Cooperation and self-regulation in a model of agents playing different games

H. Fort
Instituto de Física, Facultad de Ciencias, Universidad de la República, Iguá 4225, 11400 Montevideo, Uruguay
(Received 23 October 2002; revised manuscript received 15 April 2003; published 19 August 2003)

A simple model for cooperation between "selfish" agents, which play an extended version of the prisoner's dilemma game, in which they use arbitrary payoffs, is presented and studied. A continuous variable, representing the probability of cooperation, $p_k(t) \in [0,1]$, is assigned to each agent k at time t. At each time step t a pair of agents, chosen at random, interact by playing the game. The players update their $p_k(t)$ using a criterion based on the comparison of their utilities with the simplest estimate for expected income. The agents have no memory and do not use strategies based on direct reciprocity or "tags." Depending on the payoff matrix, the system self-organizes—after a transient—into stationary states characterized by their average probability of cooperation $\overline{p}_{eq}$ and average equilibrium per-capita income $\overline{p}_{eq}, \overline{U}_\infty$. It turns out that the model exhibits some results that contradict the intuition. In particular, some games that *a priori* seem to favor defection most, may produce a relatively high degree of cooperation. Conversely, other games, which one would bet lead to maximum cooperation, indeed are not the optimal for producing cooperation.

DOI: 10.1103/PhysRevE.68.026118 PACS number(s): 89.75.Fb, 89.65.Gh

I. INTRODUCTION

Game theory constitutes a powerful and versatile approach to analyze the collective behavior of adaptive agents, from humans to bacteria and firms. In particular, the *prisoner's dilemma* (PD) game plays in game theory a role similar to the harmonic oscillator in physics. It has been also referred to as the *Escherichia Coli* of social sciences, allowing a very large variety of studies. Indeed, this game, developed in the early 1950s, offers a very simple and intuitive approach to the problem of how cooperation emerges in societies of "selfish" individuals, i.e., individuals who pursue exclusively their own self-benefit. It was used in a series of works by Axelrod and co-workers [1] to examine the basis of cooperation between selfish agents in a wide variety of contexts. Furthermore, mechanisms of cooperation based on the PD have shown their usefulness in political science [2–4], Economics [5–11], international affairs [12–15], theoretical biology [16–18], and ecology [19,20].

The PD game consists in two players, say i and j, each confronting two choices: cooperate (C) or defect (D) and each makes its choice without knowing what the other will do. The four possible outcomes for the interaction of agent i with agent j are: (1) they can both cooperate (C,C), (2) both defect (D,D), (3) i cooperates and j defects (C,D), and (4) i defects and j cooperates (D,C). Depending on the situation (1), (2), (3), or (4), the agent i (j) gets, respectively: the "reward" $R(R)$, the "punishment" $P(P)$, the "sucker's payoff" S (the "temptation to defect" T) or $T(S)$. These four payoffs obey the following chain of inequalities:

$$T > R > P > S; \tag{1}$$

for instance, the four canonical PD payoffs are: $R=3$, $S=0$, $T=5$, and $P=1$. Clearly it pays more to defect: if one of the two players defects, say i, the other who cooperates will end up with nothing. In fact, even if agent i cooperates, agent j should defect, because in that case he will get T which is larger than R. That is, independently of what the other player does, defection D yields a higher payoff than cooperation and is the *dominant strategy* for rational agents. This is equivalent to saying, in a more technical language, that, the outcome (D,D) of both players is the Nash equilibrium [21] of the PD game. The dilemma is that if both defect, both do worse than if both had cooperated: both players get P which is smaller than R.

One can assign a *payoff matrix* M^{RSTP} to the PD game given by

$$M^{RSTP} = \begin{pmatrix} (R,R) & (S,T) \\ (T,S) & (P,P) \end{pmatrix},$$

which summarizes the payoffs for *row* actions when confronting with *column* actions.

The emergence of cooperation in PD games is generally assumed to require repeated play (and strategies such as *tit for tat* [1], involving memory of previous interactions) or features ("tags") permitting cooperators and defectors to distinguish one another [22].

In this work, I consider a simple model of selfish agents, possessing neither memory nor tags, to study the self-organized cooperative states that emerge when they play an *extended* PD game with arbitrary payoffs, i.e., payoffs that do not necessarily fulfill inequalities (1). The taxonomy of 2×2 games (one-shot games involving two players with two actions each) was constructed by Rapoport and Guyer [23], and showed that there exist exactly 78 nonequivalent games.

There are N_{ag} agents, with one variable assigned to each agent at the site or cell k and at time t: his probability of cooperation $p_k(t)$. Pairs of agents, i and j, interact by playing the PD game at each time step t. I use a mean field (MF) approach, in which all the spatial correlations in the system are neglected, and thus agents i and j are chosen at random [24]. After playing the game the players update their probability of cooperation $p_i(t)$ and $p_j(t)$ according to the same definite "measure of success" that does not vary with time. Thus all agents follow a universal and invariant strategy defined by a measure of success plus an updating rule to transform $p_i(t)$ and $p_j(t)$ into $p_i(t+1)$ and $p_j(t+1)$.

