \phi^{(1)}(0, \omega_2, \dots, \omega_n) u_0^1 u_{\omega_2}^2 \dots u_{\omega_n}^n \\ &\geq -\bar{x}_1 + \sum_{\omega_1=0} \phi^{(1)}(0, 0, \dots, 0) [1] u_{\omega_2}^2 \dots u_{\omega_n}^n = \delta_1 - \bar{x}_1 > 0. \end{aligned}$$

Since the derivative is positive,  $\bar{x}_1(t)$  converges to its limit  $\bar{x}_1(\infty)$  from below. If we assume by contradiction that  $\bar{x}_1(\infty) < \delta_1$ , then the above expression for the derivative ensures that  $\liminf_{t \rightarrow \infty} \dot{x}_1(t) \geq \delta_1 - \bar{x}_1$  for all  $\bar{x}_1 < \delta_1$ ; however, this is in contradiction with  $x_1(t) \rightarrow \bar{x}_1(\infty) < \delta_1$ .

The case  $\bar{x}_1 > \alpha_1$  can be handled similarly and, in both cases, the reasoning is the same for all players.  $\square$

Now define the  $n \times n$  matrices

$$\begin{aligned} N &= [\Phi(0, 0, \dots, 0) \ \Phi(0, 0, \dots, 0) \ \dots \ \Phi(0, 0, \dots, 0)], \\ M &= [\Phi(1, 0, \dots, 0) \ \Phi(0, 1, \dots, 0) \ \dots \ \Phi(0, 0, \dots, 1)], \\ D &= \text{diag}\{\alpha_1 - \delta_1, \alpha_2 - \delta_2, \dots, \alpha_n - \delta_n\}, \end{aligned}$$

and note that, in view of Assumption 1(a), the matrix  $[M - N]D^{-1}$  is Metzler (i.e., it has nonnegative off-diagonal entries). We remind that the dominant eigenvalue (i.e., the eigenvalue with the largest real part) of a Metzler matrix is real [12, 19].

These considerations enable us to address the following question: does there exist a *panic region*, a subset of the polytopic payoff region such that, when starting from a point inside this region, all players feel betrayed and refuse to cooperate? The panic region would hence be the domain of attraction of the non-cooperative equilibrium point within the polytopic payoff region. The following result shows that we can prevent the existence of a panic region by choosing  $p$  large enough, *provided that* matrix  $[M - N]D^{-1}$  has a positive real eigenvalue. The vanishing of the panic region for large values of  $p$  can be linked to the stability of cooperation in multiplayer social dilemmas, see, e.g., [29, p. 16428]: in fact,  $p$  can be viewed as a behavioural trait parameter that measures the altruistic nature of the players.

**THEOREM 4.4.** *Consider system (3.1) with strategy (2.4). If matrix  $[M - N]D^{-1}$  has a positive real eigenvalue, then the non-cooperative equilibrium  $\hat{x} = \Phi(0, \dots, 0)$  (i.e.,  $\hat{x}_k = \delta_k$  for all  $k$ ) is unstable for  $p$  large enough. Conversely, if  $[M - N]D^{-1}$  has no real positive eigenvalues, then, for any choice of  $p$ , the non-cooperative equilibrium is locally attractive for all initial conditions  $x_k > \delta_k$  such that  $x_k - \delta_k \leq \epsilon$ , for  $\epsilon > 0$  small enough.  $\square$*

*Proof.* If  $\hat{x}_k = \delta_k$ , then  $\sigma_k(\hat{x}_k) = 1$  and  $u_1^k = 0$ ,  $u_0^k = 1$  for all  $k$ . Hence, the system equation becomes  $\dot{x} = -\hat{x} + \Phi(0, 0, \dots, 0) = 0$ , which means that  $\hat{x}$  is an equilibrium point.

To assess stability, since the saturation function compromises differentiability at  $\hat{x}$  (this was not the case for the cooperative point  $\bar{x}$  with  $p \geq 2$ ), we need to consider the right limit of the Jacobian at the non-cooperative equilibrium (namely, the limit for  $x_k \rightarrow \hat{x}_k^+$ ), which is:  $J = -I + p[M - N]D^{-1}$ .

This expression of the Jacobian is correct if we neglect saturation in the functions  $\sigma_k(x_k)$ . In this case, evaluating the functions without saturation at the point  $\hat{x}$  gives

$$\left( \frac{\alpha_k - x_k}{\alpha_k - \delta_k} \right)^p \Big|_{x_k = \delta_k} = 1,$$

for any  $p$ . The corresponding derivative is

$$\left. \frac{d}{dx_k} \left( \frac{\alpha_k - x_k}{\alpha_k - \delta_k} \right)^p \right|_{x_k = \delta_k} = \frac{-p}{\alpha_k - \delta_k}.$$

Now consider the derivative of the term  $v$  in (3.3). At the considered point,  $u_0^k = \sigma_k = 1$  and  $u_1^k = 1 - \sigma_k = 0$ . Therefore, only the terms  $\prod_{i=1}^n \sigma_i$  and  $(1 - \sigma_k) \prod_{i=1, i \neq k}^n \sigma_i$ , for  $k = 1, \dots, n$ , yield a nonzero derivative. The derivative of  $\prod_{i=1}^n \sigma_i$  with respect to  $x_k$ , computed at the non-cooperative equilibrium, is given by

$$\left. \frac{-p}{\alpha_k - \delta_k} \prod_{i=1, i \neq k}^n \sigma_i \right|_{x = \hat{x}} = \frac{-p}{\alpha_k - \delta_k},$$

and is associated with the column vector  $\Phi(0, 0, \dots, 0)$  (which explains the structure of matrix  $N$ ). For terms of the type  $(1 - \sigma_k) \prod_{i=1, i \neq k}^n \sigma_i$ , the derivative with respect to  $x_k$ , computed at the non-cooperative equilibrium, is

$$\left. \frac{p}{\alpha_k - \delta_k} \prod_{i=1, i \neq k}^n \sigma_i \right|_{x = \hat{x}} = \frac{p}{\alpha_k - \delta_k}$$

and is associated with the term  $\Phi(0, \dots, 0, 1, 0, \dots, 0)$ , with the 1 in the  $k$ th position (which explains the structure of matrix  $M$ ), while the other derivatives are zero, because  $(1 - \sigma_k)$  is zero at the non-cooperative equilibrium.

