Adoption of simultaneous different strategies against different opponents enhances cooperation

Lucas Wardil[*] and Jafferson K. L. da Silva[†]

Departamento de Física, Universidade Federal de Minas Gerais - Caixa Postal 702,

CEP 30161-970, Belo Horizonte - MG, Brazil

Abstract

The emergence of cooperation has been widely studied in the context of game theory on structured populations. Usually the individuals adopt one strategy against all their neighbors. The structure can provide reproductive success for the cooperative strategy, at least for low values of defection tendency. Other mechanisms, such punishment, can also be responsible for cooperation emergence. But what happens if the players adopt simultaneously different strategies against each one of their opponents, not just a single one? Here we study this question in the prisoner dilemma scenario structured on a square lattice and on a ring. We show that if an update rule is defined in which the players replace the strategy that furnishes the smallest payoff, a punishment response mechanism against defectors without imputing cost to the punishers appears, cooperation dominates and, even if the tendency of defection is huge, cooperation still remains alive.

1

In biological systems cooperation is an emergent feature [1], which can be analyzed in the framework of evolutionary games on graph theory. It analyzes the evolution of strategies adopted by individuals in biological, social or other contexts, in which the interactions are given by the links of a graph [2]. The question that has been set is why cooperation can be established if it is costly for the cooperators [3]. Direct reciprocity [4], indirect reciprocity [5], spatial structure [6], in which individuals have their interaction constrained by some network of interactions, and other mechanisms have been proposed to answer this question. The spatial structure effect has been widely studied and was proved to be efficient to establish cooperation [1, 7–14]. Another mechanism is altruistic punishment where, besides the defector and cooperator, there is the altruistic punisher who acts like a cooperator but also pays for punishing defectors. Along with altruistic punishment, other mechanisms have been proposed, like introduction of volunteering participation in public good games [15] and selection among groups [16].

The cooperation problem is usually mapped into a game theory framework in which the individuals are players. Usually the player can choose one of two strategies – cooperation (C) or defection (D) – and uses this strategy during a game round. A strategy update can be set, and for the next round, the players can change their strategy taking into account the payoff earned during the round. If the players are set on the vertices of a network, they are constrained to play against their next neighbors. Note that, if some strategy is chosen, the same must be used against all of the opponents during the same round. Recently social diversity has been introduced [17] in the context of cooperation in public good games. The promotion of cooperation was analyzed by setting each individual to play different games in each neighborhood. This opens the question of what could happen if each player could choose different strategies against different opponents and adopts them during the same round. Let's state this point clearly. Suppose N individuals set on the vertices of a network. Each player can choose different strategies against each neighbor and earn different payoffs from each interaction. If the network is a square lattice, player x can choose, for example, the (C,C,C,C) strategy set, meaning that player x chooses the C strategy against all of his opponents.

In this letter, we study the emergence of cooperation if the players adopt different strategies against different opponents. We use the prisoner dilemma with population structured on a square lattice and on a ring as the scenario for the cooperation problem. The players use

different strategies against different neighbors and use the imitation rule with synchronous updating [9]. In the prisoner dilemma [18], both players receive 1 upon mutual cooperation, ϵ upon mutual defection; the defector receives b if the other cooperates, and the cooperator receives 0 if the other defects. The tendency to defect is given by b ($b > 1$), and ϵ is taken to be small ($\epsilon \ll 1$). When player x interacts with y, we assume that x has the information of (i) y's cumulative payoff and of (ii) the strategy y is using against him. In this rule players imitate the successful strategies of their neighborhoods. Each player interacts with its next neighbors and plays a round of one game with each opponent and earns a cumulative payoff. Then the player, let's call it x, chooses randomly a neighbor, y, and compares both cumulative payoffs P_x and P_y. If $P_x \geq P_y$, player x stays with its own strategy, but if $P_x < P_y$, then player x adopts the opponent's strategy with probability $(P_y - P_x)/4b$ [9]. So if player x decides to copy the strategy y is using against him, according to the rules stated above, he must also decide against which of his neighbors he will change his strategy by the new one. We have defined three different replacement rules that give rise to three different update rules. Let's call them just models A, B and C. We are going to focus on model A and use the others to highlight features of the former.

