Conformity and Network Effects in the Prisoner's Dilemma

Jorge Peña, Enea Pestelacci, Marco Tomassini and Henri Volken

Abstract— We study the evolution of cooperation using the Prisoner's Dilemma as a metaphor of the tensions between cooperators and non-cooperators, and evolutionary game theory as the mathematical framework for modeling the cultural evolutionary dynamics of imitation in a population of unrelated individuals. We investigate the interplay between network reciprocity (a mechanism that promotes cooperation in the Prisoner's Dilemma by restricting interactions to adjacent sites in spatial structures or neighbors in social networks) and conformity (the tendency of imitating common behaviors). We confirm previous results on the improved levels of cooperation when both network reciprocity and conformity are present in the model and evolution is carried on top of degree-homogeneous graphs, such as rings and grids. However, we also find that scale-free networks are no longer powerful amplifiers of cooperation when fair amounts of conformity are introduced in the imitation rules of the players. Such weakening of the cooperation-promoting abilities of scale-free networks is the result of a less biased flow of information in such topologies, making hubs more susceptible of being influenced by less-connected neighbors.

I. INTRODUCTION

Cooperation is ubiquitous in nature. From replicating molecules to humans, individuals cooperate with others even when such an option represents a cost to themselves. For this reason, understanding the emergence of cooperation is a central problem in many fields of both natural and social sciences, such as biology, sociology, anthropology and economics. Researchers of such diverse disciplines have traditionally adopted the Prisoner's Dilemma (PD) as metaphor for the tension between group welfare and individual self-ishness, and evolutionary game theory [1], [2] as a common formal framework for studying both genetic and cultural evolution.

The PD is a two-person, symmetric game in which players can be either Cooperators (Cs) or Defectors (Ds). Cs are willing to engage in cooperative tasks, while Ds prefer not to, thus exploiting Cs. If two individuals of the same type interact, they both get the reward for mutual cooperation R if they cooperate and the punishment for mutual defection P if they defect. If a D and a C interact, the D receives the temptation to defect T and the C receives the sucker's payoff S. Finally, for the game to be a PD, the pay-offs are ordered such that $T > R > P > S$. It is also assumed that $2R > T + S$, thus ensuring that mutual cooperation is preferred over an equal probability of unilateral cooperation and defection. Other games serving as metaphors for studying cooperation

Jorge Peña and Henri Volken are with the Institute of Applied Mathematics, University of Lausanne, Lausanne. Emails: {jorge.pena,henri.volken}@unil.ch. Enea Pestelacci and Marco Tomassini are with the Information Systems Department, University of Lausanne, Lausanne. Emails: {enea.pestelacci,marco.tomassini}@unil.ch.

are the Snowdrift game (SD) [3], for which $T > R > S > P$, and the Stag Hunt (SH) [4], for which $R > T > P > S$.

In evolutionary game-theoretical models, individuals interact, collect pay-offs and reproduce (genetic evolution) or imitate others (cultural evolution). As a result, the strategy profile of the population evolves in time. Whenever (i) the population is infinite, (ii) pairwise interactions are both anonymous and made at random, and (iii) selection is strictly *pay-off biased*, so that fitter individuals reproduce more (genetic evolution) or individuals have a tendency to imitate successful people (cultural evolution), the evolutionary dynamics can be analytically predicted by a set of equations called the *replicator dynamics* [2]. In the case of the PD, the only stable equilibrium of the replicator dynamics occurs when the population is entirely composed of Ds. In other words, Cs are doomed to extinction whatever their initial share of the population.

