Dynamical Game Theory

Alberto Antonioni,[1, *] Luis A. Martinez-Vaquero,[2, †] Nicholas Mathis,[3, ‡] Leto Peel,[4, §] and Massimo Stella[5, ¶]

[1]*Faculty of Business and Economics, University of Lausanne, Switzerland*
[2]*Artificial Intelligence Lab, Vrije Universiteit, Brussels, Belgium*
[3]*Department of Physics, Arizona State University, Tempe AZ, USA*
[4]*Department of Computer Science, University of Colorado, Boulder CO, USA*
[5]*Institute for Complex Systems Simulation, Southampton, UK*

In this work we introduce the new approach of Dynamical Game Theory (DGT). In this approach, individuals perceive their payoffs differently as a function of their previous gain history and intrinsic characteristics. We present the dynamical donation game as a first case in this framework and analyze its corresponding dynamics through analytical and numerical approximations. We find that a high level of cooperation can be achieved and maintained when the dynamical donation game is played. The introduction of dynamical games opens new horizons in the explanation of the evolution of cooperation in social environments.

INTRODUCTION

The emergence and the evolution of cooperation is doubtlessly one of the most intriguing open question in real-world society [1–3]. Theoretical work predicts that uncooperative behavior prevails and may easily dominate populations of rational individuals. However, many kinds of mechanisms have been proposed to explain cooperation in both animal and human societies [4–7]. From the perspective of game theory, it is generally assumed that individuals act rationally maximizing their expected payoff [8, 9] and, while mutual cooperation would be socially preferable, its accomplishment is usually difficult to achieve. Nevertheless, real life experience and multiple laboratory experiments suggest that individuals do not always follow this requirement and non-negligible amount of cooperative behaviors can be frequently observed [5, 6, 10–12].

Among the theories of cooperation, direct reciprocity has received a lot of attention, in particular from the theoretical viewpoint. In fact, reciprocity is a kind of cooperation that is far from trivial to be explained and, being such a hard challenge, it requires an important amount of work [13]. Moreover, reciprocity is one of the most suggested mechanisms that supports human cooperation [2], as reciprocity appears to be an unavoidable consequence of small group size and given the cognitive abilities of human beings [14]. Thus, the concept of direct reciprocity has been investigated in many contexts and relevant contributions can be found in [15–18].

In this paper we introduce a new type of class of games based on direct reciprocity, i.e. the *dynamical donation game*. In this framework, the perceived payoffs may change conditioned on the current state of the interacting players. We assume that this state co-evolves as a function of the cumulative payoff that the players receive during multiple one-shot interactions. This assumption can be realistic for a variety of social phenomena where individuals beliefs, social status or consciousness, can be modified as a result of their actions in the past. We have been able to show analytically and through numerical simulations how different mechanisms in the update of players' states can influence the emergence and the sustainability of high levels of cooperation. This study on the evolution of individual perceptions allowed us to investigate complex patterns of behaviors.

THE MODEL

In this section we introduce and define the dynamics of a particular class of dynamical games.

Dynamical Donation Game

Whenever two individuals encounter themselves in a Prisoner's Dilemma (PD) situation [19], they can opt for two actions: cooperation (C) and defection (D). As a result of their choices, each individual obtains a payoff. It is usually assumed that the payoff, R, that players obtain if both cooperate is higher than the payoff obtained if both defect, P. If only one player cooperates and the other defects, the defector's payoff, T, is greater than R and the cooperator's payoff, S, is smaller than P. Thus, the game is defined by the following payoff matrix:

	C	D
C	R	S
D	T	P

where inequalities $T > R > P > S$ hold. In this framework, defection strictly dominates cooperation [20] and it is always the best rational individual choice in the one-shot game, so that (D,D) is the unique Nash Equilibrium

* alberto.antonioni@unil.ch
† l.martinez.vaquero@gmail.com
‡ cole.mathis@asu.edu
§ leto.peel@colorado.edu
¶ massimo.stella@inbox.com

(NE) [21] and also the only fixed point of the replicator dynamics [8].