In the model A, player x chooses for the replacement the interaction that gives to him the smaller contribution for his cumulative payoff. The possible confrontations are (D,C),(C,C),(D,D) and (C,D), where the first entry is x's strategy and the second is the opponent strategy. Each one of these possibilities contributes to the cumulative payoff with b, 1, ϵ and 0, respectively. If player x is being exploited, he can neutralize the exploiter by changing (C,D) to (D,D) but keep other possible cooperation. It must be noted that if $\epsilon = 0$, both (D,D) and (C,D) contribute with the same amount, so the player chooses randomly one of them. In the usual game, where each player adopts the same strategy against all of his opponents, although the imitation also decreases the exploiter payoff, it cannot be thought of as a punishment act because the imitator is actually trying to improve his own benefits by probably exploiting other cooperators. But in model A, when (C,D) is changed to (D,D), just the exploiter is being punished, the other cooperation are not changed. So model A incorporates into the replacement rule both imitation and punishment without costly mechanisms. It must be mentioned that we are not using punishment as another strategy, like altruistic punishment [16]. There are only two basic strategies: cooperation and defection.

In the second replacement rule, model B, player x changes exactly the strategy he uses

3

against the player he chose to imitate. This model is trivial but interesting, because since two players reach mutual cooperation they remain forever and since an exploiter is punished there is no chance of other exploitation and they will stay forever with mutual defection. In the third replacement rule, model C, the tendency to punish exploiters is smoothed because player x chooses randomly one of his confrontations and changes the strategy used in that confrontation.

We made the simulations on the square lattice with linear size $L = 100$ and checked that the same results are still valid for $L = 200$. First we measured the mean fraction of cooperation (f_c) at stationary state for different values of the tendency for defection b. The number of cooperation (n_c) is taken to be the quantity of C strategies used in all confrontations, so $0 \leq n_c \leq 4N$ and $f_c = n_c/(4N)$. The random initial configuration is characterized by 50% of cooperation. For each initial condition we discarded the transient time needed to reach the stationary state, and then we made a time average for 1000 MCS (Monte Carlo Steps). A second average is realized over 100 different initial conditions. We used $1 \leq b \leq 5$ and $\epsilon = 0.01$. Fig. 1 shows the results for the three models and for the usual game. Note that for model C and the usual game, defection is the dominating strategy when $b > b_c$. Note also that models A and B keep cooperation alive for every b. It must be mentioned that it is possible to have some residues of defection in the case of cooperation domination or some residues of cooperation in the case of defection domination. This is due to the fact that the update rule prevents a change if the payoffs are the same. This prevention enables the existence of some stable residual configurations. For model A the fraction of cooperation for $b = 2$ is around $f_c = 0.9998$. For model C the residues of cooperation is around $f_c = 0.005$ at the region $b > b_c$.

Let's now analyze in more detail the model A. In the usual game, extensive analysis has been done and the emergence of cooperators is attributed in part to the negative effect on the cumulative payoff when an exploited individual copies a D strategy used by a successful exploiter, decreasing the exploiter payoff [6, 19]. This also happens in model A, as can be seen in Fig. 2, that depicts the typical behavior for short times. Note that when the quantity of (D,D) is a maximum, the fraction of cooperation is a minimum and starts to increase.

Another reason in the usual game for the maintenance of cooperation is the formation of cooperation clusters. This is again present in model A. Suppose we have a cluster as in Fig. 3, where all the players adopt defection in all confrontation except for a cluster
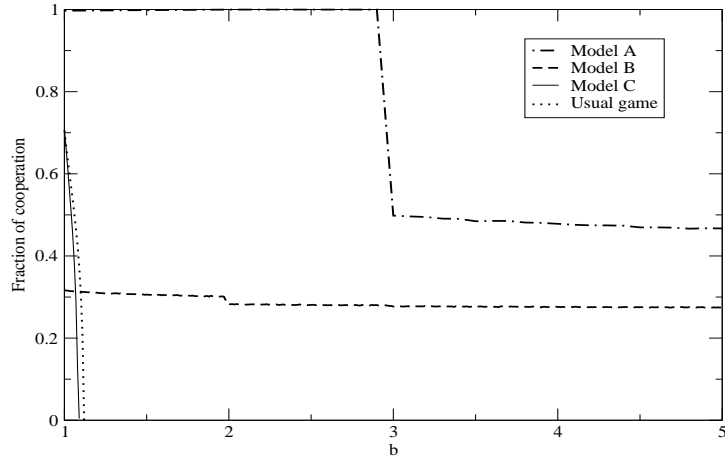
4