After a transient, the system self-organizes into a state of equilibrium characterized by the average probability of cooperation $\overline{p_{eq}}$ which depends on the payoff matrix.

Payoff matrices can be classified into subcategories according to their dominant strategy. Let us call M_D the class of those matrices, such that

$$T>R, \quad \text{and} \quad P>S, \tag{2}$$

for which the dominant strategy is *D*. This class comprises, for instance, the canonical matrix M^{3051} and M^{1053}, etc. A second class M_C corresponds to

$$R>T, \quad \text{and} \quad S>P, \tag{3}$$

for which the dominant strategy is *C*, examples of this class are the matrices M^{5310} and M^{3501}. The remaining matrices do not comply with Eq. (2) or (3) and produce situations, *a priori*, not dominated by (D,D) or (C,C).

One might wonder why bother to study matrices that imply no dilemma and are unrealistic in order to model the social behavior of the majority of individuals. There are several reasons. First, these ''unreasonable'' payoff matrices can be used by minorities of individuals which depart from the ''normal'' ones (assumed to be neutral). For instance, ''antisocial'' ''always *D*'' individuals, which cannot appreciate any advantage of cooperation, or ''altruistic'' ''always *C*'' individuals. Second, it seems interesting to test the robustness of cooperation under changes in the payoff matrix. In particular, we will see that even payoff matrices that imply a dilemma can produce either $\overline{p_{eq}}=0.5$ or $\overline{p_{eq}}=0$. Third, arbitrary payoff matrices could be also of importance in other contexts different from societies. One might envisage situations in which a definite value of $\overline{p_{eq}}$ is required or is desirable in the design of a system or is the one that optimizes the functioning of a particular mechanism, etc. For example, to understand how a market of competing firms attains self-regulation. Or, for instance, in the traffic problem, where the damage suffered from mutual *D* (crash) exceeds the damage suffered by being exploited (turn away), which is more appropriately described by the so-called *chicken game* for which $T>R>S>P$.

Fourth, we will show results for those payoff matrices that, at first glance, defy our intuition. For example, payoff matrices that, at least in principle, one would bet that favor defection and indeed produce a not so low degree of cooperation.

II. A MECHANISM TO PRODUCE COOPERATIVE EQUILIBRIUM STATES

Among the weaknesses of major approaches that have been considered to answer the question about the emergence of cooperation, two are often remarked. The first criticism is about the oversimplification in the behavior of agents: they either always cooperate (C) or always defect (D). Clearly, this is not very realistic. Indeed, the levels of cooperation of the individuals seem to exhibit a continuous gamma of values. The second objection is concerning the deterministic nature of the algorithms which seem to fail to incorporate the stochastic component of human behavior.

Both problems can be overcome by assigning to each agent *k* a probability of cooperation $p_k(t)$ (a real number in the interval $[0,1]$) instead of a definite behavior such as *C* or *D*. Concerning the first objection, $p_k(t)$ reflects the existence of a ''gray scale'' of levels of cooperation instead of just ''black'' and ''white.'' Regarding the second objection, the proposed algorithm is clearly nondeterministic: agent *k* plays *C* with probability p_k and *D* with probability $1-p_k$.

Now, let us describe the dynamics. The pairs of interacting partners, by virtue of the MF treatment, are chosen randomly instead of being restricted to some neighborhood. The implicit assumptions are that the population is sufficiently large and the system connectivity is high, i.e., the agents display high mobility or they experienced interaction at a distance (for instance, electronic transactions). In this work the population of agents will be fixed to $N_{ag}=1000$ and the number of time steps will be of order $t_f=10^5-10^6$ in such a way that both assumptions be also consistent with the fact that agents have no memory.

Starting from an initial state at $t=0$ taken as $p_k(0)$ chosen at random (in the interval $[0,1]$) for each agent *k*, the system evolves by iteration during t_f time steps following the procedure.

(1) *Selection of players.* Two agents, located at random positions *i* and *j*, are selected to interact, i.e., to play the PD game.

(2) *Playing pairwise PD.* The behavior, *C* or *D*, of each player *k* ($k=i$ or $k=j$) is decided generating a random number *r* and if $p_k(t)>r$ then he plays C and, conversely, if $p_k(t)<r$ he plays *D*.

(3) *Assessment of success.* Each of the two players compares his utilities $U_k(t)$, which is one of the four PD payoffs: *R*, *S*, *T*, or *P*, with an *estimate* $\epsilon_k(t)$ of his expected utilities. If $U_k(t)\geqslant\epsilon_k(t)$ $[U_k(t)<\epsilon_k(t)]$ the agent assumes he is doing well (badly) and therefore its level of cooperation is adequate (inadequate).