A particular scenario of the PD game is the so-called *donation game* [22], where each player can cooperate by providing a benefit, b, to her co-player at her cost, c. Here we introduce a modified version of this game that we call the Dynamical Donation Game (DDG) where the perception of the payoffs by the players changes according to their state. This game incorporates the novel feature that players may have different perceptions of cost as a function of their current state. Practically, we assume that players that have a distorted perception of the game in considering the cost c. For instance, they may interpret the game differently according to their beliefs, social status, consciousness, etc. Thus, we introduce the variable $0 \leq \alpha \leq 1$ to represent this players perceptive state. This value only influences the perceived cost such that when $\alpha = 1$ the focal player does not perceive any cost. On the other hand, when $\alpha = 0$ the cost is fully considered and the standard donation game is recovered. A crucial feature of the game is proposed in the analytical model section and it is the definition of the α-function according to players interactions during the repeated DDG. Table I shows the payoff matrix for this game:

	C	D
C	$rb - c(1 - \alpha)$	$b - c(1 - \alpha)$
D	b	0

TABLE I. Payoff matrix for the Dynamical Donation game.

First, we consider the version of the game when $\alpha = 0$. We then consider that a player who cooperates incurs in a cost $c > 0$ in order to provide a benefit $b > 0$ to her co-player and herself and that inequality $b < c$ holds. This last assumption is reasonable because the cost c is to be interpreted as the effort a player would sustain in order to to give a benefit b to both players. We also assume that this benefit is cumulative when both players cooperate and it is expressed as rb. The parameter r controls the importance of mutual cooperation and it is considered greater than one to guarantee that (C,C) situation is preferable for a cooperator than (C,D). Moreover, the inequality $r < \frac{b+c}{b}$ holds in order to tempt players towards defection. Finally, if both players defect they do not receive any benefit neither incur in any cost. According to previous assumptions, the DDG for $\alpha = 0$ can be considered as equivalent of a PD game and thus defection strictly dominates cooperation for all possible values of r, b and c. On the other hand, when $\alpha = 1$, the game is less dilemmatic and becomes analogous to the Harmony game. In fact, since $r > 1$, full cooperation is the only rational choice [8].

The α-function

In this study we consider that a population of players where individuals are matched and interact pairwise among themselves following the one-shot DDG. At each interaction the perception of the game by players is related to the their α-value. There are several possibilities to define an update function for $\alpha_{i,t+1}$ according to the payoff player i receives at time t and to player i's state, $\alpha_{i,t}$ at time t. Here we consider that the α-function is related to the amount of benefit B_C $(B_D) \in \{0, b, rb\}$ an average cooperator (defector) player receives while playing DDG. The α-function can be expressed for cooperators, α_C, and for defectors, α_D, considering the following differential equation:

$$\dot{\alpha}_X = (1 - \alpha_X)k\frac{B_X}{rb} - d\alpha_X \qquad (1)$$

where X stands for C or D. Moreover, d is a constant parameter that acts as a death rate in the payoffs perception, while k is a fixed value that indicates how efficiently the benefit B_X is affecting the player. Both k and d are positive numbers. Without loss of generality, we refer to α_X as simply α. The α-function is a convex combination in the variable α that is composed by a negative coefficient for α and a positive coefficient for $(1-\alpha)$. Variables B_C and B_D are normalized by the maximum amount of benefit, rb, a player can obtain in a single interaction and they are multiplied by the scalar factor k. We can see that these functions are bounded between 0 and 1. The lower the α the faster this level is increased for a given amount of benefit received – as we said before, the α-level measures how much the perception of the game by the players is affected.