Given such unfavorable predictions for the evolution of cooperation, several mechanisms have been invoked in order to explain why altruism can actually evolve [5]. Among these mechanisms one can cite kin selection [6], group selection [7], direct reciprocity [8], [9], indirect reciprocity [10] and network reciprocity [11], [12]. Network reciprocity emerges in models that correct the assumption of infinite and well-mixed populations by embedding individuals in the nodes of networks constraining interactions to some, but not all, of others [11], [13]. These networks are much better representations of how actual interactions occur in real biological and social systems and have been recently studied in great detail (*cf.* [14]). When the population of players possesses such a structure, Cs can survive in clusters of related individuals for a certain range of the pay-off values, as it has been evidenced since the pioneering work by Nowak and May [15]. In particular, scale-free networks [16] have received a lot of attention, since they have been found to promote cooperation to a point that Cs dominate Ds over the entire parameter space of normalized versions of both the PD and the SD [17].

In addition to the simplifying assumption of infinite, well-mixed populations, the replicator dynamics also posits that selection is entirely pay-off biased. Such premise, although obvious in genetic evolution, is less straightforward to posit in cultural evolution, where information is transmitted by means of imitation. Humans not only have a bias for imitating successful people, but also to conform, or to show a disproportionate tendency to follow the majority [18]. Recent empirical research has shown not only that conformity is an important bias in our social learning psychology [19], [20], but also that it can partially account for the results obtained in experimental social dilemmas [21], [22]. The introduction

of conformity in the framework of evolutionary game theory leads to a modified replicator dynamics featuring different equilibrium points from those predicted by the standard set of equations [23], [24], [25]. Particularly, a PD can become a SH for some parameter values, so that Cs can resist invasion from Ds and dominate them if they initially constitute the majority of the population [25].

It is natural to think of network reciprocity and conformity as mechanisms that, if simultaneously present in a population of imitating individuals, would be able to sustain cooperative behavior in cultural evolution. If (i) people imitate according to both pay-off and conformist biases, and (ii) interactions and imitations among individuals are regulated by social networks imposing a topological structure, conformity and network reciprocity could reinforce themselves to promote cooperation beyond the levels predicted when any of the two mechanisms is absent. However, to date and to the best of the authors' knowledge, only few papers have tackled the study of the interplay between conformity and local interactions and their effects on the emergence and stability of cooperation. Bravo [26] explored the outcomes of different imitation strategies in a spatially structured population. In his model, *unsatisfied* agents with a pay-off lesser than the average one in their neighborhoods imitate following *either* 100% pay-off biased or 100% conformist rules. It was found that it is pay-off biased imitation the rule that most favors the spreading of cooperation. In a second study, the first author of this paper showed that, when agents occupy the nodes of a bidimensional lattice and imitate according to *both* pay-off biased and conformist rules, cooperation in the PD is significantly promoted and that the final average fraction of Cs monotonically increases with the amount of conformity [25]. Finally, using a rigorous mathematical formalism, Mengel [27] has recently proved that conformity stabilizes cooperation when agents are arranged in rings and other simple networks.

In this paper, we extend the results presented in [25] regarding the intricate interplay between conformity and network reciprocity in the PD. We show that, depending on the graph topology, conformity can either further promote or noticeably hinder the evolution of cooperation as compared with the baseline models of network reciprocity.

II. MODEL

We focus on the rescaled version of the PD [15], for which $P = S = 0$, $R = 1$ and $1 \leq T \leq 2$ (the upper bound of 2 for T is due to due to the $2R > T + S$ constraint). The resulting game, actually lying in the boundary between the PD and the SD games, has been commonly used in the literature and found to preserve the qualitative properties of the standard PD when played in networks [15], [17], [28].

We consider a population of N individuals or players, where the i-th individual is represented by the vertex v_i of an undirected, simple graph $G(V, E)$. The open neighborhood of i, $\Gamma(i)$, is the set of all players j such that there is an edge $e_{ij} \in E$. The number of neighbors of i is thus the degree k_i

of vertex v_i. The closed neighborhood $\Gamma[i]$ is the set of i's neighbours plus i itself.

At each time step, each individual is either a C or a D. The system evolves by the successive application of interaction and imitation phases. During the interaction phase, players simultaneously engage in a single round of the PD with their neighbors. As a result, individual i collects an accumulated payoff $\Pi_i = \sum_{l \in \Gamma(i)} \pi_{il}$, where π_{il} is the pay-off that player i receives when interacting with player l. During the imitation phase, each player i randomly chooses one of its neighbors as its *cultural parent* and copies its strategy with a probability ϕ. Imitation is conformist biased with probability α and pay-off biased with probability $1-\alpha$, where α is a parameter measuring the amount of conformity in the individuals' psychology.