FIG. 1: Fraction of cooperation in the square lattice at the steady state for different values of b. It is shown the behavior of the usual game and the studied models A, B and C.
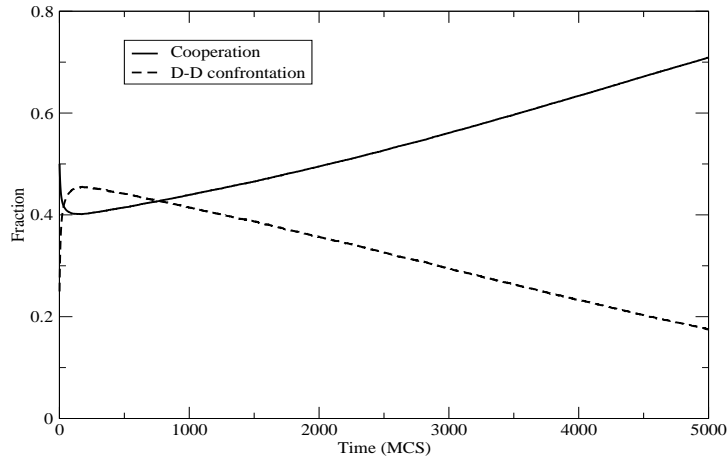

FIG. 2: Fraction of interaction of type $D - D$ and fraction of cooperation at short times for model A in the square lattice . The defect tendency is b=1.1. Each MCS (Monte Carlo Step) corresponds to a round in which every player play a game with all their neighbors.

of cooperation. One can see that for $b > 1$, the region inside the dashed square is stable in the sense that it will never be invaded by a defection. Suppose some player on the interface chooses to change his (C,D) confrontation. Since this confrontation furnishes the worst contribution, it is only that confrontation that can be changed to (D,D). After that, (D,D) is also the worst confrontation, and the D strategy cannot be copied to the interior of the square. If $b \geq 3$ the cooperation in such configuration cannot spread out, because
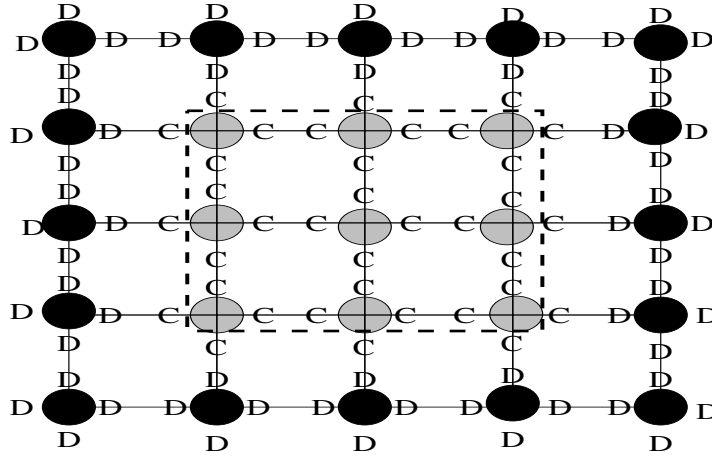
FIG. 3: Cluster of cooperation of size $S = 3$ and the region of stability shown inside the dashed square.

the exploitation of the immediately outside players give to each of them a total payoff of $b > 3$, and players immediately inside have a payoff value of only 3. But if $1 < b < 3$ the cooperation cluster can expand. Moreover a similar analysis applied to configurations like the one depicted in Fig. 4 shows that there is always a stable region of cooperation, irrespectively of the boundary and the width of the region of cooperation. For simplicity, let's put $\epsilon = 0$ in cluster analysis and evaluate numerically the probability that a square cluster of linear size S can take over a bigger square region of twice the original linear size. For $4 \leq S \leq 10$, with probability one the cluster expands. So the two mechanisms of cluster of cooperation formation and the negative effect of the exploiter being imitated are present. But the great difference of the model A is the possibility of neutralizing a defector and keeping the old cooperation. The importance of this fact can be better illustrated in a more general case. Suppose a big exploiter is introduced interacting with all individuals of a population of cooperators. After a few rounds, some players adopt D strategy against the big exploiter and still keep the cooperation with the other players. The exploitation will be almost completely neutralized. No matter how great is the exploitation, model A always prevents the invasion of defectors.

