

Effects of punishment in a mobile population playing the prisoner's dilemma game

Daniel R. Amor* and Joaquim Fort

Complex Systems Lab, Departament de Física, Universitat de Girona, E-17071 Girona, Catalonia, Spain

(Received 4 April 2011; revised manuscript received 13 October 2011; published 20 December 2011)

We deal with a system of prisoner's dilemma players undergoing continuous motion in a two-dimensional plane. In contrast to previous work, we introduce altruistic punishment after the game. We find punishing only a few of the cooperator-defector interactions is enough to lead the system to a cooperative state in environments where otherwise defection would take over the population. This happens even with soft nonsocial punishment (where both cooperators and defectors punish other players, a behavior observed in many human populations). For high enough mobilities or temptations to defect, low rates of social punishment can no longer avoid the breakdown of cooperation.

DOI: [10.1103/PhysRevE.84.066115](https://doi.org/10.1103/PhysRevE.84.066115)

PACS number(s): 89.75.Fb, 89.65.-s, 87.23.Kg

I. INTRODUCTION

The appearance and maintenance of cooperation is one of the most important enigmas set out by evolutionary biology [1]. A problem of utmost importance is that while an individual can benefit from mutual cooperation, it can often do even better by exploiting the cooperative efforts of others (and this, in turn, tends to destroy cooperation) [2]. Evolutionary game theory has proved to be a major formalization tool in this context, so different games have been used in many theoretical and experimental works, and the subject has rapidly jumped to other domains such as statistical physics and mathematics [3–7]. In order to model biological systems and human behavior [1,2], the prisoner's dilemma (PD) game has become especially well known and studied because in this game the best strategy for single players becomes the worst one for the community (see below).

The PD is a game between two individuals, where each one can choose from whether to cooperate (C) with or to defect (D) from his opponent. The corresponding payoff matrix of an interaction between two players is given by

$$\begin{pmatrix} R & S \\ T & P \end{pmatrix}, \quad (1)$$

where R denotes the payoff each player receives if both cooperate, S is what a cooperator receives when he is exploited by a defector (who gains T in this particular interaction), and P is the payoff for two defectors playing one against another. The payoffs are ordered as $T > R > P \geq S$, so rational players would always play defection (because it pays more regardless of the opponent's decision). But this rationality makes both players receive only P instead of the reward $R > P$, which both players would have accumulated by cooperating, hence the dilemma. With two individuals destined never to meet again, the only rational strategy is to defect [2].

The main idea of evolutionary game theory hinges on individuals playing multiple rounds. This simulates real-world interactions better than one-shot games and in some settings leads to the survival of cooperation. In 1981 the Axelrod computer tournaments [2] analyzed multiple strategies competing in the repeated PD game, with individuals remembering

their previous interactions. They found defection is not the only stable strategy. In contrast, some strategies lead to higher average scores than defection (the most successful strategy was *tit for tat*, i.e., simply cooperating on the first move and then doing whatever the other player did in the preceding move). These results stimulated a wealth of work that still continues today [8–10]. In 1992 Nowak and May [11] introduced the repeated PD game into a simulated population with its individuals bound to lattice sites of a two-dimensional spatial array. They found cooperators and defectors both persist indefinitely (in shifting clusters), without the need to assume the use of complicated strategies, not even that any individual remembers previous interactions (each player was either C or D, and after each round each lattice site was occupied by the player with the highest payoff among the previous owner and its neighbors). A huge amount of work has been undertaken ever since concerning evolutionary games on graphs [3], exploring many diverse combinations of realistic network topologies (link dynamics) and strategy update rules that lead to the survival of cooperation [12–20].

In a recent paper, Meloni *et al.* explored the effects of mobility in a population of PD players [21]. Moreover, the authors introduced an innovative kind of migration. In spite of the number of recent models that take into account migration processes in networks [16,20,22–24], Meloni *et al.* noted that the continuous motion of individuals (in contrast to the discontinuous jump of individuals bound to lattice sites of a grid) was an unexplored situation of practical relevance. It could also be important in several applications, e.g., in designing cooperation-based protocols for wireless devices such as robots [25] and in modeling the dynamics of interacting human populations with different cultural traits, e.g., in prehistoric transitions [26]. Therefore, in this paper we will deal with systems with continuous mobility, as proposed in Ref. [21]. Such systems exhibit only two stable attractors: those in which the whole population either cooperates or defects [21].

In addition to spatial structure, there exist other mechanisms that provide important aids to cooperators in their fight against free riders [13]. In this paper, we focus on the effects of punishing in a population with continuous motion. Punishment is a negative incentive by which some players (punishers) impose fines upon some of their coplayers. Punishers usually pay a cost to punish [27]. Why do then individuals choose

*daniel.rodriguez@udg.edu

to punish others [28,29]? Although there is not an ultimate answer available, it is widely accepted that emotions and moral sentiments play an important role in human decisions that can go beyond the maximization of their income [9,10,28,30–34]. Similarly, some humans also reward their cooperative partners [35]; thus several papers have dealt with positive incentives. Some recent models have shown, for example, that the appropriate dose of the carrot [36] or the convenient combination with the stick when agents are opportunistic [37] can notably enhance the triumph of cooperation.

Punishment can be very costly in pairwise interactions; thus in real life punishment is usually repressed by institutions holding law and order. But some pairwise interactions with punishment obviously exist in real life, both in small-scale and large-scale societies [27]. Still, in agreement with the relative rarity of pairwise punishment, traditionally, the cooperation-enhancing effects of punishment have been often analyzed within the public goods game [9,31,38,39]. Very recently, this n -person game has provided the scenario where the following effects have been studied: (i) the strategy D punishes D (instead of the usual C punishes D) also helps to defeat free riders [5], (ii) small mutation rates accelerate the spreading of costly punishment [40], and (iii) spatial structure is responsible for several kinds of coexistence between cooperators, moralists (cooperators who punish defectors), defectors, and immoralists (defectors who punish other defectors) [41]. In this paper we have opted for a simpler setup and considered pairwise interactions only. Although less commonly used, two-person games also have been employed in punishment studies, especially in experimental work. One of the most representative examples is the ultimatum game [30,42], where the rejection of an offer is indeed a kind of costly punishment. Rewarding and punishing human behavior has also been tested by using sequential PD games (often called gift exchange or trust games) [43]. In other cases, the payoff matrix of the PD game has been modified in order to include three types of players: cooperators, defectors, and punishers [44]. Moreover, the use of pairwise interactions is also present in models with incentive strategies [36].

Punishment is often considered a different strategy from pure cooperation or pure defection [5,40,41]. In this paper, we do not consider punishers to have a different strategy than other players. Players are either cooperators or defectors. After a given interaction, we introduce a probability to punish the corresponding coplayer. This is a simple way to model that a certain portion of the interactions is followed by negative incentives but players do not become obsessed in punishing every partner they are not comfortable with (since it could be extremely costly for themselves). Besides social punishment (C punishes D) we will also analyze what we call nonsocial punishment (C punishes C, D punishes C, or D punishes D), as suggested by recent experiments in many human populations around the world [33].