(4) *Probability of cooperation update.* If player *k* is doing well he keeps his probability of cooperation $p_k(t)$. On the other hand, if player *k* is doing badly he decreases (increases) his probability of cooperation $p_k(t)$ if he played $C(D)$ choosing an uniformly distributed value between $p_k(t)$ and 1 [between 0 and $p_k(t)$].

In order to introduce a simple and natural estimate $\epsilon_k(t)$ let us consider two players *i* and *j* who cooperate, at time *t*, with probabilities $p_i(t)$ and $p_j(t)$, respectively [and defect with probabilities $1-p_i(t)$ and $1-p_j(t)$], thus the expected utilities for the player *i*, $U_{ij}^{RSTP}(t)$, are given by

$$U_{ij}^{RSTP}(t)=Rp_i(t)p_j(t)+Sp_i(t)[1-p_j(t)]+T[1-p_i(t)]p_j(t)+P[1-p_i(t)][1-p_j(t)], \tag{4}$$

while the expected utilities for the player *j*, U_{ji}^{RSTP}, are obtained by interchanging *i* and *j* in the above equation.

This implies that, given the average probability of cooperation $\overline{p(t)}$, an arbitrary agent, say number *k*, can estimate his average expected utilities as

$$U_k^{RSTP}(\overline{p}(t)) = R\overline{p}(t)p_k(t) + Sp_k[1-\overline{p}(t)] + T[1-p_k(t)]\overline{p}(t) + P[1-\overline{p}(t)][1-p_k(t)]. \tag{5}$$

However, it turns out that, in general, the value of $\overline{p}$ is unknown by the agents. Hence, a simpler estimate that can be used by agent k for his expected utilities $\epsilon_k(t)$ is obtained by replacing in Eq. (5) $\overline{p}(t)$ by his own probability of cooperation $p_k(t)$:

$$\epsilon_k^{RSTP}(t) \equiv Rp_k^2(t) + (S+T)p_k[1-p_k(t)] + P[1-p_k(t)]^2 = (R-S-T+P)p_k^2(t) + (S+T-2P)p_k(t) + P. \tag{6}$$

In other words, agent k adopts the simplest possible extrapolation, i.e., that he is a "normal" individual whose probability of cooperation is representative of the average value.[1]

The rule each player follows to update his probability of cooperation is quite natural and of the type "win-stay" and "lose-shift." That is, if the player's utilities U_k are larger than his estimate, he keeps his probability of cooperation. On the other hand, if the utilities are smaller than his estimate, he changes his probability of cooperation: (a) increasing it if he played D or (b) decreasing it if he played C. From Eq. (6) the update of $p_k(t) \rightarrow p_k(t+1)$ is governed by the sign of $U_k(t) - \epsilon_k^{RSTP}(t)$, i.e., by the following inequations:

$$(S+T-R-P)p_k^2(t) - (S+T-2P)p_k(t) + \begin{pmatrix} R \\ S \\ T \\ P \end{pmatrix} - P \gtrless 0; \tag{7}$$

in the case >0 (<0) p_k is increased (decreased).

In the following section we will see that the strategy that results from the combination of the proposed measure of success and update rule for p_k—the steps (3) and (4)—produces, for a wide variety of payoff matrices, cooperative states with $\overline{p}_{eq} > 0$.

Let us end this section with a remark about the problem addressed here and its relation with the evolution of cooperation. In this approach, there is no competition of different strategies, all the agents follow the same universal strategy that does not evolve over time. However, the system is adaptive in the sense that the probabilities of cooperation of the agents do evolve.

[1]Considering more sophisticated agents, which have "good information" on the population (for instance, the value of $\overline{p}$ at time t), does not change substantially the main results obtained with these naive agents.

III. RESULTS

Depending on the payoffs R,S,T, and P the system self-organizes, after a transient, in equilibrium states with values of $\overline{p}_{eq}$ ranging from 0 to 1. The equilibrium asymptotic states can be lumped into three groups according to the degree of cooperation attained: *highly cooperative* ($\overline{p}_{eq} > 0.5$), *moderately cooperative* ($\overline{p}_{eq} \simeq 0.5$), and *poorly cooperative* ($\overline{p}_{eq} < 0.5$). The outcomes for any arbitrary payoff matrix M^{RSTP} can be understood in terms of the updating rule for the cooperation probability and the corresponding estimate ϵ^{RSTP}, i.e., from the inequalities (7).