ANALYTICAL STUDY

Let us now assume an infinite, well-mixed and homogeneous population. In other words, an infinite set of individuals who share the same d and k, and that are repeatedly randomly matched. In this scenario, we can substitute the amount of benefit received at each round by each individual by the average amount in the population. Considering that there is a fraction x of cooperators and $1-x$ of defectors, one can easily deduce from Table I that the average amount of benefit received per round by cooperators is $B_C = bx(r - 1) + b$ and by defectors is $B_D = bx$. Then, according to Eq. 1, the value of α for cooperators and defectors is, respectively:

$$\dot{\alpha}_C = (1 - \alpha_C)k\frac{x(r - 1) + 1}{rb} - d\alpha_C \qquad (2)$$

$$\dot{\alpha}_D = (1 - \alpha_D)k\frac{x}{rb} - d\alpha_D \qquad (3)$$

Concerning strategy dynamics, we assume that individuals imitate each other following the replicator dynamic equation [8] and follow with high probability strategy with higher payoff. Also:

$$\dot{x} = x(1-x)(\Pi_C - \Pi_D) \qquad (4)$$

where Π_C (Π_D) is the average payoff of cooperators (defectors) in the population. We can then estimate:

$$\Delta\Pi = \Pi_C - \Pi_D = b(rx + 1 - 2x) - c(1 - \alpha_C) \qquad (5)$$

From Eqs. 4 and 5 it follows:

$$\dot{x} = x(1-x)[\, b(rx + 1 - 2x) - c(1 - \alpha_C)\,] \qquad (6)$$

Equations 2, 3 and 6 describe the dynamics of the system. In concrete, Eqs. 2 and 6 are coupled and need to be solved together, whereas Eq. 3 can be solved after the previous ones. The equilibria in the stationary regimen are found when $\dot{x} = \dot{\alpha}_C = \dot{\alpha}_D = 0$. Considering Eq. 2, 3 and 6 we can thus obtain:

1. $x = 0$, $\alpha_C = 0$, and $\alpha_D = 0$, that is, the origin of the 3-dimensional variable space.

2. $x = 0$, $\alpha_C = k/(k + rbd)$, and $\alpha_D = 0$.

3. $x = 1$, $\alpha_C = k/(k + bd)$, and $\alpha_D = k/(k + rbd)$.

4. $x = x_p$, $\alpha_C = \alpha_{C,p}$, and $\alpha_D = \alpha_{D,p}$. This internal equilibrium is considered for $x_p, \alpha_{C,p}, \alpha_{D,p} \in (0,1)$ and it only exists for particular values of the model parameters.

Since all the parameters have been considered as positive values we deduce that first three equilibria always exist, whereas the fourth one appears only for particular conditions. Figure 1 shows the values of this last equilibrium as functions of model parameters d and k and in the case of $b = 1$, $c = 2$, and $r = 2$. We can observe that this equilibrium does not always exist since some curves are not defined in the range $(0,1)$ for the entire d-axis. Moreover, the parameter d negatively affects all values of the fourth equilibrium, the higher is d the lower is the achieved cooperation level. This is also reasonable since the d parameter acts as a death rate for both α-values. On the other hand, when k increases one can observe a positive increment for all equilibrium values. In particular, when k is large enough and d is not too small, it is likely to expect that this internal equilibrium does exist.