Under differentiability assumptions, the proof would be completed. Indeed, since matrix  $p[M - N]D^{-1}$  is Metzler, the Jacobian  $J = -I + p[M - N]D^{-1}$  is Metzler as well and has eigenvalues  $\lambda_i = -1 + p\theta_i$ , where  $\theta_i$  are the eigenvalues of matrix  $[M - N]D^{-1}$ . If there is a positive  $\theta_i$ , then  $J$  has a positive eigenvalue for  $p$  large enough, hence the equilibrium is unstable for such a choice of  $p$ .

Yet, to complete the proof, we need to address the issue of the non-differentiability of the function at  $\hat{x}$ . Note that the off-diagonal entries of  $J$  are strictly positive. Assume that  $[M - N]D^{-1}$  has a positive real eigenvalue  $\theta$ . Let  $\eta$  be the left Frobenius eigenvector associated with the positive Frobenius eigenvalue  $\lambda = -1 + p\theta > 0$  of  $J$ . Then, the inequality  $\eta > 0$  is satisfied componentwise and  $\eta^\top J = \lambda \eta^\top$ .

By changing the variables as  $z_k = x_k - \delta_k$ , we show that there exists a simplex of the form  $\mathcal{S}_\epsilon = \{z : z_k \geq 0 \text{ and } \eta^\top z \leq \epsilon\}$  such that any solution starting in the interior of  $\mathcal{S}_\epsilon$  reaches the face  $\eta^\top z = \epsilon$ . This implies that the equilibrium  $\hat{x}$  (corresponding to  $z_k = 0$  for all  $k$ ) is unstable. To this aim, let us write the system as  $\dot{z} = Jz + O(z)$ , where  $O(z)$  is an infinitesimal of order greater than 1. In view of Proposition 4.2, provided that  $\epsilon > 0$  is small enough to ensure that  $\mathcal{S}_\epsilon$  is inside the hyper-rectangle  $\mathcal{R}_{\delta, \alpha}$ , any solution starting in  $\mathcal{S}_\epsilon$  with  $z(0) > 0$  (componentwise) remains positive, because the hyper-rectangle is positively invariant. Now, consider the function  $\zeta = \eta^\top z$ . Its derivative  $\dot{\zeta} = \eta^\top \dot{z} = \eta^\top [Jz + O(z)] = \lambda \zeta + \eta^\top O(z)$  is positive in the simplex  $\mathcal{S}_\epsilon$  if  $\epsilon > 0$  is small, since the first term dominates and both  $\lambda > 0$  (by assumption) and  $\zeta = \eta^\top z > 0$  (because  $\eta$  and  $z$  are componentwise positive). Hence, the solution reaches the face  $\eta^\top z = \epsilon$ .

Conversely, if  $[M - N]D^{-1}$  has no positive real eigenvalues, then its dominant Frobenius eigenvalue is negative or zero and, for any  $p$ , all the eigenvalues of  $J$  have a negative real part. Hence,  $J$  is Hurwitz stable for all  $p$ . With the same argument as above, based on a simplex, we can prove attractivity of the non-cooperative equilibrium starting from all initial conditions  $z_k(0)$  positive and small enough.  $\square$

REMARK 3. *Even when the non-cooperative equilibrium is unstable, it admits a nontrivial region where trajectories converge to it. This is a distrust region (cf. Fig. 1), characterised by low values of the payoffs:*

$$\mathcal{D} = \{x \in \mathbb{R}^n : 0 \leq x_k \leq \delta_k \quad \forall k = 1, \dots, n\}.$$

For any initial condition in the distrust region  $\mathcal{D}$ , all the functions  $\sigma_k(x_k)$  are saturated,  $\sigma_k(x_k) = 1$ , hence no-one cooperates:  $u_1^k = 0$  and  $u_0^k = 1$ . The equations become

$$\dot{x}_k = -x_k + \phi^{(k)}(0, 0, \dots, 0) = -x_k + \delta_k,$$

hence  $x_k(t) \rightarrow \delta_k$ .

If the non-cooperative equilibrium is not stable, then an “escape region” is of the form  $x_k > \delta_k$  for all  $k$ , according to the proof of Theorem 4.4.  $\diamond$

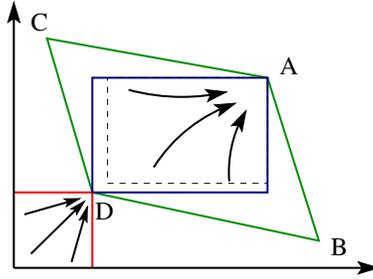


Fig. 1: The payoff polygon (green), the invariant rectangle  $\mathcal{R}_{\alpha, \delta}$  (blue), the distrust region of trajectories converging to the non-cooperative point  $D$  (red) and the region of attraction to the cooperative point  $A$  inside  $\mathcal{R}_{\alpha, \delta}$  (dashed).

**5. Dynamics in the Hyper-Rectangle  $\mathcal{R}_{\alpha, \delta}$ .** As shown in Proposition 4.3, the hyper-rectangle  $\mathcal{R}_{\alpha, \delta}$  is attractive for the system. A related issue to investigate is then the behaviour of the system when the state is inside  $\mathcal{R}_{\alpha, \delta}$ . We will prove that, for  $p$  sufficiently large, the domain of attraction to the cooperative point inside the positively invariant hyper-rectangle  $\mathcal{R}_{\alpha, \delta}$  (the dashed rectangle in Fig. 1) gets arbitrarily close to  $\mathcal{R}_{\alpha, \delta}$ . Therefore, even when the non-cooperative equilibrium is stable, its domain of attraction  $\mathcal{S}_\epsilon$  (defined in the proof Theorem 4.4) gets smaller and smaller as  $p$  increases.

To simplify the notation, we consider the change of coordinates

$$(5.1) \quad \xi_k \doteq \frac{\alpha_k - x_k}{\alpha_k - \delta_k},$$

which maps  $\mathcal{R}_{\alpha, \delta}$  into the hyper-cube

$$(5.2) \quad \mathcal{C} = \{\xi \in \mathbb{R}^n : 0 \leq \xi_k \leq 1 \quad \forall k = 1, \dots, n\}.$$

The cooperative point now corresponds to  $\xi = 0$ .