The payoff matrices that imply a dilemma—those that comply with the chain of inequalities (1)—lead either to $\overline{p}_{eq} = \frac{1}{2}$ or to $\overline{p}_{eq} = 0$. From Eq. (7) it emerges that $\overline{p}_{eq} = 0.5$ occurs in the case when $\epsilon^{RSTP} - P$ has no roots in the interval (0,1] ($p=0$ is always one of the two roots) and $\overline{p}_{eq} = 0$ in the opposite case.

Some other matrices not belonging to class M_D exhibit a tension between C and D and give rise to $\overline{p} \simeq \frac{1}{2}$. The matrices that do not embody such trade-off produce the situations that depart most from $\overline{p}_{eq} \simeq \frac{1}{2}$.

It is illustrative to consider, for a moment, the restricted subset of 24 payoff matrices obtained from permutation of the four canonical payoff values because it covers the three groups with different cooperation levels mentioned above. In fact, the system self-organizes into equilibrium states with seven values of $\overline{p}_{eq}$: 2 matrices (M^{3501} and M^{3510}) produce $\overline{p}_{eq} = 1$,2 matrices, M^{1053} and M^{0153} produce $\overline{p}_{eq} = 0$. The remaining 20 matrices produce intermediate values: $\overline{p}_{eq} \simeq 0.72$ (M^{5301}), $\overline{p}_{eq} \simeq 0.62$ (M^{3510}), $\overline{p}_{eq} \simeq 0.38$ (M^{0135}), $\overline{p}_{eq} \simeq 0.28$ (M^{1035}), and $\overline{p}_{eq} = 0.5$ (the other 16 matrices and among them the canonical payoff matrix). The 24 measures are performed over 500 simulations. Figure 1 shows the average probability of cooperation for different payoff matrices versus time for the 50 000 first time steps. The mirror symmetry with respect to the value $\overline{p} = 0.5$ between the curves for $\overline{p}(t)$ corresponding to a given matrix M^{RSTP} and its palindrome M^{PTSR} is due to the symmetry of the game when interchanging $R \leftrightarrow P$ and $S \leftrightarrow T$ simultaneously with cooperators C by defectors D. That is,

$$\overline{p}^{RSTP}(t) = \overline{(1-p)}^{PTSR}(t). \tag{8}$$

A particular interesting case study is provided by payoff matrix M^{0135} with $\overline{p}_{eq} \simeq 0.38$. This result seems, at first sight, counterintuitive: an intermediate cooperation level attained with 0 reward (and very low sucker's payoff). Nevertheless, let us show how the updating rule for the cooperation probability explains this outcome. The estimate for this matrix, given by the parabola

$$\epsilon_k^{0135} = p_k^2 - 6p_k + 5, \tag{9}$$

plotted as a solid curve in Fig. 2 (the horizontal lines at $S=1$ and $T=3$ cut the parabola at abscises $p_S = 3 - \sqrt{5}$ and $p_T = 3 - \sqrt{7}$, respectively). The cooperation update rule tells

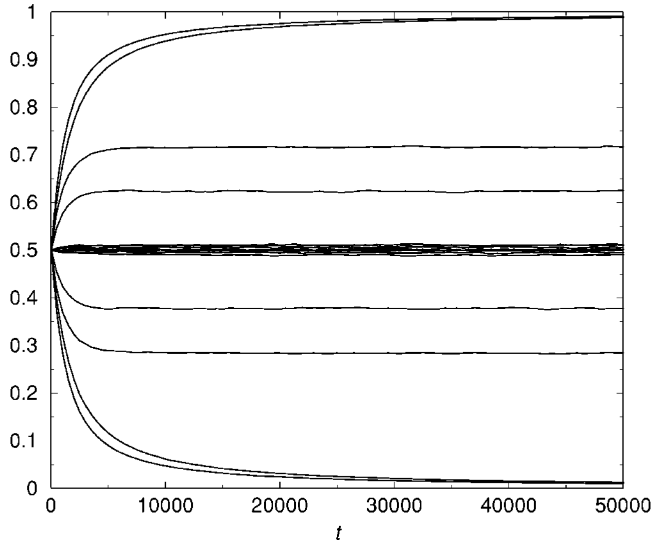

FIG. 1. Curves of $\overline{p}$ vs the number of iterations t, corresponding to the 24 payoff matrices obtained by permuting the four canonical payoffs $R=3$, $S=0$, $T=5$, and $P=1$. The system self-organizes in seven different cooperative states with: $\overline{p_{eq}}=1$, $\overline{p_{eq}}\simeq 0.72$, $\overline{p_{eq}}\simeq 0.62$, $\overline{p_{eq}}\simeq 0.5$, $\overline{p_{eq}}\simeq 0.38$, $\overline{p_{eq}}\simeq 0.28$, and $\overline{p_{eq}}\simeq 0$.