Let us denote i's cultural parent by j. If imitation is pay-off biased, $\phi = \phi_{pay-off}$ is monotonically increasing with the difference in j's and i's pay-offs and given by

$$\phi_{pay-off}(\Pi_j - \Pi_i) = \begin{cases} \frac{\Pi_j - \Pi_i}{T k_>} & \text{if } \Pi_j - \Pi_i > 0 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where $k_> = \max\{k_i, k_j\}$. Eq. 1 is a local, finite population analogue of the replicator dynamics, commonly used in the literature [17], [3]. If imitation is conformist biased, $\phi = \phi_{conf}$ is monotonically increasing with the difference in the proportions of i's and j's strategies in the local vicinity of i:

$$\phi_{conf}(p_{ij} - p_{ii}) = \begin{cases} p_{ij} - p_{ii} & \text{if } p_{ij} > p_{ii} \\ 0 & \text{otherwise} \end{cases}$$

where p_{il} is the proportion of individuals in $\Gamma[i]$ having the same strategy as l. Notice that when $\alpha = 0$ our local dynamics reduces to the strictly pay-off biased imitation rule used in previous studies [17], [3].

A. Evolutionary dynamics in infinite, well-mixed populations

In the standard case of a large, well-mixed population, the time evolution of the fraction of cooperators x for the simplified PD can be shown to be given by:

$$\dot{x} = x(1-x)\left\{\frac{1-\alpha}{T}[\pi_C - \pi_D] + \alpha(2x-1)\right\}, \quad (2)$$

where $\pi_C = x$ and $\pi_D = xT$ are the average pay-offs to Cs and Ds. Eq. 2 (or a similar formula) has been derived in related work on cultural transmission processes including both pay-off biased and conformist imitation [21], [23], [25], [24], [29]. The dynamics has the two trivial fixed points $x_0^* = 0$ and $x_0^* = 1$, as well as, if $\alpha > (T-1)/(2T-1)$, one internal non-trivial equilibrium

$$x^* = \frac{\alpha T}{(1-T) + \alpha(3T-1)}.$$

Variations in the amount of conformity can change the severity of the rescaled PD, leading to two dynamical regions:

1) Dominant defection ($\alpha < (T-1)/(2T-1)$): x_0^* is the only stable equilibrium. In this case, cooperators are doomed to extinction regardless of their initial frequency in the population.

2) Bi-stability ($\alpha > (T-1)/(2T-1)$): both x_0^* and x_1^* are stable whereas the internal fixed point x^* is unstable. In this case, the evolutionary dynamics depends on the initial frequency of Cs, x_0. For $x_0 > x^*$ cooperation prevails, whereas it vanishes for $x_0 < x^*$. Since $x^* > 1/2 \; \forall \alpha$, Cs should be initially the majority in order to have any chance to prevail. This dynamical region corresponds to the SH under the standard replicator dynamics. Thus, the effects of adding conformity to the PD can be conceptualized as a transformation of the underlying game to a SH [25].

In sum, conformity can promote cooperation in the PD to a certain degree in the mean-field limit. If in the majority (and if conformity is strong enough) Cs now have a chance of surviving invasion from Ds, and eventually take over the whole population [25]. However, in order for a minority of Cs to persist, other mechanisms should be also at work. One such mechanism is network reciprocity.