THEOREM 5.1. *Consider the scaled hyper-cube*

$$(5.3) \quad \mathcal{C}_\lambda = \{\xi \in \mathbb{R}^n : 0 \leq \xi_k \leq \lambda \quad \forall k = 1, \dots, n\}.$$

For any positive  $0 < \lambda < 1$ , there exists  $p_\lambda$  such that the state converges to the cooperative point  $\xi = 0$  for any initial condition  $\xi(0) \in \mathcal{C}_\lambda$  and for all  $p \geq p_\lambda$ .  $\square$

*Proof.* The equations (2.1), (2.3) and (2.4) in the new variable  $\xi$  become

$$\dot{\xi}_k = -\xi_k + \hat{\alpha}_k - \sum_{\omega \in \Omega} \hat{\phi}^{(k)}(\omega_1, \dots, \omega_n) \hat{u}_{\omega_1}^1 \dots \hat{u}_{\omega_n}^n,$$

where  $\hat{\alpha}_k = \frac{\alpha_k}{\alpha_k - \delta_k}$  and  $\hat{\phi}^{(k)} = \frac{\phi^{(k)}}{\alpha_k - \delta_k}$ , while

$$\begin{cases} \hat{u}_1^k = 1 - \xi_k^p, \\ \hat{u}_0^k = \xi_k^p, \end{cases}$$

since saturations are not active inside the invariant hyper-rectangle. The right-hand side is the sum of the linear term  $-\xi_k$  and a polynomial in the variables  $\xi_h^p$ :

$$\dot{\xi}_k = -\xi_k + \sum_{j_h \in \{0,1\}} \mu_{j_1, j_2, \dots, j_n} (\xi_1^p)^{j_1} (\xi_2^p)^{j_2} \dots (\xi_n^p)^{j_n}.$$

Since  $\xi = 0$  is an equilibrium, the constant term of the polynomial is zero,  $\mu_{0,0,\dots,0} = 0$ , and indeed  $\hat{\alpha}_k = \hat{\phi}^{(k)}(1, 1, \dots, 1)$ . The equation above holds for all  $\xi$  in the set  $\mathcal{C}$ , which is invariant, since it is the transformed set of  $\mathcal{R}_{\alpha,\delta}$ . To prove convergence for all initial states in  $\mathcal{C}_\lambda$ , we introduce the co-positive Lyapunov function  $V(\xi) \doteq \max_{i \in \{1, \dots, n\}} \xi_i$ . The set  $\mathcal{C}_\lambda$  is defined by  $0 \leq V(\xi) \leq \lambda$ . Define the *maximiser set* as the set of indices where the maximum is achieved:  $\mathcal{M}(\xi) = \{i : \xi_i = V(\xi)\}$ . Then, the Lyapunov derivative of  $V$  can be written in accordance with [16] as  $\dot{V}(\xi) = \max_{k \in \mathcal{M}(x)} \dot{\xi}_k$ . We show that, for  $p$  large enough,  $\dot{V}(\xi)$  is negative in  $\mathcal{C}_\lambda$  for any  $\lambda < 1$ : if  $\xi \in \mathcal{C}_\lambda$  and  $k \in \mathcal{M}(\xi)$ , then

$$\begin{aligned} \dot{\xi}_k &= -\xi_k + \sum_{j_h \in \{0,1\}} \mu_{j_1, j_2, \dots, j_n}^{(k)} \xi_1^{j_1 p} \xi_2^{j_2 p} \dots \xi_n^{j_n p} \\ &\leq -\xi_k + \sum_{j_h \in \{0,1\}} |\mu_{j_1, j_2, \dots, j_n}^{(k)}| \xi_1^{j_1 p} \xi_2^{j_2 p} \dots \xi_n^{j_n p} \\ &\leq -\xi_k + \sum_{j_h \in \{0,1\}} |\mu_{j_1, j_2, \dots, j_n}^{(k)}| \xi_k^{j_1 p} \xi_k^{j_2 p} \dots \xi_k^{j_n p} \\ (5.4) \quad &\leq -\xi_k + \sum_{h=1,2,\dots,n} \nu_h^{(k)} \xi_k^{hp} < 0, \end{aligned}$$

for some positive numbers  $\nu_h^{(k)}$ . Given  $\lambda < 1$ , we can always find  $p_\lambda$  such that, for all  $k$  and all  $p \geq p_\lambda$ , the polynomial in (5.4) is negative in the set  $\mathcal{C}_\lambda$ , where  $0 \leq \xi_k \leq \lambda$ , because the dominant linear term is negative. Hence, for  $p \geq p_\lambda$ , the Lyapunov derivative of  $V$  is negative for  $V(\xi) \leq \lambda$ ,  $V(\xi) \neq 0$ , as long as  $\xi \in \mathcal{C}_\lambda$ . Since the hyper-cube  $\mathcal{C}$  is invariant, the proof is completed.  $\square$

**REMARK 4.** *When  $p$  tends to infinity, it squashes down the function  $\sigma_k(x_k)$  to zero for all  $x_k \in (\delta_k, \alpha_k]$ , which means that the player  $k$  is numb to non-cooperation of others and cooperates. This result could be obtained alternatively by adopting the trivial open-loop strategy  $\sigma_k(x_k) \equiv 0$  for all  $x_k \in (\delta_k, \alpha_k]$ . However, an open-loop strategy is not satisfactory because it makes the players blind to defection of others: adopting a feedback strategy is fundamental, because it allows the players to monitor*

the actions of others (even though indirectly, by monitoring the evolution of their own payoffs) and react to betrayals. For this reason, we are particularly interested in finite values of  $p$ : it is always possible to find a finite value of  $p$  that ensures convergence to the cooperative point, and at the same time keeps all players sufficiently aware of the context, and able to react if needed. Indeed, the proposed strategy is conceived so that each player, exclusively based on its own payoff, can enforce a retaliation against betrayal. This crucial aspect is analysed in the next section.