us that the agent k increases his probability of cooperation when he plays D and gets $T=3$ if p_k is less than $p_T=3-\sqrt{7}<0.5$, i.e., this temptation is not enough $[T<\epsilon_k^{0135}(p_k)]$. On the other hand, he decreases his probability of cooperation when he plays C and gets $R=0$, independently of the value of p_k, or when he gets $S=1$ if p_k is less than $p_S=3-\sqrt{5}>0.5$. In the remaining situations the player keeps his probability of cooperation. Thus a value of $\overline{p_{eq}}$ between 0 and 0.5 is not surprising after all, rather it is the result of giving the two competing probabilities of cooperation flows. All this analysis for payoff matrix M^{0135} works also for any set of payoffs obeying the inequalities

$$P>T>S>R, \tag{10}$$

the only thing that changes is the value of $\overline{p_{eq}}$. We will come back over this particular payoff matrix to illustrate how p_{eq} changes under arbitrary variations of the payoffs.

The effect of changing payoffs

We are now going to analyze the effect of changing the payoff matrix in order to go beyond the 24 permutations of the canonical payoffs.

We have seen that the sign of $U_k-\epsilon_k$ controls the update of p_k. From the definition of ϵ_k, as an estimate of utilities of agent k, it is clear that it is bounded from above and from below by the largest and smallest of the four payoffs, respectively. Thus, $U_k-\epsilon_k$ may have different signs, depending on the value of p_k, only for the two intermediate payoffs. Let us denote by p_1 the value of p_k such that the estimate ϵ_k becomes equal to the larger payoff, p_2 the value of p_k such that the estimate becomes equal to the second larger payoff, and so on. Therefore, it is easy to see that the change in $\overline{p_{eq}}$ is

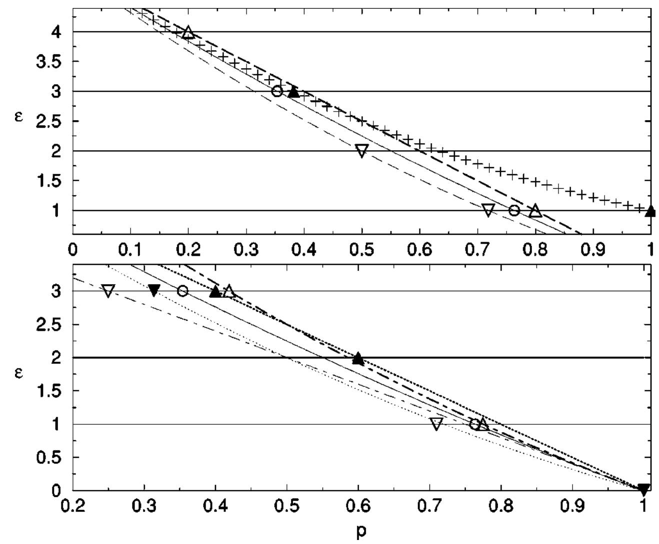

FIG. 2. (a) Below: The estimate $\epsilon^{0135}(p)$ vs p (solid line) compared with the estimates that result for independent variations of payoffs S and P, once at a time: ϵ^{0035} (dotted thin line), ϵ^{0235} (dotted thick line), ϵ^{0134} (dot-dashed thin line), and ϵ^{0136} (dot-dashed thick line). The circles correspond to the points $\epsilon_{0135}=1$ and $\epsilon_{0135}=3$. The filled up (down) triangles correspond to the points $\epsilon_{0235}=2$ and $\epsilon_{0235}=3$ ($\epsilon_{0035}=0$ and $\epsilon_{0035}=3$). The nonfilled up (down) triangles correspond to the points $\epsilon_{0136}=1$ and $\epsilon_{0136}=3$ ($\epsilon_{0134}=1$ and $\epsilon_{0134}=3$). (b) Above: The estimate $\epsilon^{0135}(p)$ vs p (solid line) compared with the estimates that result for independent variations of payoffs T and R, once at a time: ϵ^{0125} (dashed thin line), ϵ^{0145} (dashed thick line), and ϵ^{1135} (+'s). The circles correspond to the points $\epsilon_{0135}=1$ and $\epsilon^{0135}=3$. The filled up triangles correspond to the points $\epsilon^{1135}=1$ and $\epsilon^{1135}=3$. The nonfilled up (down) triangles correspond to the points $\epsilon^{0145}=1$ and $\epsilon^{0145}=4$ ($\epsilon^{0125}=1$ and $\epsilon^{0125}=2$). See text.

controlled by the displacements of p_2 and p_3 (for instance, for M^{0135}, $p_2\equiv p_T$ and $p_3\equiv p_S$). If p_2 or p_3 corresponds to the cooperative payoff, R or S, then its displacement to the right (left) decreases (increases) the proportion of C agents for whom $U_k>\epsilon_k$ which are, on average, the ones who remain C after playing the game. This in turn decreases (increases) $\overline{p_{eq}}$. On the other hand, if p_2 or p_3 corresponds to the noncooperative payoff, T or P, then its displacement to the right (left) decreases (increases) the proportion of D agents for whom $U_k>\epsilon_k$ which are, on average, the ones who remain D after playing the game. This in turn increases (decreases) $\overline{p_{eq}}$.