Considering the standard linear stability analysis for a system of differential equations [23], we can see from Eqs. 2, 3 and 6 under which conditions each equilibrium is to be considered stable. The first equilibrium, i.e. the origin, is always stable since $b < c$ and all model parameters being positive. The second equilibrium is stable when the condition $b + \frac{k}{k + rbd} < c$ is satisfied, while the third equilibrium is stable when it holds the following inequality:
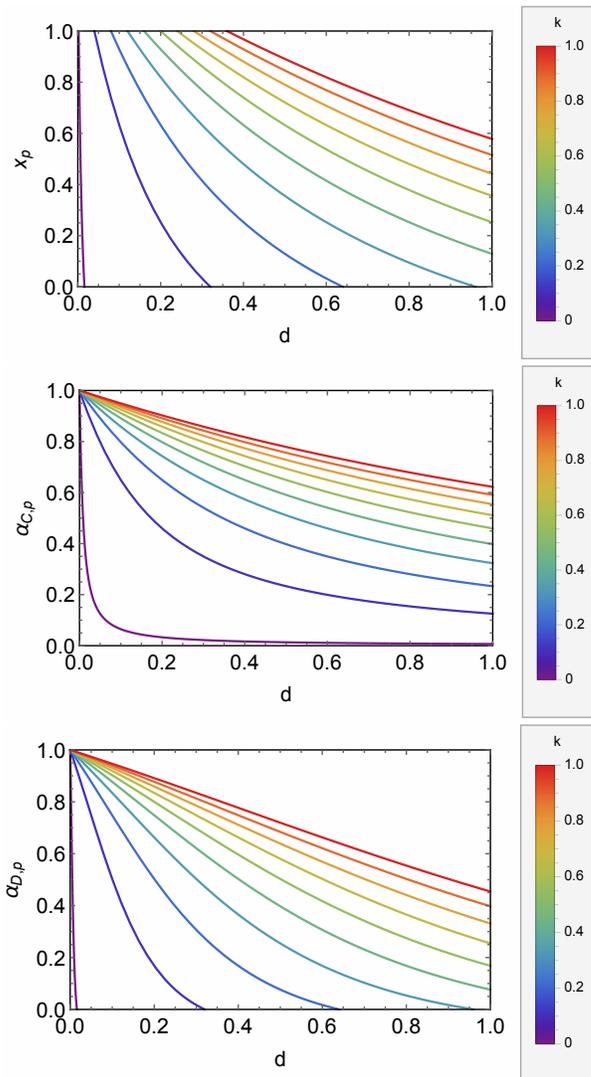
$$[b(1-r) + c](k + bd) < (2c - 1)k \qquad (7)$$

FIG. 1. Numerical values x_p, $\alpha_{C,p}$, and α_D, of the internal equilibrium as a function of model parameters k and d. The DDG in this scenario is given by $b = 1$, $c = 2$, and $r = 2$.

Finally, the stability of the fourth equilibrium cannot be easily deduced analytically and we analyze it numerically.

We first consider one of the most inspiring frameworks for this study: $b = 1$, $c = 2$, and $r = 2$ [24], where the death rate is fixed at $d = 0.05$. Figure 2 shows the numerical results for this scenario in the x-α_C plane and for different values of $k = 1.0, 0.1, 0.01$. In Fig. 2, left image, it is shown the case for $k = 1.0$ and the first three equilibria are present, while the condition for the fourth one are not satisfied here (see Fig. 1). Moreover, while both equilibria at $x = 0$ are stable but having very small basins of attraction, the equilibrium at $x = 1$ is stable (Eq. 7 is satisfied) and has an important basin of attraction which covers almost the entire space. This also means that for these parameter values the dynamics almost always leads to an equilibrium state at $x = 1$ ($\alpha_C \sim 0.95$ and $\alpha_D \sim 0.91$). Otherwise, when k decreases to 0.1, Fig. 2,
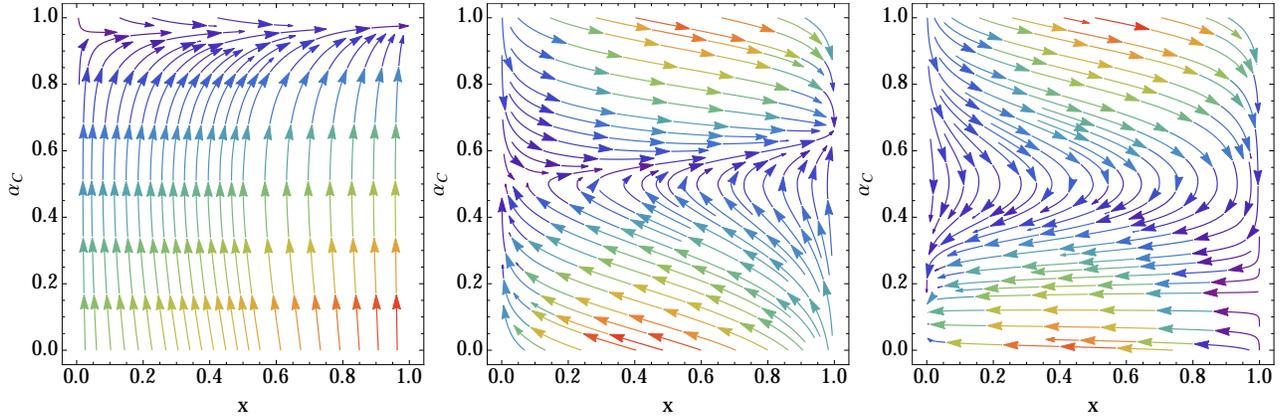
FIG. 2. Dynamics of x and α_C for $b = 1$, $c = 2$, $r = 2$, $d = 0.05$ and $k = 1$ (left), $k = 0.1$ (middle), and $k = 0.01$ (right). We can observe how the coefficient of efficiency, k, influences the dynamics of the system and the different basin of attractions for equilibria at $x = 0$ and $x = 1$.
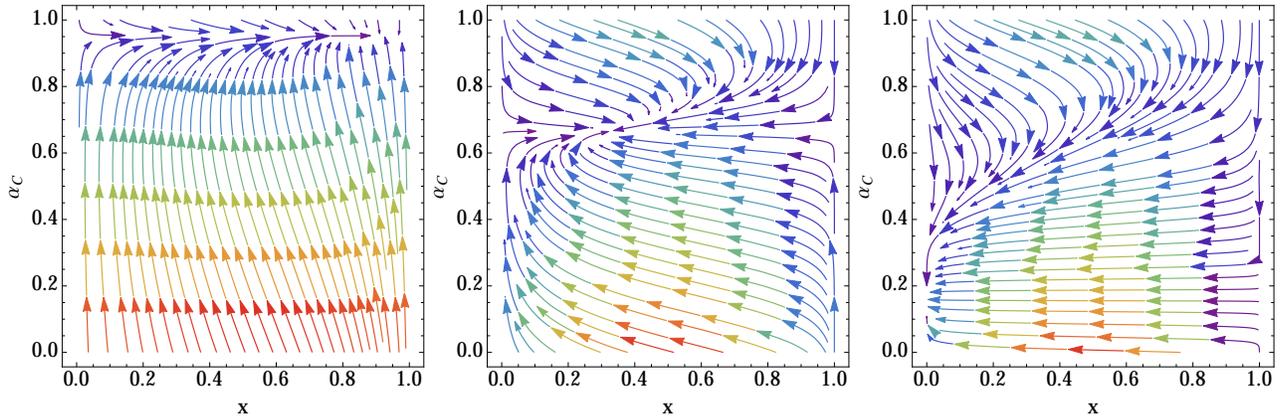