We conclude the section pointing out that, instead of the considered strategy, we could consider the more general class of sigmoidal functions, which (a) are continuous, and strictly decreasing in the interval  $(\delta_k, \alpha_k)$ , (b) have a single inflexion point, (c) satisfy  $\sigma(\alpha_k) = 0$ ,  $\sigma(\delta_k) = 1$ ,  $\sigma'(\alpha_k) = \sigma'(\delta_k) = 0$ , (d) are constant outside  $(\delta_k, \alpha_k)$ , respectively equal to 0 for  $x_k \geq \alpha_k$  and to 1 for  $x_k \leq \delta_k$ . For instance,  $\sigma(\xi_k) = \xi_k [1 - (1 - \xi_k)^q] + (1 - \xi_k) \xi_k^p$ , with  $\xi_k$  defined as in (5.1) and  $q$  and  $p$  positive integers, could be one of such functions. Remarkably, in the case of a sigmoidal function we would have no hope of destabilising the non-cooperative point. Indeed, since the derivative at the non-cooperative point (corresponding to  $\xi_k = 1$  for all  $k$ ) is zero,  $\sigma'(1) = 0$ , and the linear part of the system is stable, we would always have local stability of the non-cooperative equilibrium. Therefore, by properly choosing  $p$  and  $q$ , we could at most design a sigmoidal function so that, inside the rectangle  $\mathcal{R}_{\alpha, \delta}$ , the domain of attraction of the cooperative point is larger than the domain of attraction of the non-cooperative point.

**6. A Good Strategy?** It is interesting to assess whether the proposed strategy is good, according to the next definition.

DEFINITION 6.1. *The saturated strategy in (2.4) is good if no player benefits from abandoning the strategy and choosing  $u_0^k$  (steady non-cooperation).*  $\diamond$

We assume that, if at least one player does not cooperate, then all other players benefit from full non-cooperation.

ASSUMPTION 2. *If  $\omega_i = 0$  for some indices  $i \in \mathcal{I}$ , then*

$$\phi^{(k)}(0, 0, \dots, 0) > \phi^{(k)}(\omega_1, \dots, \underbrace{0}_{=\omega_i}, \dots, \omega_n) \quad \text{for all } k \notin \mathcal{I}. \quad \diamond$$

THEOREM 6.2. *Given system (3.1) under Assumption 2, if one of the players defects and the others adopt the strategy (2.4), then for any choice of  $p$  the system converges to the non-cooperative equilibrium:  $x_k = \delta_k$  for all  $k$ .*  $\square$

*Proof.* If player 1 defects, then, in view of Corollary 3.2,  $x(t)$  converges to the convex hull of the points  $\Phi(0, \omega_2, \dots, \omega_n)$ . If player  $k$  does not defect, due to Assumption 2, her/his payoff  $x_k(t)$  decreases to values below  $\delta_k$ :  $\limsup_{t \rightarrow \infty} x_k(t) \leq \phi^{(k)}(0, 0, \dots, 0) = \delta_k$ . On the other hand, when  $x_k \leq \delta_k$ , the strategy saturates to  $\sigma_k(\delta_k) = 1$  for any  $p$ . Therefore, asymptotically,  $\dot{x} = -x + \phi^{(k)}(0, 0, \dots, 0)$ , which yields  $x_k(t) \rightarrow \delta_k$  for all  $k$ .  $\square$

Hence, the saturated strategy (2.4) is good. Remarkably, the players have no information beyond their own current payoff  $x_k$  and their own  $\alpha_k$  and  $\delta_k$ : their reaction against the defector relies on this knowledge only.

**7. A Population-Game Perspective.** Consider a random player  $i$  playing against a random player  $j$  and let player  $i$ 's payoff matrix be

$$(7.1) \quad F = \begin{bmatrix} \alpha & \gamma \\ \beta & \delta \end{bmatrix}.$$

The above entries represent the payoff of player  $i$  when both players cooperate ( $\alpha$ ), player  $i$  only defects ( $\beta$ ), player  $j$  only defects ( $\gamma$ ) and both players defect ( $\delta$ ). Player  $i$  chooses the row, while player  $j$  chooses the column. The matrix of player  $j$  is symmetric and can be ignored. Assumption 1 leads to the payoff ordering

$$(7.2) \quad \beta > \alpha > \delta > \gamma.$$

Let the distribution of cooperators and defectors in the population be  $u_1(x)$  and  $u_0(x)$ , respectively, where  $u_1 + u_0 = 1$ , with  $u_1, u_0 \geq 0$ , and consider the mixed strategy

$$u = u(x) = \begin{bmatrix} u_1(x) \\ u_0(x) \end{bmatrix}.$$

The average payoff across the population at time  $t$  is then  $u(x(t))^\top F u(x(t))$ . Also, the average payoff across the population and across time, namely from 0 to  $t$ , is

$$(7.3) \quad \dot{x}(t) = -x(t) + u(x(t))^\top F u(x(t)).$$

Indeed, consider the time-average expected (over opponent's play) payoff defined as

$$\Gamma(s) = \frac{1}{s} \int_0^s (u^\top F v) d\tau \in \mathbb{R}.$$

If we rescale the time window using  $s = e^t$ , take  $x(t) = \Gamma(e^t)$  and differentiate with respect to  $t$ , we obtain the differential equation (7.3). Note that, after rescaling the time window, we have

$$x(0) = \int_0^1 (u^\top F v) d\tau \in \mathbb{R}.$$

Equation (7.3) is in the same spirit as in Hart and Mas-Colell's paper [25] on continuous-time approachability (with the calculation of the average payoff analogous to the one in (2.1)). Adopting the strategy (2.4), system (7.3) becomes

$$\dot{x} = -x + \alpha + (\beta + \gamma - 2\alpha)\sigma(x) + (\alpha + \delta - \beta - \gamma)\sigma(x)^2.$$

The result in Theorem 4.4 translates into conditions for the diffusion of cooperation in a population of individuals. To see this, consider that, because of the symmetry in the payoff matrices of the game, it holds that

$$p[M - N]D^{-1} = \frac{p}{r} \begin{bmatrix} -g & l \\ l & -g \end{bmatrix},$$

where  $r \doteq \alpha - \delta > 0$ ,  $g \doteq \delta - \gamma > 0$ , and  $l \doteq \beta - \delta > 0$ . Conditions for the existence of positive eigenvalues for  $[M - N]D^{-1}$ , then translate into conditions on the trace and determinant of the above matrix. In particular, we have positive eigenvalues if  $l > g$ . This condition means that, when a cooperator meets a defector, the damage for the cooperator is less than the benefit for the defector. Under this condition, in a population of "patient" and "farsighted" cooperators (this means that  $p$  is large enough), the cost of cooperating with defectors is relatively small, and therefore it turns out that all players end up cooperating.