The payoff matrix M^{0135} will serve to illustrate the effect the changes in the values of the payoffs have on p_{eq}. We will proceed by modifying one of the four payoffs at a time and keeping fixed the remaining three in such a way that the chain of inequalities (10) is preserved. This variation of a quantity that results when the payoff X is modified and the other three payoffs remain fixed is denoted by δ_X. The estimates that result from these changes are the curves shown in Fig. 2. Let us consider first the changes δ_{S^+}, produced by an increment in the sucker's payoff from $S=1$ to $S=2$ (which transforms M^{0135} into M^{0235}), δ_{S^-}, produced by a decrease from $S=1$ to $S=0$ (which transforms M^{0135} into M^{0035}). For

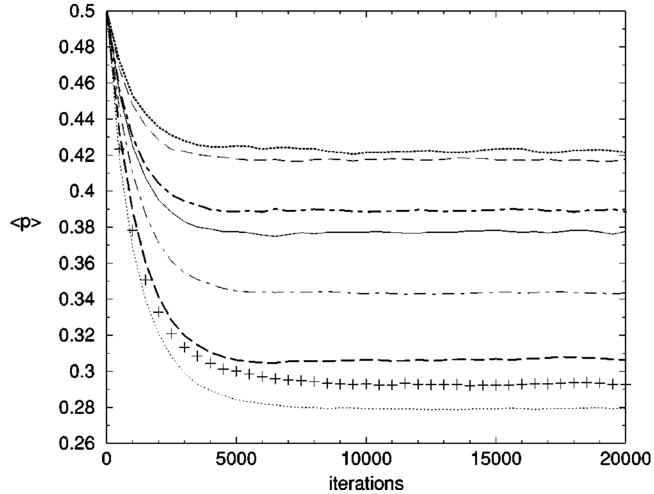

FIG. 3. The effect of independent variations of payoffs S,T, and P (the reward remains fixed at $R=0$), once at a time, from payoff matrix M^{0135}. Variations of S: $S=2$ (dotted thick line), $S=0$ (dotted thin line). Variations of P: $P=6$ (dot-dashed thick line), $P=4$ (dot-dashed thin line). Variations of T: $T=4$ (dashed thick line), $T=2$ (dashed thin line). Variations of R: $R=1$ ("+").

M^{0235}, p_3 is the abscise of the point $\epsilon_k \equiv S=2$ [filled up triangle in Fig. 2(a)] and for M^{0035}, p_3 is the abscise of the point $\epsilon_k \equiv S=0$ [filled down triangle in Fig. 2(a)], while the corresponding p_2 are the abscises of the points $\epsilon_k \equiv T=3$ [filled triangles: up for M^{0235} and down for M^{0035} in Fig. 2(a)]. We can see that increasing (decreasing) the sucker's payoff, from $S=1$ to $S=2$ $(S=0)$, produces a displacement of p_2 to the right (left), from $3-\sqrt{7}\simeq 0.354$ to 0.4 [to $(7-\sqrt{33})/4\simeq 0.314$], and of p_3 to the left (right), from $3-\sqrt{5}\simeq 0.764$ to 0.6 (to 1). Hence, both changes point in the same direction increasing (decreasing) p_{eq} as can be observed in Fig. 3 (dotted lines versus solid lines). In other words,

$$\delta_{S^+}(p_2-p_3)\simeq(0.4-0.354)-(0.6-0.764)=0.21>0,$$
$$\delta_{S^-}(p_2-p_3)\simeq(0.314-0.354)-(1-0.764)=-0.276<0. \tag{11}$$

Similarly, we denote by δ_{P^+} the variations produced by an increment in the punishment, from $P=5$ to $P=6$ (which transforms M^{0135} into M^{0136}), and by δ_{P^-} the variations produced by a decrease in the punishment, from $P=5$ to $P=4$ (which transforms M^{0135} into M^{0134}). For both matrices, the corresponding p_2 and p_3 are the abscises of the points $\epsilon_k \equiv T=3$ and $\epsilon_k \equiv S=1$ (nonfilled triangles in Fig. 2: up for M^{0136} and down for M^{0134}), respectively. Also in Fig. 2(a) we see that changing the punishment, from $P=5$ to $P=6$ $(P=4)$, produces a displacement of p_2 to the right (left), from $3-\sqrt{7}\simeq 0.354$ to $(4-\sqrt{10})/2\simeq 0.419$ (to 0.25) , and of p_3 to the right (left), from $3-\sqrt{5}\simeq 0.764$ to $(4-\sqrt{6})/2\simeq 0.775$ (to 0.75), hence the two changes point in opposite directions: the first tends to increase (decrease) p_{eq} and the second to decrease (increase) it. As the first displacement is larger it dominates, and the net result is an increase (decrease) of p_{eq} as can be observed in Fig. 3 (dot-dashed lines versus solid line). That is,

$$\delta_{P^+}(p_2-p_3)\simeq(0.419-0.354)-(0.775-0.764)=0.054>0,$$
$$\delta_{P^-}(p_2-p_3)\simeq(0.25-0.354)-(0.75-0.764)=-0.09<0. \tag{12}$$