FIG. 3. Dynamics of x and α_C for $b = 1$, $c = 2$, $r = 1$, $d = 0.05$ and $k = 1$ (left), $k = 0.1$ (middle), and $k = 0.01$ (right). Modified perception of the benefit given by interactions cooperator-cooperator, $r = 1$. Here, cooperation is more difficult to achieve since there is not much incentive in being cooperative as the payoff for cooperators is the same against defectors or themselves. Thus, cooperation is more risky.
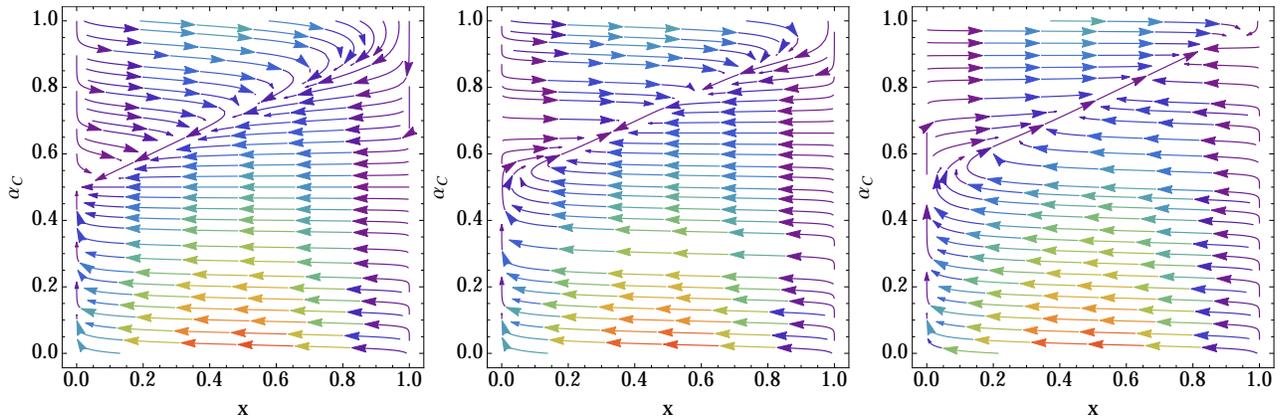