To summarize, in this paper we introduce social and nonsocial punishment effects into a model where players can move continuously in a two-dimensional (2D) world. Moreover, no punishment strategies are considered; instead, we deal with probabilities that different interactions between agents are followed by punishment. In particular, we find that social punishment helps to maintain cooperation in extreme

environments, even with such a high mobility or temptation to defect that the system would otherwise be completely invaded by defectors. We shall also find that the benefits of social punishment are remarkable, even when some degree of nonsocial punishment is present in the game.

This paper is organized as follows. In Sec. II we present the main features of the model (including mobility rules, network of interactions, evolutionary dynamics, social punishment, and nonsocial punishment). Section III is devoted to explaining the main effects of considering a nonmobile population in our model. In Sec. IV we present the general results of our simulations (phase diagrams showing the dependence on the relevant parameters of the model). Finally, our concluding remarks are presented in Sec. V.

II. THE MODEL

In this section we explain the rules that drive the evolution of the system. We study a population of $N = 1000$ individuals living in a square plane of size L . As in Ref. [21], periodic boundary conditions are imposed at the ends of the square (this makes the square equivalent to a toroidal surface, thus avoiding border effects). Simulations are governed by three groups of rules concerning the motion rules, the network of interactions, and the evolutionary dynamics. The first two groups of rules we use here have the same properties of the original model [21], but the evolutionary dynamics are notably different because we introduce a probability of punishing the opponent. Simulations perform sequentially the three sets of rules at each time step t . In the following three sections, we detail these three types of rules in the model.

A. Motion rules

At the beginning of each round, every player moves a fixed distance in a random direction. Hence, the position of a given individual i ($i = 1, 2, \dots, N$) is changed as

$$\mathbf{x}_i(t+1) = \mathbf{x}_i(t) + \mathbf{v}_i(t), \quad (2)$$

where $\mathbf{v}_i(t) = [v \cos \theta_i(t), v \sin \theta_i(t)]$ is the speed of the player i . The direction of the speed is determined randomly as

$$\theta_i(t+1) = \eta_i, \quad (3)$$

where η_i are N independent random variables chosen at each time with uniform probability in the interval $[-\pi; \pi]$.

As in Ref. [21], we consider that the module of the speed v is constant for all agents, so v is one of the relevant parameters of the system. At $t = 0$, both a random position in the square and a random initial direction of movement is assigned to every player.

B. Network of interactions

After movement, every player establishes its network of interactions. An agent i will consider the player j as his or her neighbor if j is within a certain radius of interaction r . First, the Euclidean distance d_{ij} between players is computed. Then, if $d_{ij} < r$, we say that i and j are neighbors (and will interact following the rules in the next section). Without loss of generality, in all the simulations presented here we have set

$r = 1$. The instant network of contacts can be defined by a graph that links its nodes (individuals) with the current web of neighbors. Note that such a graph changes at every round t due to the motion and neighborhood updating of the agents. The mean degree of the graph is $\langle k \rangle = \rho \pi r^2 = \rho \pi$, where $\rho = N/L^2$ is the population density (see the work by Meloni *et al.* [21] for this and other topological features of the graph). The dependence of the system properties on the value of ρ was already analyzed in Ref. [21], Fig. 2 [45]. Therefore, in this paper we have used the value $\rho = 1.30$ (as in Ref. [21], Figs. 1, 3, and 4) [45].

C. Evolutionary dynamics

In this section we summarize the rules regarding the interaction between agents and their strategy updates.

At each time step t , every individual plays once a PD game with each of his neighbors. The payoff matrix of the game is presented in Eq. (1). As usually done in recent studies (see, e.g., Ref. [21]), we choose $R = 1$, $P = S = 0$, and $T = b > 1$. The payoffs obtained from the multiple interactions are accumulated by each player during the round. In the first round ($t = 0$), the two possible strategies (C or D) are equally distributed among the population. Although in other works each individual can play different strategies against his multiple opponents, following Ref. [21] here we consider the simple case where agents can only choose to cooperate with or to defect from all of his neighbors. This will make it easier to focus our attention on the effect of punishment.

In contrast to Ref. [21], after every PD interaction is performed, we consider that each player has the opportunity to punish his opponent if he is not satisfied with the outcome of the game. In order to simulate this circumstance, in Sec. IID we will allow cooperators to punish defectors with probability p_s , representing social punishment, i.e., the intention to promote a cooperative society where defectors have a bad reputation. In Sec. IIE, we will extend the model to allow a nonsocial type of punishment: with probability p_a , not only will defectors punish their coplayers (irrespective of their strategy) but cooperators will also act against other cooperators (representing other human motives like revenge or preventive strikes [9,33]). When a punishment action occurs (either in the social or in the nonsocial case), the punisher pays a cost of 1 unit of his accumulated payoff in order to reduce the payoff of the punished player by 3 units (this rate has been used in previous human experiments, e.g., in Ref. [33], although harder rates where the punished player loses 4 units instead of 3 have been also studied [9]).

Finally, after all games have been played and the corresponding punishments have been executed, it is time for the agents to update their strategies. As in Ref. [21], every agent will compare his own payoff with that of a randomly chosen neighbor, and then he will decide whether he keeps playing the same strategy in the next round or not, as follows. If individual i and the chosen neighbor j use the same strategy, nothing happens. In the opposite case, and provided that j has accumulated higher gains in the current time step, individual i will adopt the strategy of j with the following probability [21]:

$$\Pi_{ij} = \frac{P_j - P_i}{\max\{k_j, k_i\}b}, \quad (4)$$

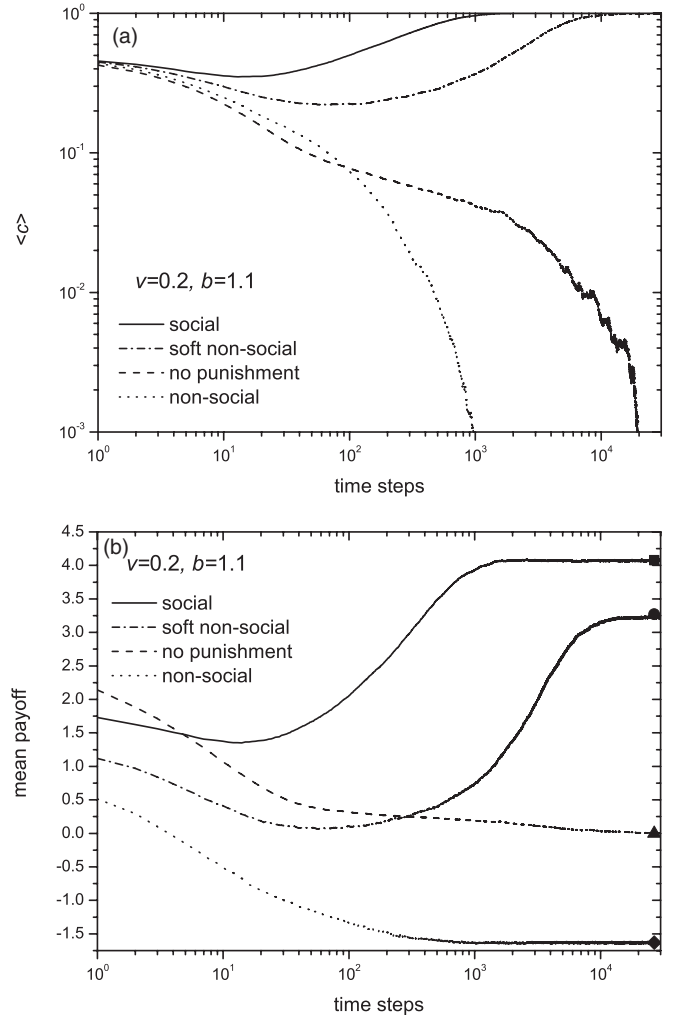


FIG. 1. Average level of cooperation $\langle c \rangle$ (defined as the average fraction of cooperators) as a function of time (or rounds of the game). Simulations have been performed for a population of $N = 1000$ individuals, $\rho = 1.30$, $b = 1.1$, and $v = 0.2$. The solid line depicts the evolution for a population where social punishment is allowed (with $p_s = 0.1$). The dash-dotted line shows the evolution of the same population in a soft nonsocial environment, where $p_s = 0.1$ and $p_a = 0.05$. The dashed line stands for the case where punishment is not allowed at all. Finally, the dotted line shows a nonsocial system where $p_s = 0.1$ and $p_a = 0.1$. Results are averaged for 100 simulations. (b) Average payoff per player as a function of time for the same cases presented in Fig. (a). The symbols depict the asymptotic theoretical value for the average payoff. The square is for the social case, the circle stands for the soft nonsocial system, the triangle is for the case without punishment, and the rhombus is for the nonsocial system.

where P_j and P_i are the payoffs accumulated by players j and i , respectively. In Eq. (4) k_j and k_i stand for the instantaneous number of neighbors (i.e., number of players within the circle of radius r) that players j and i have, respectively. This updating process is done synchronously for all individuals in the system. Finally, the payoffs of all individuals are reset to zero, and the next round can start as explained above.

In our simulations, we have found that the system has only two stable states, as in the original model without

punishment [21]. These two attractors are reached when either cooperation or defection is played by all agents. Moreover, high values of v or b destroy cooperation when no punishment is considered, in agreement with Ref. [21].

D. Social punishment

In order to introduce social punishment into our model, we have allowed cooperators to punish altruistically defectors as a separate action after the PD interaction. We do not assume that cooperators always punish their opponents, just that there is a probability of punishing socially after each cooperator-against-defector (C-D) interaction. With probability p_s the cooperator will incur a cost of 1 payoff unit to diminish the defector payoff by 3 units.

Figure 1 sheds some light on the implications of such a behavior acting directly against freeloaders. Figure 1(a) shows the evolution of the average level of cooperation (c), defined as the average fraction of cooperators, for $v = 0.2$ and $b = 1.1$. The dashed line shows how a population where no punishment is allowed evolves to a completely defector system. However, the population can be driven to the maximum cooperation fraction if some degree of social punishment is present (solid line). Note that only a 10% of the C-D interactions are followed by a punishing action ($p_s = 0.1$) in this case. Hence, we have seen that a relatively low dose of social punishment can introduce critical advantages for cooperators.

E. Nonsocial punishment

Anger, revenge, and preventive strikes are only a few motives [33] that can inspire humans to punish, even if this action plays against the benefits of the community and the punisher himself. In this section, we will study the effects of this kind of punishment on the evolution of the population. We thus consider a type of punishment different from that explained in Sec. II C. We will call it nonsocial punishment.

Analogously to the social punishing probability p_s introduced above, we define the probability of punishing nonsocially as p_a . While p_s is the probability that a cooperator punishes a defector after a C-D interaction, p_a applies to any interaction different from a C-D interaction (see Sec. II C). This agrees with the definition of antisocial punishment in the paper of Herrmann *et al.* [33] that has especially inspired our work, but in other papers antisocial punishment is strictly defined as the punishment of cooperators [38,46]. Because of this, we use the term nonsocial rather than antisocial. The dotted line in Fig. 1(a) depicts the evolution of the population in a highly nonsocial environment where $p_s = p_a = 0.1$: no matter which players are engaged, 10% of the interactions are followed by a punishing action. We can see how in this case the system favors defection, as this strategy is quickly adopted by the whole population. Therefore, nonsocial punishment works directly against the benefits of social punishment, as it would be expected intuitively [compare with the social case represented by the solid line in Fig. 1(a)].

Although both social and nonsocial punishment are found in human communities, the first one is perceived to be more rational. Imposing a fine on someone who contributes less than you (social punishment) is easily justified because you find

her or his behavior unfair. In contrast, punishing nonsocially means that you attack someone that has conducted himself (at least) as properly as you in the game; hence you do not tolerate your own behavior. As a result, human players are usually found to invest higher amounts in social punishment [46,47]. To simulate this effect, we have performed simulations with $p_s \neq p_a$. Note that this does not mean the total number of social punishing actions in the system exceeds the nonsocial one (because the first one is strictly dependent on the number of C-D interactions). The dash-dotted line in Fig. 1(a) depicts the evolution for $p_s = 0.1$ and $p_a = 0.05$. In this case, we find another victory of cooperation. As in the system containing only social punishers (solid line), the evolution starts with a soft fall of cooperators, and then (after approximately 100 rounds) the population starts a constant progression to the maximum cooperation degree. The transient decline of cooperation is related to the defeat of those cooperators who were initially placed in a neighborhood with plenty of defectors. However, aggregates (clusters) of cooperators are able to survive and flourish in this notably changing ($v = 0.2$) environment, provided that the benefits of social punishment exceed those of nonsocial punishment. Indeed, the evolution of the nonsocial system [dotted line in Fig. 1(a)] is an example where these benefits are insufficient to lead the population to cooperation. Although here we focus on systems with mobile players, an analog to Fig. 1(a) for $v = 0$ is included in Sec. III.

It could be sensible to think that the soft nonsocial case [dash-dotted line in Fig. 1(a)] models a system where social punishment dominates and, consequently, could be analogously represented by a system containing only social punishment. However, this case does not present the same features of the soft nonsocial system. Indeed, whereas in both cases the final state is a population composed by cooperators exclusively, the mean payoff per individual is notably lower in the nonsocial case. To see this, the mean payoff per individual in the final state can be predicted as

$$w = k(f_c R - p_a - 3p_a), \quad (5)$$

where f_c is the fraction of cooperators in the final stable state (in our system, either 0 or 1). The mean degree of the graph in all of our simulations is $k = \rho\pi = 4.08$. In Eq. (5), the mean number of neighbors that an individual has k is multiplied by the reward of an interaction R provided that the final state is all cooperators (i.e., $f_c = 1$) minus the mean costs of the punishments performed or received by the individual ($-p_a$ and $-3p_a$, respectively). Note that the social punishment probability p_s does not appear in Eq. (5) because C-D interactions are absent in the asymptotic state, in which all players are either C or D.