The population scenario suggests a second interpretation of the results if we view  $p$  as the number of repeated plays between player  $i$  and  $p$  different opponents  $j$ ,

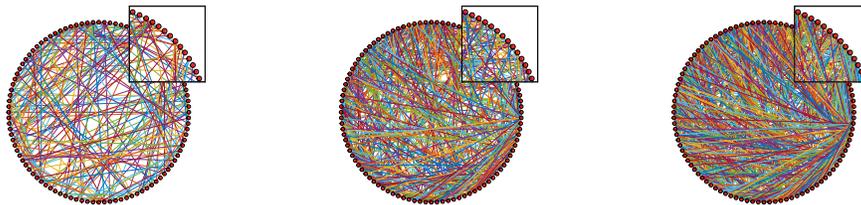


Fig. 2: Co-evolving networks for  $p = 5$  (left),  $p = 10$  (center) and  $p = 15$  (right).

assuming that, at each iteration, player  $i$  defects with probability  $\left(\frac{\alpha_k - x_k}{\alpha_k - \delta_k}\right)$ . If we view the repeated plays as  $p$  independent identically distributed experiments, the joint probability of defecting in all  $p$  plays (or equivalently, one minus the probability of cooperating at least once out of  $p$  plays) is given by

$$\underbrace{\left(\frac{\alpha_k - x_k}{\alpha_k - \delta_k}\right)}_{\text{iteration 1}} \cdot \underbrace{\left(\frac{\alpha_k - x_k}{\alpha_k - \delta_k}\right)}_{\text{iteration 2}} \cdots \underbrace{\left(\frac{\alpha_k - x_k}{\alpha_k - \delta_k}\right)}_{\text{iteration } p}.$$

From the invariance of the payoff polytope, the above joint probability is equal to the mixed strategy (2.3). The convergence to the cooperative equilibrium means that every player ends up cooperating with all of her/his  $p$  neighbour players. If we imagine a network where the nodes are the players, and we add a link  $(i, j)$  any time player  $i$  cooperates with player  $j$ , we can construct a co-evolving network which, in the case of convergence to the cooperative equilibrium, has degree  $p$  for each node. Fig. 2 depicts three different co-evolving networks for  $p = 5$ ,  $p = 10$  and  $p = 15$ . By ‘co-evolving network’ we mean that the topology of the network evolves together with the probability of cooperation. In this context, the probability of cooperation between any two players determines the probability of forming a link between the corresponding nodes. Thus, Fig. 2 displays the steady-state topology configuration when all players cooperate with probability 1, and each player has 5, 10 or 15 neighbours.

REMARK 5. *It is interesting to compare the proposed saturated strategy with replicator dynamics in evolutionary games, where the cooperation frequency in a population increases if the cooperators perform better than the defectors. In particular, the continuous-time replicator dynamics is given by*

$$\dot{x}_s(t) = x_s(t) \frac{[u(s, x(t)) - \bar{u}(x(t))]}{\bar{u}(x(t))},$$

where  $s$  is a generic strategy,  $x_s$  is the portion of the population playing that strategy,  $\bar{u}(x(t))$  is the average payoff across the population, and  $u(s, x(t))$  is the instantaneous payoff obtained by playing  $s$  in a population whose strategies are distributed according to  $x(t)$ . A first difference is that the proposed strategy is a feedback on the current payoff which is averaged over time, while in the replicator dynamics the average across the population is considered, but not the average over time. A second difference is that the dynamics (7.3) describes the evolution over time of the average payoff (over time and across the population), while the replicator dynamics captures the dynamics on the portion of cooperators and defectors: therefore, the two dynamics model different

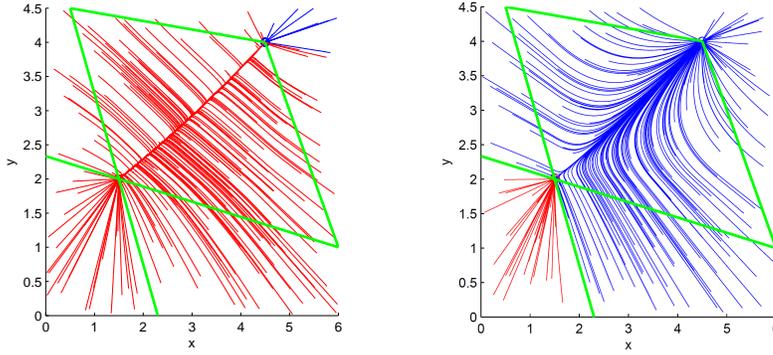


Fig. 3: Evolution of the system with payoff matrix (8.1) and strategy (2.4). Blue trajectories converge to the cooperative equilibrium, red trajectories to the non-cooperative equilibrium. Left,  $p = 1$ . Right,  $p = 3$ .

quantities (the payoff in one case, the portion of cooperators/defectors in the other). A third difference is that, in our case, the portion of cooperators increases or decreases based on the current average payoff over time and across the population, while in the replicator dynamics the instantaneous payoff obtained from cooperating is compared with the average payoff across the population. Hence, although the idea of averaging across the population is a common point, it plays a different role in the two cases.

## 8. Numerical Examples.

**8.1. A 2-player game.** Consider a 2-player game with the payoff matrix

$$(8.1) \quad \left[ \begin{array}{c|c} (\phi^{(1)}(1,1), \phi^{(2)}(1,1)) & (\phi^{(1)}(1,0), \phi^{(2)}(1,0)) \\ \hline (\phi^{(1)}(0,1), \phi^{(2)}(0,1)) & (\phi^{(1)}(0,0), \phi^{(2)}(0,0)) \end{array} \right] = \left[ \begin{array}{c|c} (4.5, 4.0) & (0.5, 4.5) \\ \hline (6.0, 1.0) & (1.5, 2.0) \end{array} \right]$$

and adopt the strategy (2.4). If  $p = 1$ , the cooperative equilibrium  $A = (\alpha_1, \alpha_2) = (4.5, 4.0)$  is unstable, while the non-cooperative equilibrium  $D = (\delta_1, \delta_2) = (1.5, 2.0)$  is asymptotically stable, as illustrated in Fig. 3, left. Note that there are trajectories that converge to the cooperative equilibrium, for some high initial values.