FIG. 4. Dynamics of x and α_C for $b = 1$, $c = 2$, $r = 1$, $k = 0.01$, and $d = 0.01$ (left), $d = 0.005$ (middle) and $d = 0.001$ (right). This scenario can be also interpreted as an increased speed of interactions or as a smaller efficiency of the death rate, d. Level of cooperation can be positively promoted when d tends to zero.
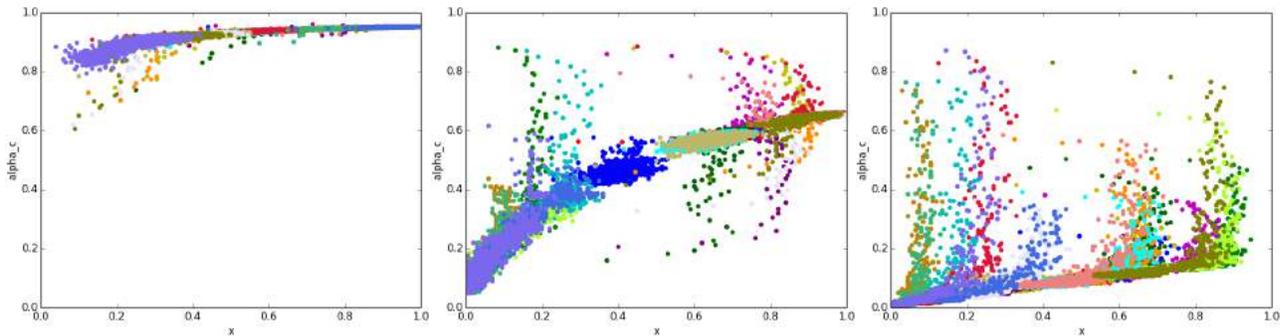
FIG. 5. Agent-based model dynamics of x and α_C for $b = 1$, $c = 2$, $r = 2$, $d = 0.05$ and $k = 1$ (left), $k = 0.1$ (middle), and $k = 0.01$ (right) and for an agents population size of 200. Same colors stand for the same run at different time steps. Different runs have also different initial conditions. In left and middle images dynamics almost always converges to the equilibrium at $x = 1$, while in the right image, even after a period at $x > 0$, the system is then attracted to the equilibrium in $x = 0$.

middle image, and to 0.01, Fig. 2, right image, dynamics does change. For $k = 0.1$, the equilibrium at $x = 1$ is again stable but for smaller values of α_C (and α_D), while for $k = 0.01$ this equilibrium becomes unstable and the system dynamics is attracted by equilibria at $x = 0$. For these model parameters, in order to obtain an attractor at $x = 1$, it is sufficient to have $k > 0.025$.

In Fig. 3 we show the dynamics of the system when the multiplication factor, r is decreased to one respect to the previous case: $b = 1$, $c = 2$, $r = 1$, and $d = 0.05$. Under these conditions, it is possible to observe the internal equilibrium in Fig. 3, left image, and Fig. 3, middle image. The conditions for the existence of the fourth equilibrium are not satisfied in Fig. 3, right image. We note also that, when this equilibrium exists, it is also stable and it has the largest basin of attraction in the variable space.

We now study the influence of d for different values in Fig. 4 for the case: $b = 1$, $c = 2$, $r = 1$, and $k = 0.01$, see also Fig. 3, right image, for comparison. Figure 4, left image, shows how the non-trivial equilibrium of $x = 0$ is affected by smaller values of d. In fact, this equilibrium remains stable with the larger basin of attraction but it possess a higher value of α_C (and α_D). The internal equilibrium does not exist for this scenario. On the other hand, when d decreases even more, middle and right images, it is possible to obtain the stable internal equilibrium and thus observe coexistence of cooperators and defectors.