Figure 1(b) shows the evolution of the mean payoff per individual [for the same cases presented in Fig. 1(a)], and the symbols stand for the final-state values predicted by Eq. (5). The four curves tend asymptotically to the analytical value for the final stable state. In all the cases, the mean payoff shows a rapid decay during the first rounds of the game (every simulation ends or smooths this decay at a different moment before 100 rounds). Again, this transient effect corresponds to the defeat of those cooperators who have been initially placed in a neighborhood with many defectors. After this transient, the

surviving cooperators are clustered and resist more efficiently the exploitation of the defector population. In the no-punishers and nonsocial cases (dashed and dotted lines, respectively), the environment is so extreme that even clusters of cooperators cannot endure the exploitation, and the population continues its fall toward defection. When the final state is reached, the mean payoff is 0 in the no-punishers case, as there is no one that contributes any amount in the game. The situation is even worse in the nonsocial case, as nonsocial punishment makes the individuals accumulate negative payoffs. In contrast, the negative effect of social punishment on the mean payoff decreases as the number of C-D interactions decays. Hence the social punishers system (solid line) reaches the maximum payoff level available when the system is composed only of cooperators, namely, $w = k = 4.08$ (because $p_a = 0$ for

social punishment, and we have chosen $R = 1$ as in, e.g., Ref. [21]). Therefore, social punishment is very effective for promoting cooperation. Finally, the soft nonsocial scenario allows the survival and maintenance of cooperation [as shown in Fig. 1(a)], but the average payoff is lower [compare the dash-dotted line with the solid line in Fig. 1(b)]. Below, we will present several figures that analyze the relevant parameters of the model. Although it will not be shown, the computational and the analytical results of Eq. (5) for the mean payoff agree as well as in Fig. 1(b) for all the simulations presented in this paper (when the corresponding steady state is reached).

Some of the major evolutionary characteristics of the system, detailed above, can be identified by looking a time series of snapshots for a single simulation. Figure 2 captures the snapshots at four times for a simulation with $v = 0.1$,

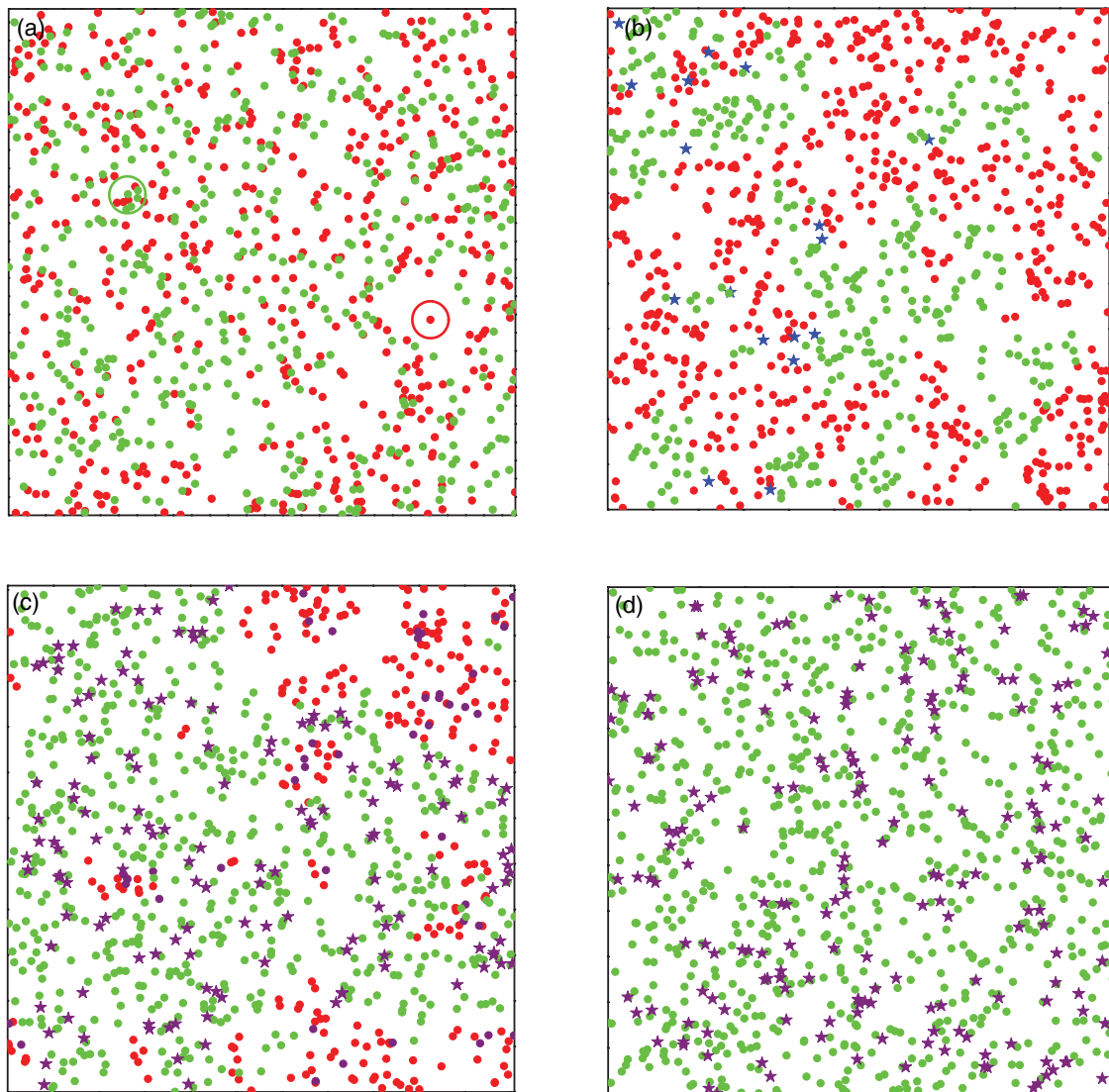


FIG. 2. (Color online) Spatiotemporal evolution of a soft nonsocial population where $p_s = 0.1$ and $p_a = 0.05$. This specific simulation has been performed for a population of $N = 1000$ individuals, $\rho = 1.30$, $b = 1.1$, and $v = 0.1$. Green (light gray) dots correspond to cooperator players, and red (dark gray) dots correspond to defectors. (a) Snapshot at $t = 0$. The population has been randomly distributed in the two-dimensional square. The two circles show the radius of interaction of a single cooperator [green (light gray) circle] and a single defector [red (dark gray) circle]. (b) Snapshot at $t = 100$. Blue stars correspond to cooperators that punished socially in the last round. (c) Snapshot at $t = 1000$. Purple stars (dots) correspond to cooperators (defectors) that punished nonsocially in the last round. (d) The stable all-C state is reached at $t = 5693$. Purple stars show the cooperators that still punish their partners when the whole population is cooperator.

$b = 1.1$, and the soft nonsocial environment used in Fig. 1 (i.e., $p_s = 0.1$ and $p_a = 0.05$). At $t = 0$ [Fig. 2(a)], both cooperator [green (light gray) dots] and defector [red (dark gray) dots] players are randomly placed in the system, with the same fraction ($\langle c \rangle = 0.5$). In the left part of Fig. 2(a), we show a green (light gray) circle that represents the radius of interaction of a specific cooperator player who has been initially surrounded by several neighbors [during the first round, he will play the PD game with the eight coplayers who lie within the green (light gray) circle, with half of them being cooperators in this case]. In contrast, some other players may start the game completely isolated, as is the case of the defector in the center of the red (dark gray) circle [right in Fig. 2(a)]. Because of mobility, the number of connections for each player will change in time. Thus simulations with no mobility (i.e., $v = 0$) are essentially different from the ones where $v > 0$. Indeed, isolated players [like the one in the center of the red (dark gray) circle in Fig. 2(a)] will never play with other individuals in a simulation where $v = 0$, and consequently, the system will not reach a final all-C or all-D state (see Sec. III for further details of the case without mobility).