To destabilise the non-cooperative equilibrium  $D$ , let us derive the matrix  $\Theta = [M - N]D^{-1}$ , which in the present instance is given by

$$\Theta = \left( \left[ \begin{array}{cc} 0.5 & 6.0 \\ 4.5 & 1.0 \end{array} \right] - \left[ \begin{array}{cc} 1.5 & 1.5 \\ 2.0 & 2.0 \end{array} \right] \right) \left[ \begin{array}{cc} 3.0 & 0.0 \\ 0.0 & 2.0 \end{array} \right]^{-1}.$$

To investigate stability, let us consider the matrix

$$-I + p\Theta = \left[ \begin{array}{cc} -1 - \frac{p}{3} & \frac{2p}{3} \\ \frac{5p}{4} & -1 - \frac{p}{2} \end{array} \right],$$

which becomes unstable with a strictly positive eigenvalue when  $p > 2$ ; hence,  $D$  becomes unstable. We simulate the system for  $p = 3$ : for this value of the parameter that captures the patience of the players,  $D$  becomes unstable and, as expected, the proposed saturated strategy guarantees convergence to the cooperative equilibrium for all trajectories originating in the invariant payoff polygon having vertices  $(4.5, 4)$ ,

(0.5, 4.5), (1.5, 2) and (6, 1). Some trajectories originating from low initial payoffs for both players still converge to the non-cooperative equilibrium. This is shown in Fig. 3, right. For details on the two-player game, the reader is referred to [21].

**8.2. Channel sharing in wireless communication.** Wireless communication is an important field to which control theory can provide several contributions, for instance to control how a wireless channel is shared among several transmitters (see e.g. [14] and the references therein); the main problem is to maximise transmission while avoiding congestion problems. Consider the case of Frequency Division Multiple Access (FDMA), in which the radio frequency channel is split into several smaller sub-channels, each with its own frequency range and allocated to a single transmitter only. Problems may arise due to greedy transmitters that start using more sub-channels than the allocated one: this leads to degradation of the success rate in transmission, for all messages that are sent using a sub-channel that is used by more than one transmitter (we call this phenomenon *congestion* for simplicity). In terms of quantity of messages that are successfully sent, all transmitters have lower performance in the case of full congestion (each transmitter uses the whole bandwidth) than in the ideal situation (each transmitter uses its allocated sub-channel only). On the other hand, a single transmitter can see its performance definitely improve if it unilaterally decides to use more sub-channels, while all others use their assigned sub-channel only.

A game-theoretic approach to this problem is very popular (surveys are for instance [6, 22]). Adopting the paradigm in [6, Example 2 and Figure 4], we can distinguish between two transmission modes for each transmitter.

**0: Broadband: non-cooperation.** This selfish transmission mode (using the whole bandwidth instead of the assigned sub-channel) would benefit a single transmitter that adopts it, but would cause performance degradation for the others and rapidly cause congestion (typically due to conflicts caused by simultaneous packet transmission [14]) if adopted by several transmitters.

**1: Narrowband: cooperation.** In this transmission mode, each transmitter uses exclusively the sub-channel assigned by a network supervisor to optimise the channel. This leads to a fair and efficient transmission.

For the  $n$ -transmitter problem, all the conditions in Assumption 1 are verified:

- (a) is met because the cooperation of transmitter  $h$  is more profitable than non-cooperation for all other transmitters (there is less congestion);
- (b) is verified, because a transmitter that uses more bandwidth than the assigned sub-channel is advantaged, if the others stick to their assigned sub-channel only;
- (c) full cooperation is more profitable than full non-cooperation, because full congestion induces a severe performance drop.

The problem fits nicely in our setup. Each transmitter can be reasonably assumed to be aware exclusively of the expected performance in the fully cooperative and in the fully non-cooperative case, and of the current performance. Hence, our saturated strategy can be employed. The decision variables  $u_1^k$  and  $u_0^k = 1 - u_1^k$  represent how prone transmitter  $k$  is to transmitting in Narrowband/Broadband mode. (The strategy can be implemented also in a P-persistent transmission scheme, by suitably altering the transmission probability; see [14] for details.)

Given  $n$  equal transmitters, consider their payoff in terms of successful transmission rate (typically megabits/sec). If  $c$  transmitters cooperate and  $d = n - c$  defect, then the unitary payoff will be respectively

$$P_1 = \frac{a}{n} f(d)$$

Fraction of Narrowband transmitters	1	3/4	1/2	1/4	0
Narrowband transmitter payoff	1	0.4444	0.2857	0.2105	-
Broadband transmitter payoff	-	1.3333	0.8571	0.6316	0.5

Table 1: Payoffs for Narrowband transmitters (cooperating players) and Broadband transmitters (non-cooperating players), as a function of the fraction of Narrowband transmitters.

for the transmitters that cooperate and

$$P_0 = \frac{b}{n}f(d)$$

for the transmitters that defect, with  $b > a$ . Function  $f(d)$  represents the performance degradation due to congestion and conflicts. It is a strictly decreasing function, with  $f(0) = 1$ , so that full cooperation means an equal performance  $a/n$  for all transmitters, and with  $bf(n) \ll a$ . This means that, if all transmitters operate in Broadband mode, then the individual payoff is strongly reduced for all of them. We assume that, for all  $d = 0, 1, \dots, n - 1$ ,

$$\frac{b}{n}f(d+1) > \frac{a}{n}f(d),$$

hence any single transmitter has the interest to switch to Broadband mode (so that  $d$  increases by 1) for any fixed decision of the others. We can rewrite the condition as

$$\frac{b}{a} > \frac{f(d)}{f(d+1)}.$$

REMARK 6. *The knowledge of the function  $f$  is not required to implement the saturated strategy. All information needed by the transmitters is given by the fully cooperative payoff  $\alpha_i = \frac{a}{n}$ , the fully non-cooperative payoff  $\delta_i = \frac{b}{n}f(n)$  and the coefficient  $p$  of the control function.*

We now provide a simulation of the system evolution with the saturated strategy in the case of  $n = 4$  transmitters, with  $a = 4$ ,  $b = 12$  and

$$f(d) = \frac{n}{n + 5d},$$

leading to  $\alpha_i = 1$  and  $\delta_i = 0.5$ . Table 1 shows the payoffs for both Narrowband and Broadband transmitters as a function of the fraction of Narrowband transmitters.