B. Evolutionary dynamics in finite, structured populations

The infinite, well-mixed case is important since it is easy to analyze and it can be described by means of simple differential equations. However, real populations are both finite and structured, so that interactions are constrained to local neighbors. Games on graphs have been traditionally carried on the top of degree-homogeneous (or regular) graphs such as one- and two-dimensional lattices. These graphs are characterized by a single-peaked degree distribution. In other words, the number of neighbors is the same for all the individuals. More recently, several researchers (*cf.* [30], [17], [28], [31]) have also studied evolutionary game dynamics on top of degree-heterogeneous graphs characterized by broad-scale degree distributions. Particularly, Erdős-Renyi random graphs [32] (with Poissonian degree distributions) and Barabasi-Albert networks [16] (with scale-free degree distributions) have received a lot of attention due to the beneficial effect of such degree heterogeneity in the evolution of cooperation [17], [33], [28].

In order to study the interplay between conformity and network reciprocity on both degree-homogeneous and degree-heterogeneous population structures, we consider four different topologies: (i) Rings (regular 1D-lattices with cyclic boundary conditions), (ii) Grids (regular 2D-lattices with cyclic boundary conditions), (iii) Erdős-Renyi random graphs [32] and (iv) Barabasi-Albert scale-free networks [16]. For cases (i), (iii) and (iv) we generated networks with sizes $N = 10^3$, $N = 3 \times 10^3$ and $N = 10^4$, and average degrees $\bar{k} = 4$ and $\bar{k} = 8$. The random graphs were generated using the standard method of independently creating $(\bar{k} N)/2$ edges between pairs of nodes chosen uniformly at random [32]. With the values of $\bar{k}$ used in this study, it is mathematically certain that there is a giant component in the graph [32]. Besides, we ensure that the whole graph is connected. The scale-free networks were generated using the standard Barabasi-Albert growing method starting with a clique of $\bar{k}/2$ nodes and then creating $\bar{k}/2$ edges for each

new node that joins the graph [16]. For grids, we used sizes of $N = 961$ (31×31), $N = 2916$ (54×54) and $N = 10^4$ (100×100) and neighborhoods of the Von Neumann ($\bar{k} = 4$) and Moore ($\bar{k} = 8$) types.

Populations where randomly initialized with 50% Cs and 50% Ds. The probability α of using the conformist transmission rule is set to different values between 0 and 0.5 in steps of 0.1. The temptation to defect T was varied in steps of 0.05. For each experiment, we carried out 50 runs of 10^4 steps each, using a fresh graph realization in each run. The final proportions of Cs and Ds were obtained by averaging over 10^3 steps after a relaxation time of 10^4 steps.

III. RESULTS

A. Degree-Homogeneous Graphs

The final proportion of Cs obtained on grids and rings of 10^4 nodes for different amounts of conformity are shown in Fig. 1 and Fig. 2. These plots confirm results previously obtained both for the $\alpha = 0$ case (*cf.* [15], [34], [17]) and the $\alpha > 0$ case [25]. Even without conformity, Cs are able to survive for low values of T by forming clusters within which they interact more often with their own strategy than what is expected in mixing populations. Cs can thus greatly benefit from mutual cooperation and counterbalance the exploitation of Ds at the borders [34].

As it was already observed in [25] for smaller grids, conformity enhances cooperation in these regular graphs, moving the threshold of T for which Cs become extinct farther to the right. Furthermore, the different curves for each regular graph are ordered such that the higher the amount of conformity, the higher the steady-state fraction of Cs in the population and the larger the value of T for which defection dominates. Conformity has a positive effect on regular structures because it helps to promote the creation of clusters of individuals of the same type. Such clustering is both favorable for Cs and detrimental for Ds, since cooperator-cooperator interactions yield higher pay-offs than defector-defector interactions.