On the other hand, let us consider the variations produced by the increment of the temptation δ_{T^+}, from $T=3$ to $T=4$ (which transforms M^{0135} into M^{0145}), and by its decrease δ_{T^-}, from $T=3$ to $T=2$ (which transforms M^{0135} into M^{0125}). For M^{0145}, p_2 is the abscise corresponding to the point $\epsilon_k \equiv T=4$ (nonfilled up triangle) and for M^{0125}, p_2 is the abscise of the point $\epsilon_k \equiv T=2$ (nonfilled down triangle), while the corresponding p_3 are the abscises of the points $\epsilon_k \equiv S=1$ (nonfilled triangles: up for M^{0145} and down for M^{0125}). In Fig. 2(b) we can see that increasing (decreasing) the sucker's payoff, from $T=3$ to $T=4$ $(T=2)$, produces a displacement of p_2 to the left (right), from $3-\sqrt{7}\simeq 0.354$ to 0.2 (to 0.5) and of p_3 to the right (left), from $3-\sqrt{5}\simeq 0.764$ to 0.8 [to $(7-\sqrt{17})/2\simeq 0.719$]. Hence, both changes point in the same direction decreasing (increasing) p_{eq} as can be observed in Fig. 3 (dashed lines versus solid line). That is

$$\delta_{T^+}(p_2-p_3)\simeq(0.2-0.354)-(0.8-0.764)=-0.19<0,$$
$$\delta_{T^-}(p_2-p_3)\simeq(0.5-0.354)-(0.719-0.764)=0.19>0. \tag{13}$$

With a similar argument one realizes that increasing (decreasing) the reward $R=0$ p_{eq} decreases (increases).

In summary, for payoff matrices like M^{0135}, which obey the chain of inequalities (10), we found two expected results: a higher value of p_{eq} can be reached by increasing the sucker's payoff S (which makes C agents more altruistic) or decreasing the temptation T (reducing the incentives to free ride). Additionally, we found two *a priori* unexpected results: a higher value of p_{eq} can also be reached by increasing the punishment P or decreasing the reward R. By an inspection of Fig. 2(a) the effect of an increment of P can be understood as a rising the expectations of the D agents, which in turn diminishes the fraction of agents that are satisfied after playing the game. Similarly, from Fig. 2(b) we can see that decrease of R makes the C agents less ambitious and increases the fraction of altruists.

It is worth remarking that, for the case of payoffs obeying Eq. (10), something which at first seems as innocent as to interchange the two noncooperative payoffs T and P has a dramatic consequence: it transforms a system with an intermediate level of cooperation into one with null cooperation. This can be understood by comparing the estimate (9) for payoff matrix M^{0135} to the one for M^{0153}, which is given by

$$\epsilon_k^{0153}=-3p_k^2+3. \tag{14}$$

Both estimates have maximum value of P (5 and 3, respectively, at $p_k=0$), but the important difference is that in the first case P is the maximum payoff while in the second one $P<T$. Thus in this second case, only the agents that play C can do badly, and then the only possible change for p_k (according to its update rule) is a reduction till it reaches zero value.

Finally, let us include a note regarding the efficiency to attain cooperative regimes. The state of maximum cooperation $\overline{p}_{eq}=1$ is reached for payoff matrices such that $S \geqslant R > \max\{T,P\}$ plus the condition that equation

$$(S+T-R-P)p^2-(S+T-2P)p+R-P=0 \qquad (15)$$

has no roots in the interval [0,1] different from $p=1$ [which is always a root of Eq. (15)]. This condition on the roots is because in the opposite case, when there is a root p_x in between 0 and 1, it follows easily from inequations (7) that $\overline{p}$ converges to the semisum of p_x and 1. It can be checked by elemental algebra that this is the case of, for instance, payoff matrices M^{3501}, M^{3510}.

IV. CONCLUSIONS

The success of the strategy to attain cooperative regimes for a wide variety of games (payoff matrices)—mainly those that imply dilemmas or clearly favor D—relies on the combination of the proposed measure of success and update rule for the probability of cooperation. Basically, it works by tuning the agent's cooperation guided by a trade-off between efficiency (increase of utilities) and equity (indirect reciprocity). If the agent is doing well he maintains his probability of cooperation, otherwise he changes it. When he is doing badly playing D he becomes more cooperative, i.e., he increases his probability of cooperation, attempting to change to behavior C and explore this alternative behavior. Conversely, if he is doing badly playing C then he decreases his probability of cooperation attempting to change to behavior D and see what happens.