If $b > 3$, the strategies on the boundary of the stable region fluctuate. By using a rate equation approach we obtain an analytical expression for the fraction of cooperation ρ inside
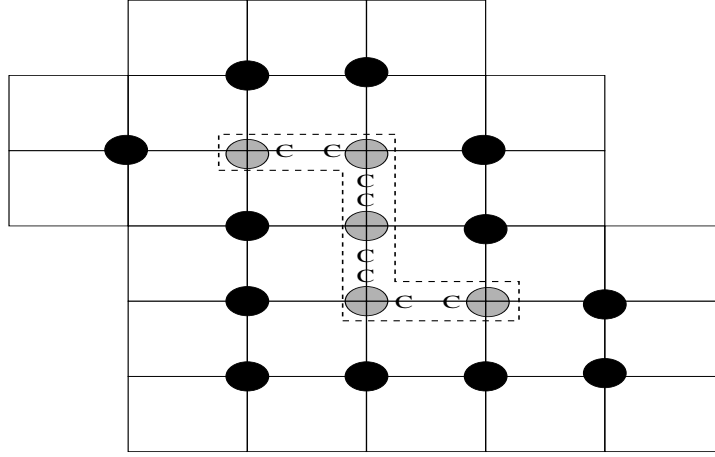
6

FIG. 4: Mutual cooperation configurations in model A prevent invasion by defections. Note that no matter the boundary nor the region width, there is always a stable region inside the dashed line.
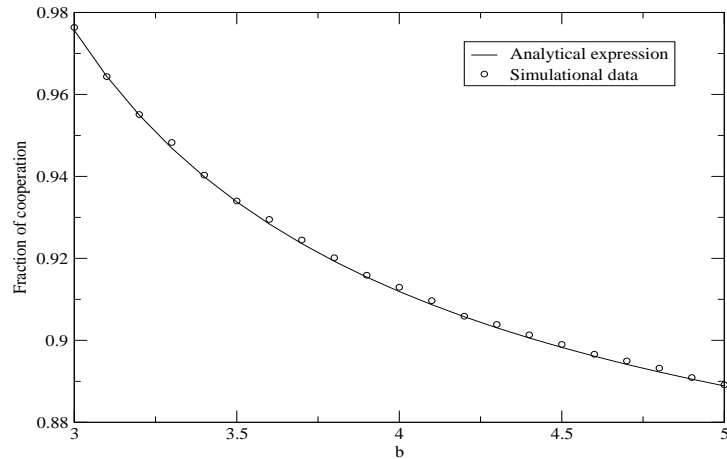

FIG. 5: Fraction of cooperation inside a square of linear size $S = 6$ when $b > 3$ for model A. The continuous line represents the analytical expression and the circles the simulations data.

a square of size S:

$$\rho = 1 - \frac{1}{S} + \frac{-b + \sqrt{b^2 + 8(b-2)} + (S-2)}{(b-2)S^2} \quad . \tag{1}$$

Fig. 5 shows that the simulation and the analytical expression for $S = 6$ agree very well.

Let's analyze the other models. The model B is a trivial one, because since a player has updated one of his confrontation, this confrontation remains the same forever. Since

in the random initial condition we have, in average, 50% of cooperation, there is at the beginning, in average, 25% of each one of the four possible confrontations. So there is an inferior bound for the fraction of cooperators, namely 25%. Since all confrontation are initially equally probable, every player that has more than one (D,C) confrontation, will have a bigger cumulative payoff. The players that are exploited by these exploiters will replace the (C,D)confrontation by a (D,D) one. All these (C,D) confrontations will change to (D,D) ones while this unbalanced exploitations are not neutralized. So the cooperation that remain will be only those that were initially present in (C,C) confrontations, that is, the fraction will be around 25% as it is shown in Fig. 1. Note that after the exploitations are neutralized, they cannot be converted to mutual cooperation. This feature is in contrast with model A, because if a (D,D) confrontation appears in model A, it can be changed to a (C,C) one when both players x and y imitate at the same round other players that are cooperating with them.