Figure 2(b) shows the state of the system after 100 iterations. At this time, the cooperation fraction has decreased to $\langle c \rangle \simeq 0.35$ due to the defeat of those cooperators initially placed in hostile neighborhoods. As in Fig. 2(a), red (dark gray) dots correspond to defectors. In Fig. 2(b) we explicitly show those players who have punished socially at least once in the last round; hence cooperators have been divided in two types: green (light gray) dots show the players who have not punished socially in the last round, and blue stars indicate cooperators who punished at least one defector neighbor. We can see that the fraction of cooperators that punishes socially in a given round [blue stars in Fig. 2(b)] is relatively small. Nevertheless, we will see how the benefits provided by social punishment exceed the handicap due to nonsocial punishment in this system, and the all-C state will be finally reached. In Fig. 2(c) we can see the state of the system at $t = 1000$. Here we explicitly identify the players who used nonsocial punishment in the last round: purple stars correspond to cooperators who punished at least one other cooperator, and purple (darkest gray) dots correspond to defectors who punished at least one coplayer. In this case, the fraction of cooperation is $\langle c \rangle \simeq 0.7$, so clusters of cooperators have spread since $t = 100$ [compare to Fig. 2(b)]. Finally, Fig. 2(d) shows the final all-C state reached after 5693 iterations. In order to explicitly show that nonsocial punishment still takes place in the game, purple stars indicate those cooperators who still punish their cooperator partners. This has a direct effect on the mean payoff attained by players, as we have shown above [see the discussion of Fig. 1(a)].

III. THE KEY ROLE OF MOBILITY

In the model we have analyzed, a certain mobility rate of the players is present in most of the simulations. In the original model where punishment was not considered [21], an extended study concerning all the relevant parameters (including the mobility rate) of the corresponding model was presented. Roughly summarizing the conclusions by Meloni *et al.* [21], the authors noticed cooperation could flourish in their model

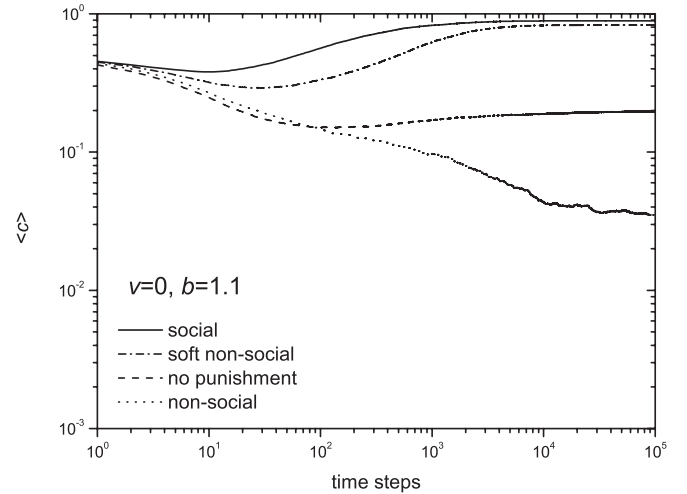


FIG. 3. Average level of cooperation $\langle c \rangle$ as a function of time for a population without mobility ($v = 0$) for $N = 1000$ individuals, $\rho = 1.30$, and $b = 1.1$. The solid line depicts the evolution for a population where social punishment is allowed (with $p_s = 0.1$). The dash-dotted line shows the evolution of the same population in a soft nonsocial environment, where $p_s = 0.1$ and $p_a = 0.05$. The dashed line stands for the case where punishment is not allowed at all. Finally, the dotted line shows a nonsocial system where $p_s = 0.1$ and $p_a = 0.1$. Results are averaged over 100 simulations.

when the parameters b and v were not too high. Moreover, they pointed out that the mobility led the system to only two stable states (i.e., all cooperators or all defectors). In contrast, when $v = 0$ the graph corresponds to a random geometric graph [48], so the stabilization of a population containing a mixture of the two strategies is expected.

The question of how punishment could affect the system when $v = 0$ arises. Thus in Fig. 3 we have rerun our simulations in Fig. 1(a) for a population that does not move ($v = 0$). In this case, clusters of cooperators manage to stabilize and survive in the system if no punishment is applied (the dashed line in Fig. 3 shows that the average cooperation in the system is above 20%). However, social punishment strongly supports cooperators in this immobile system, as shown by the solid line in Fig. 3 (the cooperation fraction reaches 90% of the population in this case, where $p_s = 0.1$ and $p_a = 0$). Furthermore, the final outcome is much the same when the frequency of nonsocial punishment is increased up to $p_a = 0.05$ (the dash-dotted line depicts this soft nonsocial scenario, where the stable cooperation fraction is also high). Nevertheless, the stabilization of this latter case (dash-dotted line) comes later than in the social case (solid line) and remains at lower cooperation levels at any time. These results are no longer reached if we explore a nonsocial system where $p_s = p_a = 0.1$ (dotted line in Fig. 3) since in this case defection is practically extended to the whole population (the final cooperation fraction is above 3%).

If we compare the results of Fig. 1(a) ($v = 0.2$) with those presented in Fig. 3 ($v = 0$), we can draw similar conclusions regarding punishment. In both cases, social punishment helps cooperation to flourish. But this conclusion breaks down if sufficiently strong nonsocial punishment is present [in both Figs. 1(a) and 3, the soft nonsocial case with $p_a = 0.05$ leads to

generalized cooperation, but the nonsocial case with $p_a = 0.1$ does not]. Furthermore, the comparison of Figs. 1(a) and 3 suggests that the effect of mobility can be very important in the case of nonsocial punishment. In the case of nonmobile and nonpunishing populations (dashed line in Fig. 3), the final stable state permits a coexistence of cooperators and defectors. The final fraction of cooperators depends on the value of b . While cooperators are able to survive in clusters for moderate values of b (as is the case represented by the dashed line in Fig. 3), only a few isolated cooperators remain in an almost all-defector population for high values of b (not shown). In contrast, in the case of mobile populations, the value of v influences the final outcome for a given value of b . Here we have shown how clusters of cooperators that survive in nonmobile populations (dashed line in Fig. 3) are not able to resist in a population with $v = 0.2$ [dashed line in Fig. 1(a)]. In this case the final population is composed exclusively by defectors. However, the system would have turned into a full cooperation state if we had chosen a slower velocity of $v = 0.01$ for the same value of $b = 1.1$ (see, for example, Fig. 1 in Ref. [21]).

In this paper, a huge number of configurations of the parameters v , b , p_s , and p_a has been tested. Generally, we have focused our work on mobile environments with $v > 0$. However, the case $v = 0$ has been simulated and analyzed for all of the environments presented above (even if not shown in the figures). In every case, the conclusions from the simulations with $v = 0$ were very close to the conclusions from those with low speeds like $v = 0.01$ [and sometimes much higher values like $v = 0.2$, such as those presented above by comparing Figs. 1(a) and 3]. In Figs. 4, 5, and 6 we have taken care that the minimum speed presented does not sensibly differ from the results obtained for $v = 0$.

IV. GENERAL RESULTS

This section is devoted to exploring the limits of punishment-enhanced cooperation. Hence, we will extend the results in order to analyze all of the relevant parameters of the model.

In addition to N and ρ (see Sec. II B), the model in Ref. [21] has two relevant parameters, namely, the mobility rate v and the temptation to defect b . Here we have extended that model to include punishment, so we will analyze the role of the parameters p_s and p_a in addition to v and b .