AGENT-BASED SIMULATIONS

In this section we present an agent-based model for the DDG and for well-mixed but finite and heterogeneous populations of individuals. At the beginning of the simulation, we assign cooperation (C) as a strategy to a fraction x of agents in the population of size N. The remaining agents are assigned to defection (D). The dynamics of the population is then simulated by the following stochastic process at each time step t:

1. *random encounter*: two agents i and j are chosen uniformly at random among the population.

2. *pairwise interaction*: two agents play the DDG being cooperative, or not, according to their previously assigned strategy and their α-value is updated according to the discrete version of the α-function:

$$\alpha_{t+1} = (1 - \alpha_t)k\frac{B_c}{rb} + (1 - d)\alpha_t$$

Again, $B_c \in \{0, b, rb\}$, according to the number of cooperators in the interaction.

3. *strategy update*: focal agent i undergoes a strategy revision phase according to the local replicator rule based on his current payoff and the payoff of his current partner j [8]. The average payoff of i is calculated as follows:

$$\Pi_i = \sum_{k \in N} M[\sigma_i, \sigma_k]/N$$

where M indicates the payoffs matrix in Table I using α_i of the focal player i, σ_i is a vector giving the strategy profile of player i, using $C = 1$ and $D = 2$. The payoff of player j is calculated in the same way. Then, player i's strategy is replaced with player j's strategy with probability:

$$p_{i,j} = \frac{\Pi_j - \Pi_i}{K}$$

where $K = rb$, i.e. the maximum payoff a player can get in a game interaction.

4. Steps 1, 2, and 3 are repeated until the maximum number of timesteps is reached.

The above process follows the replicator dynamics equation and is the standard approximation in evolutionary game theory for finite well-mixed and structured populations [8, 20].

Corresponding results for this agent-based model are shown in Fig. 5 and they must be compared to those in Fig. 2, since the model parameters are the same. Same colors represent the same run at different time steps. Different runs have also different initial conditions in order to obtain and analyze different patterns of the corresponding dynamics. For instance, in Fig. 5, left image, the $k = 1.0$ system is always attracted towards the equilibrium in $x = 1$ and every run converges to the line at $\alpha_C = 0.95$, as predicted by numerical results in previous section. The dynamics quickly goes to very high α-level for any starting point in the space. Subsequently, the α-level remains very high and the fraction of cooperators x increases as time goes by. On the other hand, when we analyze the $k = 0.1$ case, Fig. 5, middle image, we observe that the system is attracted by both equilibria in $x = 0$ and $x = 1$. Here, the initial conditions play a more important role and the finite size of agent populations do not completely allow to reproduce numerical results for homogeneous and infinite size populations. Finally, the scenario for $k = 0.01$, Fig. 5, right image, is much more interesting. In fact, systems that possess initial conditions in the upper or upper-left region of the space increase very fast the level of cooperators jointly with a slow decrease in the α-level. When the α-level becomes too small ($\alpha_C < 0.2$), the system slowly goes the stable equilibrium in $x = 0$. This feature is also well-represented and predicted in Fig. 2, right image.

DISCUSSION

In this work we have introduced the approach of Dynamical Game Theory (DGT). We have focused our attention on a particular scenario of this new framework, – the Dynamical Donation Game (DDG). In this game agents start their interactions with a given convention, cooperation or defection, and they are matched randomly with another player in the population. The important feature of this game is that the payoffs can be interpreted differently by the players according to their perceptive state. This state changes in the subsequent interactions as a function of the cumulative gain players receive during the game and it is represented by the variable that we call α-level.

Our main finding is that the definition of the α-function, the way in which the α-level of the players is updated after each interaction, importantly influences the dynamics of the system. This can lead to very cooperative systems in which individuals fully cooperate with their partners as a result of their high α-level, which also leads to a perception of the game more prone to cooperation. On the contrary, when model parameters are not sufficient to obtain this stable equilibrium, the system is bound to reach smaller cooperation levels. However, we observed that an internal equilibrium for the level of cooperation $0 < x < 1$ is present and it is almost always stable. The last equilibrium we obtained is $x = 0$, where defection dominates the population. These results have been found analytically and by numerical simulations considering well-mixed, infinite size and homogenous populations and well-mixed, finite size and homogeneous populations.