Model C does not support the same capability of survival for the cooperators, as can be seen in Fig. 1. The mechanism of defection neutralization is smoothed, because if there is a (C,D) confrontation, the D strategy can be copied into any confrontation and the exploitation has a probability to remain alive. Let's analyze again the cluster shown in Fig. 3. This cluster is no longer stable. If $b < 2$, the cluster will expand. But there is no guarantee of stability. Suppose some player is reached by the expanding cooperation cluster and changes all of his strategies to C but one is still kept with D. Suppose also that this player is surrounded by cooperation. The cluster can go on expanding, but now this D strategy can start a defection invasion from inside the cluster because it will give to that player a bigger payoff. This does not happen in model A. The random replacement does not provide the certainty of neutralizing an exploitation and still keeping cooperation with the others that are cooperating

Further simulations show that if $b \gg 1$ the fraction of cooperation at steady state in model A tends to the same value as that in model B. But this is expected, because if b is huge, every (C,D) confrontation will give to the exploiter a huge advantage and every (C,D) confrontation change to (D,D) one.

It must be mentioned that the steady state fraction of cooperation depends on the initial configuration. The parameter we used to vary the initial conditions is the probability p of having a cooperation. For model A, as the cooperation survival depends on cluster

8

formation, if p is small, there is no cluster of cooperation at the beginning, but if $p > 0.4$ and $1 < b < 3$, the stationary value does not depend on the initial condition. When $b > 3$, cooperation clusters no longer spread out and the stationary values strongly depends on the initial configuration of cooperation reinforcements. After discarding finite size effect, model C does not depend on initial conditions.

Let's now consider another interest point: the robustness of model A. So far we have assumed that the players always judge correctly what is their worst confrontation. But misjudgment is a relevant parameter in real behavior. Instead of always choosing the worst confrontation, now the players can misjudge and choose a confrontation that does not give the worst contribution to the payoff. For this we introduce a misjudgment probability (p_m). Let's state this point clearly. The individuals can update their strategies in every round. For each individual update we introduce the possibility of a random replacement, given by a misjudgment probability p_m. By this definition, with probability $1 - p_m$ the replacement follows the original model A and with probability p_m follows model C. If $p_m = 0$ and $p_m = 1$ model A and model C are recovered, respectively. We say that model A is robust if the fraction of cooperation in the case $p_m \neq 0$ still remain close to the fraction in the case $p_m = 0$ if we vary the p_m values. We say that for some b value model A is more robust than for another b value if the decreasing of the cooperation fraction is lower for the first b value. The robustness analysis shows that there are three typical behaviors that depend on the defector tendency. This defines three b regions: i) $1 \leq b < b_c$, ii) $b_c \leq b < 3$ and iii) $b > 3$. In the first region both models A and C exhibits cooperation. In the second and third region, just model A exhibits cooperation. Note that the fraction of cooperation in the third region is lower that in the second. We choose $b = 1$, $b = 2$, and $b = 4$ as representative values for the three region. The simulational results are shown in Fig. 6. For $b = 1$ model A is indeed robust: up to $p_m = 0.1$ there is no decrease on the cooperation fraction. This means that if $N = 10^4$, even in the presence of 1000 misjudgments in every round, nothing happens to the cooperations. For $b = 2$, model A does not change the fraction of cooperation for p_m values up to $p_m = 0.001$. For $b = 4$ the fraction of cooperation reduces even for $p_m = 0.0001$. Note that for $b > b_c$ in model C the inherent randomness on the replacement reduces drastically the cooperation. This b dependency is also present when some misjudgment introduces randomness on the replacements of model A. So the results supports the conclusion that for low defector tendency model A is robust although for large values of b the randomness has
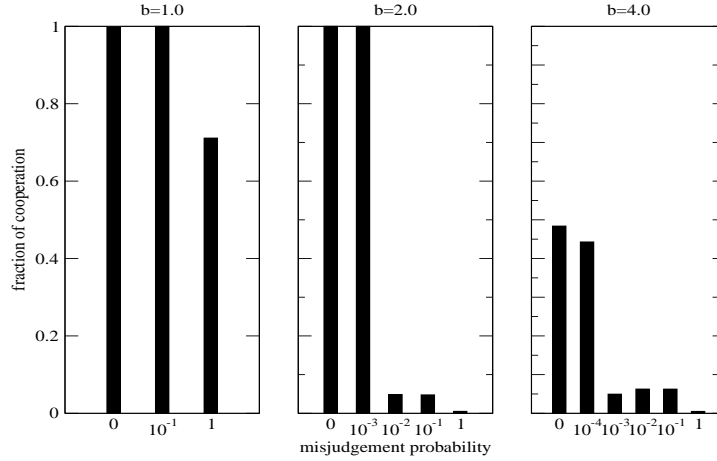
9