B. Degree-Heterogeneous Graphs

Fig. 3 and 4 show the results obtained for the considered degree-heterogeneous graphs. These graphs have been found to enhance cooperation thanks to the colonization of the more connected nodes by Cs [33]. In particular, scale-free networks, which are highly degree-heterogeneous, have been shown to greatly promote the emergence and sustainability of cooperation [17], [28]. We have confirmed that this is the case when imitation is strictly pay-off biased, as it can be seen from the thick curves of Fig. 3 and 4. For random graphs, Cs are still able to survive for values of T as large as 1.8. The effect of degree-inhomogeneity in the sustainability of cooperation is much more evident in scale-free networks. Indeed, for this type of graphs, the proportion of Cs is always above 0.7 for $\bar{k} = 4$ and above 0.35 for $\bar{k} = 8$, for all values of T. When comparing these results to those obtained in homogeneous graphs, it is clear that degree heterogeneity
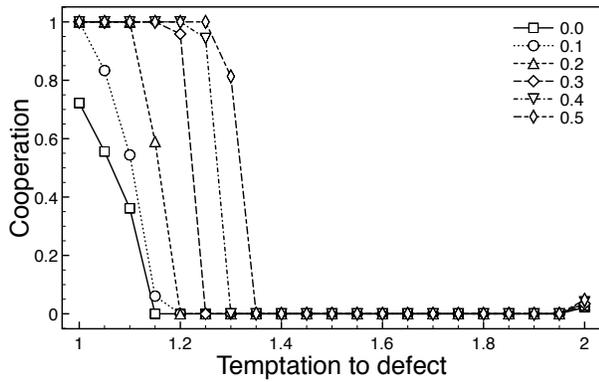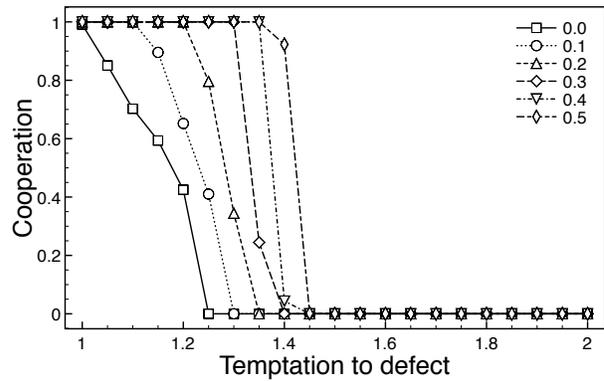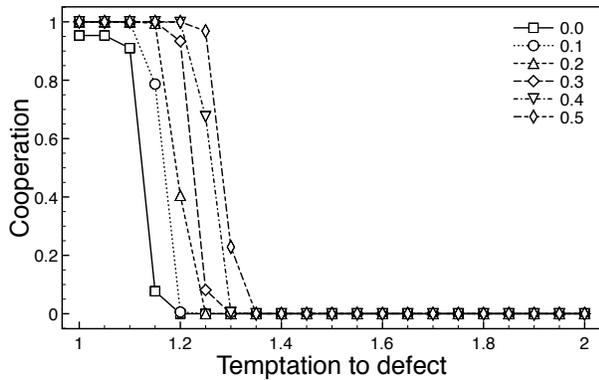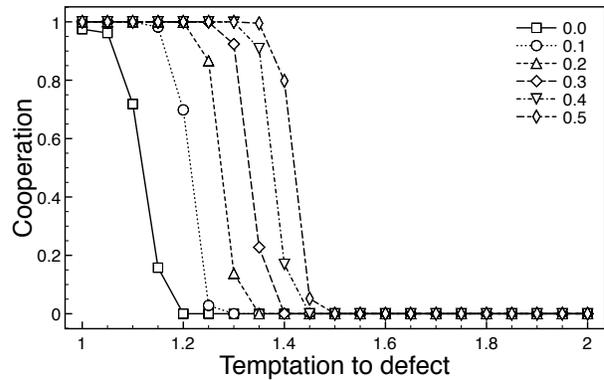
Fig. 1. Final proportion of Cs on a grid with $N = 10^4$. (a) $\bar{k} = 4$ (Von Neumann). (b) $\bar{k} = 8$ (Moore). In each figure, six different curves are shown, one for each value of $\alpha \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5\}$.
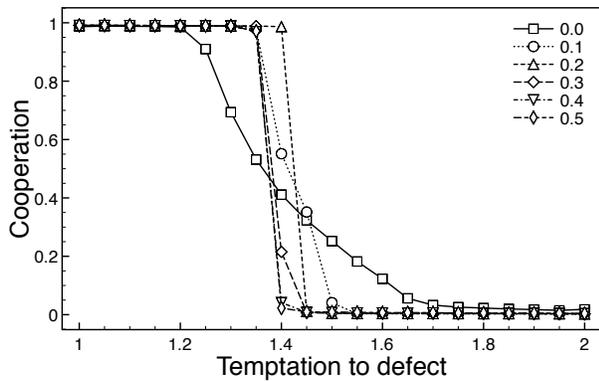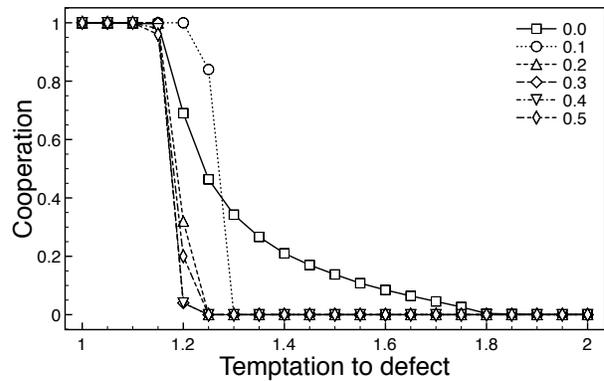