Figure 4 shows a phase diagram where the effects of the mobility v and the temptation to defect b are explored (for the same sets of values for p_s and p_a as in Fig. 1). The dashed line corresponds to the case where punishment is not present (i.e., $p_s = p_a = 0$). The region where simulations end up with a population of all cooperators falls below the dashed line (whereas an all-D state is found above the dashed line). This limit agrees well with the results presented in Ref. [21]. The solid line in Fig. 4 indicates the frontier between cooperation and defection when social punishment is introduced with probability $p_s = 0.1$. Comparing the dashed line with the solid line, it follows that social punishment clearly expands the parameter region where cooperation is reached. When no punishment is allowed in the game (dashed line), the maximum temptation to defect that cooperation can endure is slightly

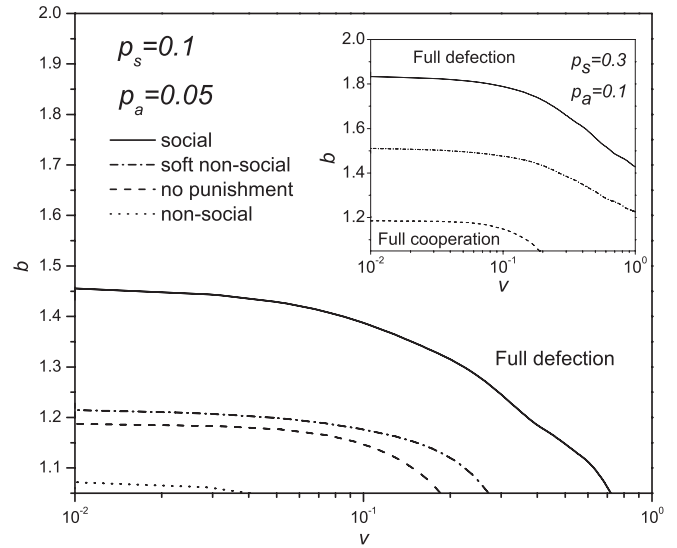


FIG. 4. Phase diagrams for different versions of punishment as a function of v and b . Results have been obtained for a population of $N = 1000$ individuals, with $\rho = 1.30$. The solid line depicts the phase transition for a population where social punishment is allowed (with $p_s = 0.1$). The dash-dotted line corresponds to the case of a soft nonsocial environment, where $p_s = 0.1$ and $p_a = 0.05$. The dashed line stands for the phase transition where punishment is not allowed at all. Finally, the dotted line corresponds to a nonsocial system where $p_s = 0.1$ and $p_a = 0.1$. In the inset, the lines depict the transition phases for the following cases: the solid line corresponds to $p_s = 0.3$, the dash-dotted line corresponds to $p_s = 0.3$ and $p_a = 0.1$, and the dashed line corresponds to the case $p_s = p_a = 0$. No full cooperation phase was found for the nonsocial case where $p_s = p_a = 0.3$.

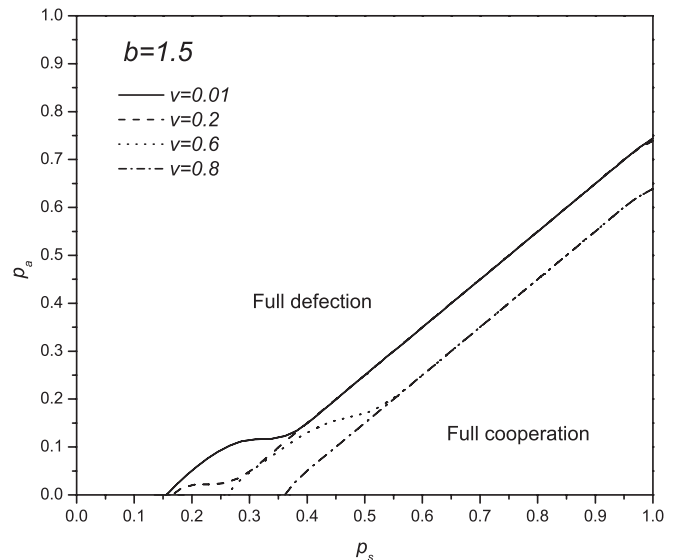


FIG. 5. Phase diagrams for different mobility rates as a function of p_s and p_a in the specific case with $b = 1.5$. Results have been obtained for a population of $N = 1000$ individuals, with $\rho = 1.30$. The solid line depicts the phase transition when the mobility of the agents is $v = 0.01$, the dashed line stands for the case $v = 0.2$, the dotted line is for $v = 0.6$, and the dash-dotted line represents the case of $v = 0.8$.

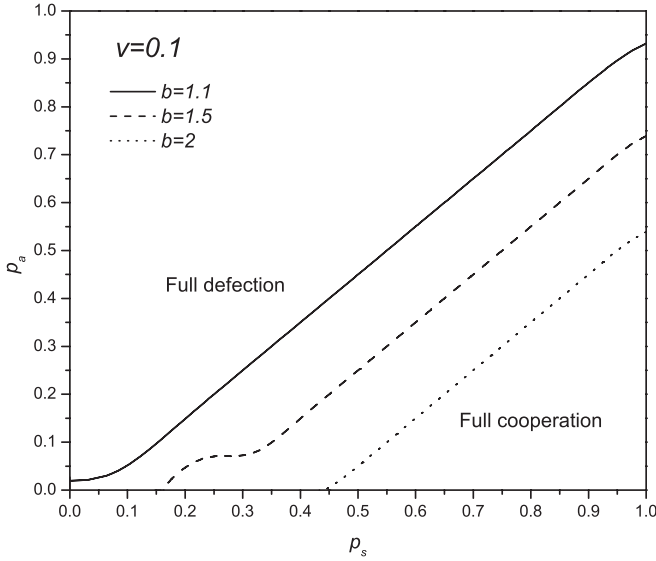


FIG. 6. Phase diagrams for several temptation to defect values as a function of p_s and p_a and $v = 0.1$. Results have been obtained for a population of $N = 1000$ individuals, with $\rho = 1.30$. The solid line depicts the phase transition when the temptation to defect is $b = 1.1$, the dashed line stands for the case $b = 1.5$, and the dotted line presents the case $b = 2.0$.

under the value $b = 1.2$. In contrast, when social punishment is introduced (solid line), the transition from cooperation to defection occurs at a maximum b value of 1.45. These values of b remain approximately independent of the mobility rate when $v < 0.05$ (see Fig. 4). This indicates that the clustering of cooperators is easily attainable below $v \simeq 0.05$. Nevertheless, Fig. 4 shows that at higher speeds ($v > 0.05$) the parameter region where cooperation is available becomes smaller: both solid and dashed lines present a gradual decay (the higher the value of v is, the lower the temptation to defect b that cooperation can resist is). This is because agglomerations of cooperators are hard to maintain in highly mobile systems. Without punishment (dashed line), the system cannot achieve the full cooperation regime when $v \gtrsim 0.2$ (independently of the value of b), whereas in the social environment (solid line) full cooperation is possible up to $v \approx 0.7$.