FIG. 6: Fraction of cooperation in the square lattice for model A at the steady state with different probabilities of misjudgments. It is shown the results using $b = 1$, $b = 2$ and $b = 4$.

a more eminent role reducing the cooperation fraction.

In order to have a more complete picture of the models with the possibilities of different strategies against different opponents, let's study the one-dimensional lattice with periodic boundary conditions (ring). The ring has $N = 10^4$ nodes and each node is linked only to their two next neighbors. The fraction of cooperation is shown in Fig. 7. Note that for the usual game there is no cooperation and model A is again able to sustain some cooperation. Note that model C is also able to keep some cooperations alive even for huge defector tendency, what does not happen in the square lattice. The influence of connectivity on the cooperation fraction was already noted for the regular network with the usual game [9]. The robustness analysis shows the same b dependency, although the ring is less robust than the square lattice. Fig. 8 shows the robustness simulational results for the ring.

In summary, we have introduced the adoption of multiple strategies by the same player which gives to the players the possibility of cooperating with some opponents and of defecting with others. If the player wants to punish the exploiter but keep the old cooperations, he can use simultaneous multiple strategies with the "replacing the worst confrontation" replacement rule that model A provides. Model A incorporates a true punishment, because the players do not take advantage in imitating the defector strategy, but just decrease the payoff of who wants to be an exploiter. Model A also incorporates the reestablishment of mutual cooperation by not keeping a frozen interaction. The usual game on the square
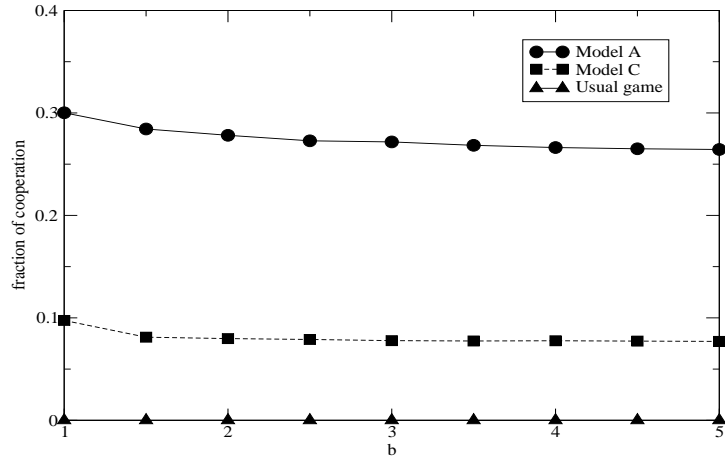
FIG. 7: Fraction of cooperation in the ring. It is shown the behavior at the steady state of the usual game and the studied models A and C for different values of b.
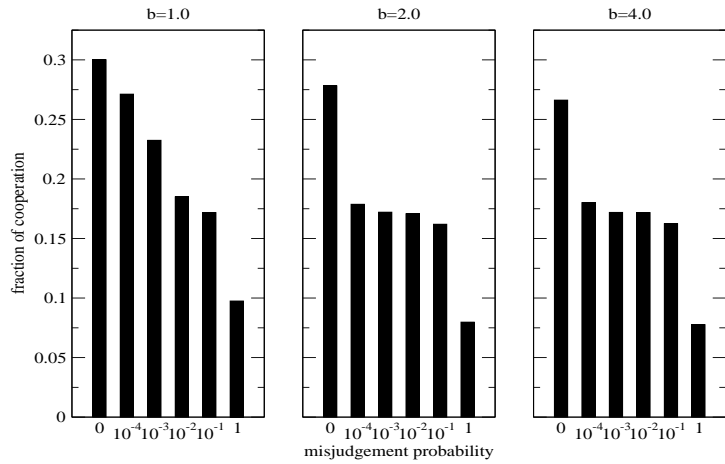