Fig. 2. Final proportion of Cs on a ring with $N = 10^4$. (a) $\bar{k} = 4$. (b) $\bar{k} = 8$. In each figure six different curves are shown, one for each value of $\alpha \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5\}$.

Fig. 3. Final proportion of Cs on random graphs with $N = 10^4$. (a) $\bar{k} = 4$. (b) $\bar{k} = 8$. In each figure six different curves are shown, one for each value of $\alpha \in \{0.0, 0.1, 0.2, 0.3, 0.4, 0.5\}$.

enhances network reciprocity and, for the case of scale-free networks, greatly supports cooperative behavior.

When the population is structured according to degree-heterogeneous graphs, the addition of conformity has some-what counterintuitive consequences. Only for a scale-free topology with $\bar{k} = 4$ and $\alpha \leq 0.2$ does conformity improve the final proportion of Cs beyond what is obtained with $\alpha = 0$ for all values of T. In the rest of the cases, conformity

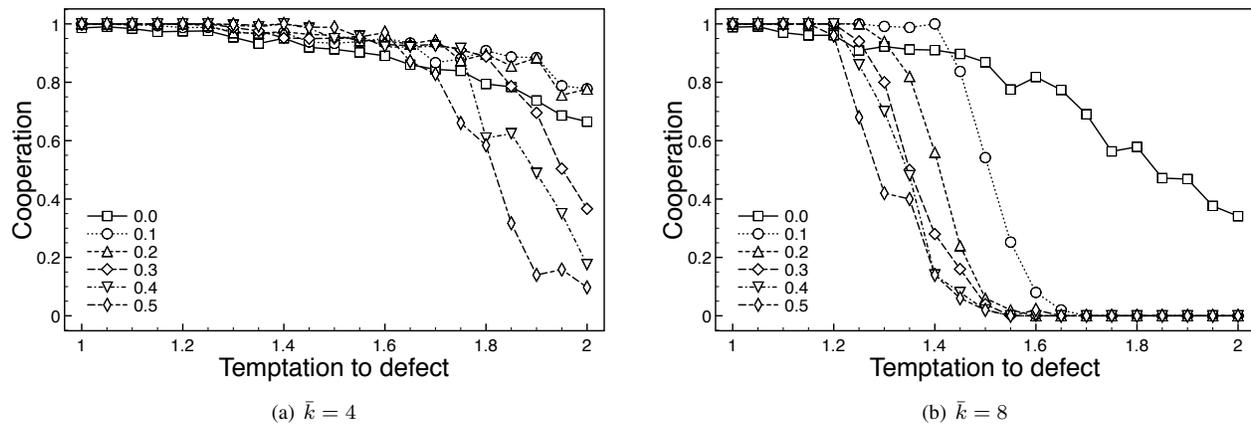(a) $\bar{k} = 4$ (b) $\bar{k} = 8$

Fig. 4. Final proportion of Cs on Barabasi-Albert scale-free networks with $N = 10^4$. (a) $\bar{k} = 4$. (b) $\bar{k} = 8$. In each figure six different curves are shown, one for each value of $\alpha \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5\}$.