Some future directions are still to be investigated. In particular, the theoretical approach of DGT can be useful in many disciplines. To begin with, the application and the study of other plausible α-functions is to be considered as an important follow-up work. real-life scenarios, that can be observed in social interactions, can be interpreted in the framework of dynamical games. Moreover, we would like to study also the influence of structured populations instead of well-mixed ones. As a matter of fact, the individual condition state may not only influence the perception of the game by the interacting player but also, it can alter and modify the frequency of interactions with some kind of players. It would be very interesting to see if particular correlations can be observed among individuals that play this kind of dynamical games.

[1] R. L. Trivers. The evolution of reciprocal altruism. *Quarterly Review of Biology*, pages 35–57, 1971.

[2] R. Axelrod. *The Evolution of Cooperation*. Basic Books, Inc., New-York, 1984.

[3] K. Sigmund. *The Calculus of Selfishness*. Princeton University Press, 2010.

[4] W. D. Hamilton. The genetical evolution of social behaviour I. *Journal of theoretical biology*, 7(1):1–16, 1964.

[5] E. Fehr and S. Gächter. Cooperation and punishment in public goods experiments. *American Economic Review*, pages 980–994, 2000.

[6] M. Milinski, D. Semmann, and H.-J. Krambeck. Reputation helps solve the "tragedy of the commons". *Nature*, 415(6870):424–426, 2002.

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

[8] J. Hofbauer and K. Sigmund. *Evolutionary Games and Population Dynamics*. Cambridge, N. Y., 1998.

[9] F. Vega-Redondo. *Economics and the Theory of Games*. Cambridge University Press, Cambridge, UK, 2003.

[10] J. Andreoni. Cooperation in public-goods experiments: kindness or confusion? *American Economic Review*, pages 891–904, 1995.

[11] C. Gracia-Lázaro, A. Ferrer, G. Ruiz, A. Tarancón, J. A. Cuesta, A. Sánchez, and Y. Moreno. Heterogeneous networks do not promote cooperation when humans play a Prisoner's Dilemma. *Proc. Natl. Acad. Sci. USA*, 109:12922–12926, 2012.

[12] J. Grujić, C. Fosco, L. Araujo, J. A. Cuesta, and A. Sánchez. Social experiments in the mesoscale: Humans playing a spatial Prisoner's Dilemma. *PLoS ONE*, 5(11):e13749, 2010.

[13] L. A. Dugatkin. *Cooperation among animals: An evolutionary perspective.* Oxford University Press, Oxford, 1997.

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

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

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

[17] L. A. Imhof, D. Fudenberg, and Martin A. Nowak. Tit-for-tat or win-stay, lose-shift? *Journal of theoretical biology*, 247:574–580, 2007.

[18] L. A. Martinez-Vaquero, J. A. Cuesta, and A. Sánchez. Generosity pays in the presence of direct reciprocity: A comprehensive study of 2x2 repeated games. *PLoS ONE*, 7:e35135, 2012.

[19] W. Poundstone. *Prisoner's Dilemma.* Random House LLC, 2011.

[20] J. W. Weibull. *Evolutionary game theory.* MIT press, 1997.

[21] J. F. Nash. Equilibrium points in n-person games. *Proc. Natl. Acad. Sci. USA*, 36(1):48–49, 1950.

[22] C. Hilbe, M. A. Nowak, and K. Sigmund. Evolution of extortion in iterated Prisoner's Dilemma games. *Proc. Natl. Acad. Sci. USA*, 110(17):6913–6918, 2013.

[23] S. H. Strogatz. *Nonlinear dynamics and chaos: with applications to physics, biology and chemistry.* Perseus publishing, 2001.

[24] L. A. Guaifi, N. Matador, M. W. D. Mathematica, L. Pirate, and A. de Pyramid. Drunk game theory. *to appear*, 2015.