When the transmitters decide their behaviour based on the saturated strategy with  $p = 5$ , their payoff evolution is shown in Fig. 4. Starting from the randomly taken initial condition  $x(0) = [1.4449 \ 0.2997 \ 1.3192 \ 1.0372]^\top$ , as expected, the four transmitters converge to the cooperative point  $\bar{x}_\alpha = [1 \ 1 \ 1 \ 1]^\top$ . However, at time  $t = 10$  the first transmitter switches to Broadband mode and starts defecting,  $u_0^1 = 1$ . Initially it has an advantage. Then, all the other transmitters detect a reduction of their performance and start switching to Broadband mode as well. As a consequence, the payoffs converge to the non-cooperative equilibrium  $\bar{x}_\delta = [0.5 \ 0.5 \ 0.5 \ 0.5]^\top$  and all transmitters are penalised, including the “traitor”.

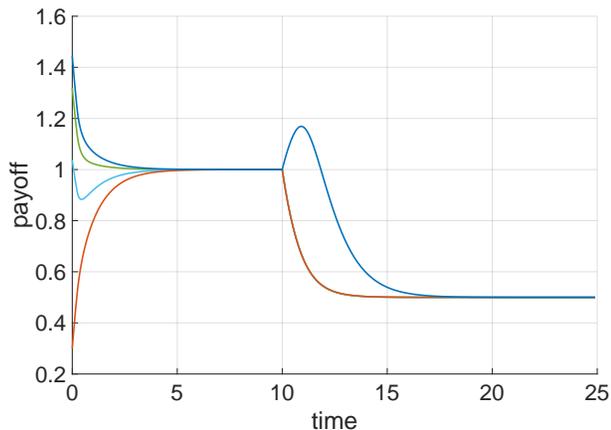


Fig. 4: Evolution of the wireless problem with  $n = 4$  and  $p = 5$ . The system initially converges to the cooperative point  $\bar{x}_\alpha = [1 \ 1 \ 1 \ 1]^\top$ . At time  $t = 10$ , player 1 starts defecting ( $w_0^1 = 1$ ) and the system eventually converges to the non-cooperative point  $\bar{x}_\delta = [0.5 \ 0.5 \ 0.5 \ 0.5]^\top$ .

**9. Concluding Discussion.** We have studied diffusion of cooperation among selfish players in a problem modelled as a repeated  $n$ -player Prisoner's dilemma in continuous time. For the same model, we have provided an interpretation as an  $n$ -population game involving at each time  $n$  random individuals, one from each population. We have shown that the game can be turned into a positive dynamical system and we have proposed a parametrised strategy that enforces cooperation for suitable values of the parameter. The strategy is original, as it builds on minimum knowledge on the part of the players, and has an explicit expression that allows us to provide convergence proofs, exploiting Lyapunov methods and Metzler matrix theory.

Future work includes: (i) extending the analysis to heterogeneous populations, which use the strategy with different parameters (different level of patience or farsightedness); (ii) studying the co-evolving network and its stability, as well as connections with random networks, scale free networks, small world networks; (iii) investigating different scenarios that involve neutral players in addition to cooperators and defectors, and where players can transition across the three categories, in accordance to controlled or uncontrolled transition rates.

Results from the above settings can be then applied to different domains, including i) cyber-security with distributed detection and counter-attack strategies, ii) demand side management in which the consumers (players) adopt different levels of responsiveness to cooperation, and iii) coalitional aggregation of power producers, in which the producers (players) may opt to cooperate and join coalitions; they can then sign contracts with an aggregator that is responsible of managing production schedule and cash flow from and to the producers.

#### REFERENCES

- [1] J. P. AUBIN, *Viability Theory*, Birkhäuser, 1991.
- [2] J. P. AUBIN AND A. CELLINA, *Differential Inclusions: Set-Valued Maps and Viability Theory*,