does not hamper cooperation for small to medium values of T but is detrimental for large values of T. Furthermore, the threshold above which the final fraction of Cs is higher than in the case without conformity seems to be a monotonically decreasing function of both α and $\bar{k}$, so that the higher the amount of conformity and the degree of the graph, the smaller the level of temptation for which conformity is no more favorable for Cs as compared to the standard case. Particularly, for scale-free networks with $\bar{k} = 8$ and $\alpha \geq 0.2$, conformity weakens the advantage of scale-free networks in promoting cooperation. Indeed, for such parameter values, the final fraction of Cs in these highly heterogeneous graphs is in general not much larger than the corresponding fraction obtained in grids or rings. For $T \leq 1.4$, cooperation is even better sustained by regular graphs than by scale-free networks (or random graphs) when $\alpha \geq 0.3$.

Apart from the effects on the average level of cooperation, the addition of conformity can also lead to a different dynamical organization of cooperation in heterogeneous graphs. When individuals imitate exclusively according to a pay-off bias, one observes that Cs and Ds coexist in quasi-equilibrium, with some nodes fixed in cooperative or defective behavior and others where there is no fixation and cycles of invasion follow indefinitely [35]. In this case, the gradual drop in cooperation seen in the thick curves of Fig. 3 and 4 is mostly due to fluctuating individuals spending less and less time engaging in cooperative behavior. Such dynamical picture changes when individuals not only imitate according to a pay-off bias, but also follow a conformist rule. In this case, for $\bar{k} = 8$, the population almost always reaches one of the two absorbing states, so that in the limit only one strategy gets fixed: Cs for low values of T, Cs or Ds (with a certain probability) for intermediate values of T, and Ds for large values of T. In general, and contrary to what happens without conformity, intermediate levels of cooperation for the cases with $\alpha > 0$ are not the result of the coexistence or fluctuation of different strategies but of the fact that, for a narrow window of the temptation T,

the system converges some times to the pure cooperator state and some times to the pure defector state. Additionally, evolutionary dynamics develop much faster in the presence of conformity. Fig. 5 illustrate these observations for the case of scale-free networks with $N = 10^4$, $\bar{k} = 8$ and $T = 1.35$. Without conformity (Fig. 5(a)) the fraction of Cs for each run slowly increases until, eventually, it stabilizes around a value of 0.9 (not shown in the figure). Conversely, with $\alpha = 0.3$ (Fig. 5(b)), very early in the evolutionary process the population goes either to full cooperation or to full defection.

C. Scaling

Any numerical study must make use of a finite (and not excesively large) number of individuals. This causes finite-size effects that should be gauged in order to ensure the statistical significance of the obtained results. We have performed numerical simulations for three different population sizes: $N = 10^3$, $N = 3 \times 10^3$, and $N = 10^4$. The figures shown in the previous sections referred to the $N = 10^4$ case. Here we show how results change when smaller networks are used. For reasons of space, we only show the cases of rings and scale-free graphs with $\bar{k} = 4$ and $\alpha = 0.0$ and 0.5. In the case of rings (Fig. 6(a)) the curves are almost superimposed and $N = 10^3$ is already sufficient to get stable results. In the case of scale-free graphs (Fig. 6(b)) the behavior is slightly noisier but the general trends are maintained for all sizes with and without conformity. This is expected since regular graphs do not change their local topology when the graph size is changed, while in scale-free graphs inter-hub connections are affected to some extent.