Figure 4 also shows the transition from full cooperation to full defection for a soft nonsocial environment (dash-dotted line, i.e., $p_a = 0.05$ in addition to $p_s = 0.1$). For this parametrization, the cooperative region covers a higher area of v and b than in the case without punishment (dashed line) but smaller than in the social environment (solid line). The last curve shown in Fig. 4 corresponds to a nonsocial environment where $p_s = p_a = 0.1$ (dotted line). In this case, nonsocial punishment plays an important role against cooperators: the majority of the parameter range explored is dominated by defection, whereas cooperation only manages to survive at very low values of v and b ($v < 0.03$ and $b = 1.05$).

The curves in Fig. 4 clearly show that the system is sensitive to the addition of both social and nonsocial punishment. Thus we now analyze what values of p_s and p_a drive the population to cooperation. Before an in-depth analysis, in the inset of Fig. 4 we consider the case in which social punishment is as frequent as $p_s = 0.3$ and nonsocial punishment is not allowed

(full line). Then cooperation wins for most values of b if the mobility is not high ($v < 0.1$). When increasing the mobility, the cooperative state displays its usual decay to more modest values of b . However, under a critical value of the temptation to defect (in this specific case for $b < 1.45$) cooperation prevails at any speed in the range $v \in [0, 1]$, indicating that the positive effects of such extended social punishment ($p_s = 0.3$) sustain cooperation even when the system is close to a well-mixed population. This behavior remains qualitatively the same if a moderate degree of nonsocial punishment is introduced in the system (dash-dotted line in the inset of Fig. 4, i.e., $p_a = 0.1$ in addition to $p_s = 0.3$). In this latter situation the range of b cooperation can endure is always below the case where only social punishment is present (solid line in Fig. 4, inset). For comparison, we have also included the system without punishment (dashed line in Fig. 4, inset). Finally, the nonsocial environment $p_s = p_a = 0.3$ has also been explored, but no full cooperation phase has been found in the parameter ranges explored. This indicates that increasing the value of p_a leads to a harsher effect against cooperation (compare to the case $p_s = p_a = 0.1$ in Fig. 4 where a small cooperative region is found). Summarizing, we have shown in Fig. 4 that social punishment enhances notably the v - b region available for the cooperative phase. Furthermore, social punishment produce benefits for cooperation even if some degree of nonsocial punishment is considered. However, when nonsocial punishment becomes substantially frequent, it leads to the triumph of freeloaders.

In order to further analyze the role of p_s and p_a , Fig. 5 shows their effect on the transition between cooperative and defective regimes (for $b = 1.5$). The solid line in Fig. 5 shows computational results for a mobility rate $v = 0.01$. Note that cooperation is not possible below a critical social punishment probability ($p_s \lesssim 0.15$). For $p_s > 0.15$, the all-C stable state is reached provided that p_a takes moderate values (full cooperation is reached if $p_s > p_a$). It follows that the more extended the nonsocial punishment is in the system, the more frequent the social punishment has to be in order to sustain cooperation. There is an interesting region ($0.15 < p_s < 0.4$) where social punishment seems to resist better the presence of nonsocial punishment. The dashed line depicts the case $v = 0.2$, where in response to the higher mobility, the system is more sensible to nonsocial punishment in the region $0.15 < p_s < 0.4$. The dotted line corresponds to a substantially higher mobility, $v = 0.6$. Now the system needs more frequent social punishment to make cooperation successful. Indeed, whereas for $v < 0.2$ cooperation is found at $p_s \gtrsim 0.15$ (solid and dashed lines), for $v = 0.6$ no cooperation phase exists below $p_s = 0.25$ (dotted line). Furthermore, there is again a region where social punishment is more efficient against nonsocial punishment (dotted line, $0.25 < p_s < 0.5$). Finally, the dash-dotted line in Fig. 5 displays a highly mobile system ($v = 0.8$), which is closer to a well-mixed population than the case $v = 0.6$ (dotted line). Under very high mobility, the enhanced efficiency of social punishment in the region $0.25 < p_s < 0.5$ disappears, and cooperation is not sustainable below more frequent social punishment ($p_s = 0.35$).

Although the solid line in Fig. 5 depicts the phase transition for the case $v = 0.01$, we have also performed simulations for several values in the range $v \in [0, 0.01]$, and the corresponding phase transitions are independent of v in this range. On the

other hand, the outcome of simulations performed for $v > 0.8$ agree well with the case $v = 0.8$.

Whereas in Fig. 5 we have shown the dependence of the phase transitions on the parameters p_s and p_a for different mobilities and a fixed value of $b = 1.5$, in Fig. 6 we vary the temptation to defect b for a fixed value of v ($v = 0.1$). The dashed line in Fig. 6 ($b = 1.5$, as in Fig. 5) shows similar behavior to the solid line in Fig. 5 ($v = 0.01$), including a region where social punishment is especially effective ($0.15 < p_s < 0.30$). On the other hand, the solid line in Fig. 6 represents a much more social environment since the temptation to defect is now $b = 1.1$ (cooperation is available even in the absence of punishment, $p_s = p_a = 0$). The dotted line in Fig. 6 corresponds to a high temptation to defect ($b = 2$). In this case, the region available for cooperation is limited to high values of social punishment (cooperation is not sustainable for $p_s < 0.45$), similar to what happens under high mobility (dash-dotted line in Fig. 5).

In Figs. 5 and 6, we have shown that for very high values of v or b , defection is the dominant strategy, and cooperation is only possible for extreme social punishment frequencies (in contrast, for low values of v and b cooperation is more sustainable). This extends the results by Meloni *et al.* [21] to systems under social and nonsocial punishment, as observed in many human populations [33]. Moreover, when v and b take moderate values, low doses of social punishment are especially effective in counteracting the effects of nonsocial punishment. This is an important result because it means that players do not need to be rude punishers in order to promote cooperation; thus in the fight against defection, lower mean expenditures on altruistic punishment are necessary.

V. CONCLUSIONS

We have built a model that introduces altruistic punishment options in a population with mobile players of the PD game. Players move continuously in a two-dimensional world, a case of practical relevance with potential applications, such as the design of cooperation-based protocols for communication [25] and the modelization of transitions in human prehistory [26].

In our model, punishment is not a strategy but an action that players may perform against their partners with a certain probability after each round of the game. We have found that punishing after only 10% of the cooperator-defector interactions is enough to lead the system to a world of cooperation, in some environments where otherwise defection would take over the population. Furthermore, this conclusion holds even if some degree of nonsocial punishment (an action that is commonly performed by human players) is present in the system. Our analytical predictions for the mean payoff of the final state agree with simulations. We have also found that, although soft nonsocial punishment can lead to a cooperative state, it yields lower payoffs than social punishment.

We have extensively analyzed the role of the relevant model parameters: the mobility of the players v , the temptation to defect b , the social punishment probability p_s , and the nonsocial one p_a . The phase diagrams have shown that social punishment increases the values of v and b where cooperation is available. We have found that this result resists some degree of nonsocial punishment. Moreover, the full-cooperation region is sensible to the increment of p_a . Finally, we have shown that the benefits of social punishment are limited, and defection asymptotically prevails in harsh environments for cooperation (represented here by high mobilities, high temptation to defect values, and extended practices of nonsocial punishment).

The model in this paper takes into account simple mobility rules and strategies. Additional degrees of complexity could be added in order to closely study human abilities to face defector (or cooperator) neighbors. For example, success-driven migration could be studied (this has been recently proposed in square lattices [16] but not yet for continuous motion).