- Springer, 1991.
- [3] J. P. AUBIN AND H. FRANKOWSKA, *Set-Valued Analysis*, Birkhäuser, 1990.
  - [4] R. J. AUMANN AND M. B. MASCHLER, *Repeated Games with Incomplete Information*, MIT Press, 1995.
  - [5] J. BERGIN AND W. B. MACLEOD, *Continuous time repeated games*, *International Economic Review*, 34(1):21–37, 1993.
  - [6] G. BACCI, S. LASAULCE, W. SAAD AND L. SANGUINETTI, *Game Theory for Networks: A tutorial on game-theoretic tools for emerging signal processing applications*, *IEEE Signal Proc. Mag.*, 33(1):94–119, 2016.
  - [7] F. BAGAGIOLO AND D. BAUSO, *Objective function design for robust optimality of linear control under state-constraints and uncertainty*, *ESAIM Control Optim. Calc. Var.*, 17:155–177, 2011.
  - [8] D. BAUSO, E. LEHRER, E. SOLAN, X. VENEL, *Attainability in repeated games with vector payoffs*, *INFORMS Math. Operations Res.*, 40(3):739–755, 2015.
  - [9] K. BEHRSTOCK, M. BENAÏM AND M. W. HIRSCH, *Smale strategies for network prisoner’s dilemma games*, *J. Dyn. Games*, 2:141–155, 2015.
  - [10] M. BENAÏM, J. HOFBAUER, AND S. SORIN, *Stochastic approximations and differential inclusions*, *SIAM J. Control Optim.*, 44(1):338–348, 2005.
  - [11] M. BENAÏM AND M. W. HIRSCH, *Asymptotic pseudotrajectories and chain recurrent flows, with applications*, *J. Dynam. Differential Equations*, 8:141–176, 1996.
  - [12] A. BERMAN AND R. J. PLEMMONS, *Nonnegative matrices in the mathematical sciences*, *Classics in Applied Mathematics*, SIAM Press, Philadelphia (PA), 1994.
  - [13] D. BLACKWELL, *An analog of the minimax theorem for vector payoffs*, *Pacific J. Math.*, 6(1):1–8, 1956.
  - [14] F. BLANCHINI, D. CASAGRANDE, G. GIORDANO, P. L. MONTESSORO, *A robust decentralized control for channel sharing communication*, *IEEE Trans. Control Netw. Syst.*, 4(2):336–346, 2017.
  - [15] F. BLANCHINI, *Set invariance in control: a survey*, *Automatica*, 35(11):1747–1768, 1999.
  - [16] F. BLANCHINI AND S. MIANI, *Set-theoretic methods in control—Second edition*, ser. *Systems & Control: Foundations & Applications*. Boston: Birkhäuser, 2015.
  - [17] N. CESA-BIANCHI AND G. LUGOSI, *Prediction, Learning and Games*, Cambridge University Press, 2006.
  - [18] N. J. ELLIOT AND N.J. KALTON, *The existence of value in differential games of pursuit and evasion*, *J. Differential Equations*, 12:504–523, 1972.
  - [19] L. FARINA, S. RINALDI, *Positive Linear Systems; Theory and Appl.*, J. Wiley & Sons, 2000.
  - [20] D. FOSTER, R. VOHRA, *Regret in the on-line decision problem*, *Games and Economic Behavior*, 29:7–35, 1999.
  - [21] G. GIORDANO, D. BAUSO AND F. BLANCHINI, *A saturated strategy robustly ensures stability of the cooperative equilibrium for Prisoner’s dilemma*, in *Proc. 55th IEEE Conf. Dec. Control*, pp. 4427–4432, 2016.
  - [22] Z. HAN, D. NIYATO, W. SAAD, T. BASAR AND A. HJRUNGNES, *Game Theory in Wireless and Communication Networks: Theory, Models, and Applications* (1st ed.), Cambridge University Press, New York, NY, USA, 2012.
  - [23] S. HART, *Adaptive heuristics*, *Econometrica*, 73:1401–1430, 2005.
  - [24] S. HART AND A. MAS-COLELL, *A general class of adaptive strategies*, *J. Econom. Theory*, 98:26–54, 2001.
  - [25] S. HART AND A. MAS-COLELL, *Regret-based continuous-time dynamics*, *Games Econom. Behav.*, 45:375–394, 2003.
  - [26] C. HILBE, M. A. NOWAK, AND K. SIGMUND, *Evolution of extortion in Iterated Prisoner’s Dilemma games*, *PNAS*, 110(17):6913–6918, 2013.
  - [27] C. HILBE, M. A. NOWAK, AND A. TRAUlsen, *Adaptive dynamics of extortion and compliance*, *PLOS One*, 8(11):e77886, 2013.
  - [28] C. HILBE, T. RÖHL, AND M. MILINSKI, *Extortion subdues human players but is finally punished in the prisoner’s dilemma*, *Nature Communications*, 5:3976, 2014.
  - [29] C. HILBE, B. WU, A. TRAUlsen, AND M. A. NOWAK, *Cooperation and control in multiplayer social dilemmas*, *PNAS*, 111(46):16425–16430, 2014.
  - [30] C. HILBE, A. TRAUlsen, AND K. SIGMUND, *Partners or rivals? Strategies for the iterated prisoner’s dilemma*, *Games Econom. Behav.*, 92:41–52, 2015.
  - [31] C. HILBE, B. WU, A. TRAUlsen, AND M. A. NOWAK, *Evolutionary performance of zero-determinant strategies in multiplayer games*, *Journal of Theoretical Biology*, 374:115–124, 2015.
  - [32] S. HUCK, H.-T. NORMANN, J. OECHSSLER, *GLAD: a simple adaptive strategy that yields coop-*

- eration in dilemma games*, Physica D, 200:133–138, 2005.
- [33] D. KAHNEMAN AND A. TVERSKY, *Prospect theory: An analysis of decision under risk*, Econometrica, 47: 263–291, 1979.
  - [34] E. LEHRER, *Allocation processes in cooperative games*, Internat. J. Game Theory, 31:341–351, 2002.
  - [35] E. LEHRER, *Approachability in infinite dimensional spaces*, Internat. J. Game Theory, 31(2):253–268, 2002.
  - [36] E. LEHRER, *A wide range no-regret theorem*, Games Econom. Behav., 42, 2003.
  - [37] E. LEHRER AND E. SOLAN, *Excludability and bounded computational capacity strategies*, Mat. Oper. Res., 31(3):637–648, 2006.
  - [38] E. LEHRER, E. SOLAN, AND D. BAUSO, *Repeated games over networks with vector payoffs: the notion of attainability*, in Proc. NetGCoop, 2011.
  - [39] E. LEHRER, S. SORIN, *Minmax via differential inclusion*, Convex Analysis, 14(2):271–273, 2007.
  - [40] D. G. LUENBERGER, *Optimization by vector space methods*, John Wiley & Sons Inc., New York, New York, USA, 1969.
  - [41] A. NEYMAN, *Continuous-time stochastic games*, Games and Economic Behavior, 104:92–130, 2017.
  - [42] E. ROXIN, *The axiomatic approach in differential games*, J. Optim. Th. Appl., 3:153–163, 1969.
  - [43] W. H. SANDHOLM, *Population Games and Evolutionary Dynamics*, MIT Press, 2010.
  - [44] S. SMALE, *The prisoner's dilemma and dynamical systems associated to non-cooperative games*, Econometrica, 48:1617–1634, 1980.
  - [45] A. S. SOULAIMANI, M. QUINCAMPOIX, S. SORIN, *Approachability theory, discriminating domain and differential games*, SIAM J. Control Optim., 48(4):2461–2479, 2009.
  - [46] P. VARAIYA, *The existence of solution to a differential game*, SIAM J. Control Optim., 5:153–162, 1967.
  - [47] N. VIEILLE, *Weak approachability*, Math. Oper. Res., 17:781–791, 1992.
  - [48] J. VON NEUMANN, *Zur Theorie der Gesellschaftsspiele*, Math. Annalen, 100:295–320, 1928.