FIG. 8: Fraction of cooperation in the ring for model A at the steady state with different probabilities of misjudgments. It is shown the results using $b = 1$, $b = 2$ and $b = 4$.

lattice, with the same synchronous update rule, provide a smaller fraction of cooperations for low b values and has a transition at a low b value when the cooperators die out. We have shown that model A provides a mechanism in which cooperation dominates if $1 < b < 3$. Moreover even for huge defection tendency ($b > 3$) model A ensures cooperation survival. For the ring topology model A also exhibited capability of sustaining cooperation. Finally we showed that model A is robust against misjudgments, at least for low defector tendency, for both lattices. We stress that this simple rule can be easily set in other game contexts

11

what opens the possibility of further researches.

* Electronic address: wardil@fisica.ufmg.br

† Electronic address: jaff@fisica.ufmg.br

---

[1] H. Ohtsuki, C. Hauert, E. Lieberman, and M. A. Nowak, Nature **441**, 502 (2006).

[2] M. A. Nowak, in ''Evolutionary Dynamics: Exploring the Equations of Life'' (Harvard University Press, 2007).

[3] M. A. Nowak, Five rules for the evolution of cooperation, Science **314**, 1560 (2006).

[4] H. Ohtsukia and M. A. Nowak,Direct reciprocity on graphs, J. Theor. Biol. **247**, 462 (2007).

[5] M. A. Nowak and K. Sigmund, Evolution of indirect reciprocity, Nature **437**, 1291 (2005).

[6] G. Szaboa and G. Fathb, Evolutionary games on graphs, Physics Reports **446**, 97 (2007).

[7] M. A. Nowak, S. Bonhoeffer and R. M. May, More spatial games, Int. J. Bifurcat. and Chaos **4**, 33 (1994).

[8] G. Szab, J. Vukov and A. Szolnoki, Phase diagrams for an evolutionary prisoner's dilemma game on two-dimensional lattices, Phys. Rev. E **72**, 047107 (2005).

[9] F.C Santos and J. M. Pacheco, Scale-free networks provide a unifying framework for the emergence of cooperation, Phys. Rev. Lett. **95**, 098104 (2005).

[10] S. Assenza, J. Gomez-Gardenes and V. Latora, Enhancement of cooperation in highly clustered scale-free networks, Phys. Rev. E **78**, 017101 (2008).

[11] M. Nakamaru, H. Matsuda and Y. Iwasa, The evolution of cooperation in a lattice! Structured population, J. Theor. Biol. **184**, 65 (1997).

[12] J. Vukov, G. Szab and A. Szolnoki, Evolutionary prisoner's dilemma game on Newman-Watts networks, Phys. Rev. E **77**, 026109 (2008).

[13] E. Lieberman, C. Hauert and M. A. Nowak, Evolutionary dynamics on graphs, Nature

433, 312 (2005).

[14] J. Gomez-Gardees, M. Campillo, L. M. Flor and Y. Moreno, `Dynamical organization of cooperation in complex topologies`, Phys. Rev. Lett. **98**, 108103 (2007).

[15] C. Hauert, A. Traulsen, H. Brandt, M. A. Nowak and K. Sigmund, `Via freedom to coercion:  The emergence of costly punishment`, Science **316**, 1905 (2007).

[16] R. Boyd, H. Gintis, S. Bowles and P. J. Richerson, `The evolution of altruistic punishment`, P. Nat. Acad. Sci. USA **100**, 3531 (2003).

[17] F. C. Santos, M. D. Santos and J. M. Pacheco, `Social diversity promotes the emergence of cooperation in public good games`, Nature **454**, 213 (2008).

[18] M. A. Nowak and R. M. May, `Evolutionary games and spatial chaos`, Nature **359**, 826 (1992).

[19] P. Langer, M. A. Nowak and C. Hauert, `Spatial invasion of cooperation` J. Theor. Biol. **250**, 634 (2008).