ACKNOWLEDGMENTS

This work was funded by the Ministry of Science (Grants No. SimulPast-Consolider-CSD-2010-00034 and No. FIS-2009-13050) and the Generalitat de Catalunya (Grant No. 2009-SGR-374). D.C. was supported by the Ministry of Education under the FPU program.

-
- [1] M. A. Nowak, *Evolutionary Dynamics* (Harvard University Press, Cambridge, MA, 2006).
 - [2] R. Axelrod and W. D. Hamilton, *Science* **211**, 1390 (1981).
 - [3] G. Szabó and G. Fáth, *Phys. Rep.* **446**, 97 (2007).
 - [4] C. P. Roca, J. A. Cuesta, and A. Sánchez, *Phys. Rev. E* **80**, 046106 (2009).
 - [5] D. Helbing, A. Szolnoki, M. Perc, and G. Szabó, *New J. Phys.* **12**, 083005 (2010).
 - [6] A. Ule, A. Schram, A. Riedl, and T. N. Cason, *Science* **326**, 1701 (2009).
 - [7] X. Li and L. Cao, *Phys. Rev. E* **80**, 066101 (2009).
 - [8] A. Traulsen, D. Semmann, R. D. Sommerfeld, H. Krambeck, and M. Milinski, *Proc. Natl. Acad. Sci. USA* **107**, 2962 (2010).
 - [9] A. Dreber, D. G. Rand, D. Fudenberg, and M. A. Nowak, *Nature (London)* **452**, 348 (2008).
 - [10] B. Rockenbach and M. Milinski, *Nature (London)* **444**, 718 (2006).
 - [11] M. A. Nowak and R. M. May, *Nature (London)* **359**, 826 (1992).
 - [12] M. Perc and A. Szolnoki, *Biosystems* **99**, 109 (2010).
 - [13] M. A. Nowak, *Science* **314**, 1560 (2006).
 - [14] C. Hauert, *Proc. R. Soc. London, Ser. B* **268**, 761 (2001).
 - [15] J. M. Pacheco, A. Traulsen, and M. A. Nowak, *Phys. Rev. Lett.* **97**, 258103 (2006).
 - [16] D. Helbing and W. Yu, *Proc. Natl. Acad. Sci. USA* **106**, 3680 (2009).
 - [17] S. Van Segbroeck, F. C. Santos, T. Lenaerts, and J. M. Pacheco, *Phys. Rev. Lett.* **102**, 058105 (2009).
 - [18] T. Killingback, M. Doebeli, and N. M. Knowlton, *Proc. R. Soc. London, Ser. B* **266**, 1723 (1999).
 - [19] F. Fu, T. Wu, and L. Wang, *Phys. Rev. E* **79**, 036101 (2009).

- [20] H. X. Yang, Z. X. Wu, and B. H. Wang, *Phys. Rev. E* **81**, 065101 (2010).
- [21] S. Meloni, A. Buscarino, L. Fortuna, M. Frasca, J. Gómez-Gardeñes, V. Latora, and Y. Moreno, *Phys. Rev. E* **79**, 067101 (2009).
- [22] C. A. Aktipis, *J. Theor. Biol.* **231**, 249 (2004).
- [23] M. H. Vainstein, A. T. C. Silva, and J. J. Arenzon, *J. Theor. Biol.* **244**, 722 (2007).
- [24] L. L. Jiang, W. X. Wang, Y. C. Lai, and B. H. Wang, *Phys. Rev. E* **81**, 036108 (2010).
- [25] V. Srivastava, J. Neel, A. B. Mackenzie, R. Menon, L. A. DaSilva, J. E. Hicks, J. H. Reed, and R. P. Gilles, *IEEE Commun. Surv. Tutorials* **7**, 46 (2005).
- [26] J. Fort and T. Pujol, *Rep. Prog. Phys.* **71**, 086001 (2008).
- [27] K. Sigmund, *Trends Ecol. Evol.* **22**, 593 (2007).
- [28] R. Boyd, H. Gintis, S. Bowles, and P. J. Richerson, *Proc. Natl. Acad. Sci. USA* **100**, 3531 (2003).
- [29] A. M. Colman, *Nature (London)* **440**, 744 (2006).
- [30] *Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life*, edited by H. Gintis *et al.* (MIT Press, Cambridge, MA, 2005).
- [31] E. Fehr and S. Gächter, *Nature (London)* **415**, 137 (2002).
- [32] M. Milinski and B. Rockenbach, *Nature (London)* **452**, 297 (2008).
- [33] B. Herrmann, C. Thöni, and S. Gächter, *Science* **319**, 1362 (2008).
- [34] A. Szolnoki, G. Szabó, and M. Perc, *Phys. Rev. E* **83**, 036101 (2011).
- [35] J. Andreoni *et al.*, *Am. Econ. Rev.* **93**, 893 (2003).
- [36] A. Szolnoki and M. Perc, *Europhys. Lett.* **92**, 38003 (2010).
- [37] C. Hilbe and K. Sigmund, *Proc. R. Soc. B* **277**, 2427 (2010).
- [38] S. Gächter and B. Herrmann, *Philos. Trans. R. Soc. B* **364**, 791 (2009).
- [39] R. Boyd, H. Gintis, and S. Bowles, *Science* **328**, 617 (2010).
- [40] D. Helbing, A. Szolnoki, M. Perc, and G. Szabó, *Phys. Rev. E* **81**, 057104 (2010).
- [41] D. Helbing, A. Szolnoki, M. Perc, and G. Szabó, *PLoS Comput. Biol.* **6**, e1000758 (2010).
- [42] *Genetic and Cultural Evolution of Cooperation*, edited by P. Hammerstein (MIT Press, Cambridge, MA, 2003).
- [43] E. Fehr and S. Gächter, *Eur. Econ. Rev.* **42**, 845 (1992).
- [44] J. Wu, B. Zhang, Z. Zhou, Q. He, X. Zheng, R. Cressman, and Y. Tao, *Proc. Natl. Acad. Sci. USA* **106**, 17448 (2009).
- [45] For too low values of ρ , cooperators cannot form clusters, and in the final state only defectors survive. For too high values of ρ the agent's neighborhoods resemble a well-mixed population in which cooperation is again eventually destroyed. See Fig. 2 in Ref. [21].
- [46] The definition of nonsocial punishment in this article coincides with the definition of antisocial punishment in Ref. [33]. In the work of Herrmann *et al.* [33], punishing antisocially embraces penalizations against those who contributed equally or more than you in the game. The public goods game was used in Ref. [33]. Note that when playing the PD game, this definition covers only three possibilities: a cooperator who punishes another cooperator or a defector who punishes any coplayer (either another defector or a cooperator). These are the only three possibilities that embrace the nonsocial punishment defined in our paper.
- [47] See, for example, Fig. 1 in Ref. [33]. Among the 16 comparable participant pools around the world presented there, 13 of them showed notably higher mean expenditures in social punishment than in nonsocial one. Indeed, 11 of those participant pools invested in social punishment more than twice that invested in nonsocial punishment (compare the first 11 participant pools in Ref. [33], Fig. 1, with the last 5 pools).
- [48] J. Dall and M. Christensen, *Phys. Rev. E* **66**, 016121 (2002).