An interesting extension of the model would be to allow competition of different strategies to promote their evolution, i.e., players who imitate the best-performing ones in such a way that lower scoring strategies decrease in number and the higher scoring ones increase.

Another possibility would be to allow the use of distinct payoff matrices. For instance, individuals inclined to cooperate (defect) might be represented by agents using the payoff matrix M^{5301} (M^{1035}) while "neutral" ordinary agents by those using the canonical payoff matrix M^{3051}. This would make possible to study if mutants inclined to D can invade a group of neutral individuals or individuals inclined to C and drive out all cooperation.

Here I considered a MF approximation that neglects all the spatial correlations. One virtue of this simplification is that it shows that the model does not require that agents interact only with those within some geographical proximity in order to sustain cooperation. Playing with fixed neighbors is sometimes considered as an important ingredient to successfully maintain the cooperative regime [25,26]. However, the quality of this MF approximation depends on the nature of the system one desires to model, and varies whether one deals with human societies, viruses [27], cultures of bacteria [28], or market of providers of different products. In order to consider situations in which the effect of geographic closeness cannot be neglected, an alternative version of this model might include spatial games, in which individuals interact only (or mainly) with those within some geographical proximity. In that case, the study of spatial patterns seems an interesting issue to address. Work is in progress in that direction.

ACKNOWLEDGMENTS

I thank R. Axelrod for valuable opinions and R. Donangelo for useful discussions and criticism.

[1] R. Axelrod, *The Evolution of Cooperation* (Basic Books, New York, 1984); *The Complexity of Cooperation* (Princeton University Press, Princeton, NJ, 1997). These two volumes include lots of useful references. Also illuminating is R. Axelrod and M. Cohen, *Harnessing Complexity: Organization Implications of a Scientific Frontier* (The Free Press, Ann Arbor, MI, 2000), Chap. 3.
[2] H.A. Simon, Science **260**, 1665 (1990).
[3] R. Axelrod, Am. Polit. Sci. Rev. **75**, 306 (1981).
[4] J.M. Grieco, J. Polit. **50**, 600 (1988).
[5] E. Fehr and U. Fischbacher, Econom. J. **112**, 478 (2002).
[6] K. Clark and M. Sefton, Econom. J. **111**, 51 (2001).
[7] I. Bohnet and B. Frey, J. Econ. Behav. Organ. **38**, 43 (1999).
[8] J. Andreoni and J.H. Miller, Econom. J. **103**, 570 (1993).
[9] K. Binmore and L. Samuelson, J. Econ. Theory **57**, 278 (1992).
[10] B. Kogut, J. Industrial Economics **38**, 183 (1989).
[11] D. Warsh, Harvard Bus. Rev. **67**, 26 (1989).
[12] R. Hausmann, Foreign Policy **122**, 44 (2001).
[13] J. Goldstein, International Studies Quarterly **35**, 195 (1991).
[14] R. Powell, Am. Polit. Sci. Rev. **85**, 1303 (1991).
[15] G.H. Snyder, International Studies Quarterly **15**, 66 (1971).
[16] L.M. Wahl and M.A. Nowak, J. Theor. Biol. **200**, 307 (1999); **200**, 323 (1999).
[17] M.A. Nowak and K. Sigmund, J. Theor. Biol. **168**, 219 (1994).
[18] M.A. Nowak, Theor Popul. Biol. **38**, 93 (1990)
[19] M. Mesterton-Gibbons and L.A. Dugatkin, Anim. Behav. **54**, 551 (1997).
[20] L.A. Dugatkin, M. Mesterton-Gibbons and A.I. Houston, Trends in Evolutionary Ecology **7**, 202 (1992).
[21] J. Nash, Ann. Math. **54**, 286 (1951).
[22] J. Epstein, *Zones of Cooperation in Demographic Prisoner's Dilemma*, Complexity **4**(2) (1998) (available at www.santafe.edu/sfi/publications/wpabstract/199712094).
[23] A. Rapoport and M. Guyer, General Systems **11**, 205 (1966).
[24] A similar study restricted to the canonical prisoner's dilemma matrix was performed by R. Donangelo and H. Fort.
[25] M.D. Cohen, R.L. Riolo, and R. Axelrod, Rationality and Society **13**, 5 (2001).
[26] M.A. Nowak and R. May, Nature (London) **359**, 826 (1992).
[27] P.E. Turner and L. Cho, Nature (London) **398**, 441 (1999).
[28] R. Dawkins, *The Selfish Gene: New Edition* (Oxford University Press, New York, 1989).