IV. DISCUSSION

We have confirmed that, as it has been suggested by previous studies [25], [27], conformity further promotes cooperation when the network regulating interactions and imitations is a regular graph. In this case, conformity and network reciprocity reinforce each other, thus favoring the formation of clusters of individuals of the same type that allow Cs to maximize their fitness and resist exploitation by
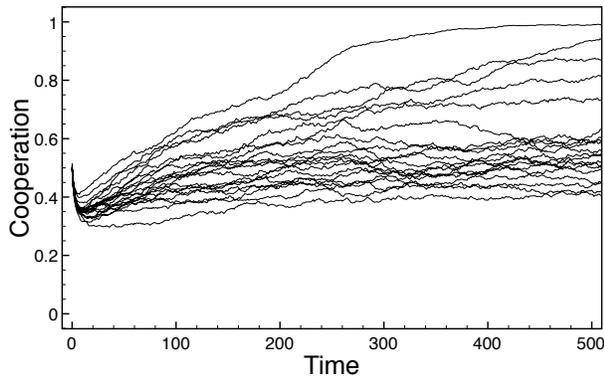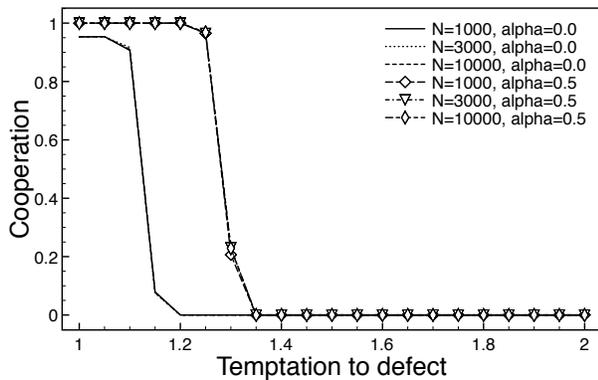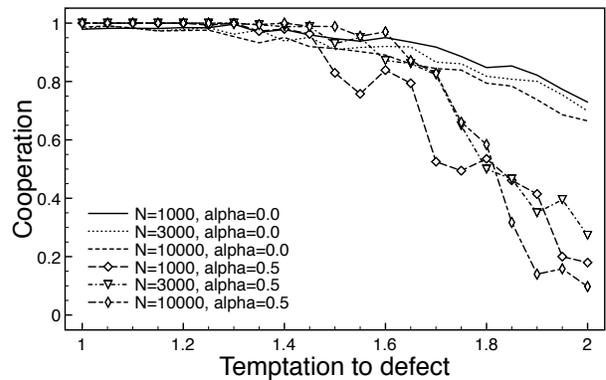
(a) $\alpha = 0.0$
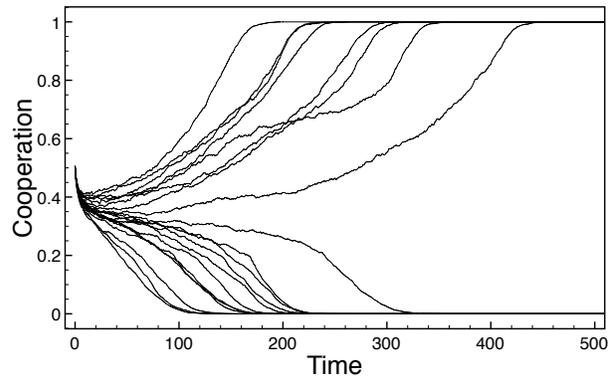
(b) $\alpha = 0.3$

Fig. 5. Evolution of the proportion of Cs on Barabasi-Albert scale-free graphs with $N = 10^4$ during the first 500 time steps. In both figures, $T = 1.35$. (a) $\alpha = 0.0$. (b) $\alpha = 0.3$. In both figures 20 distinct curves are shown.

(a) Ring, $\bar{k} = 4$

(b) Scale-free, $\bar{k} = 4$

Fig. 6. Final proportions of Cs for $\alpha = 0.0$ and $\alpha = 0.5$ and three different population sizes ($N = 10^3$, $N = 3 \times 10^3$ and $N = 10^4$). (a) Rings. (b) Barabasi-Albert scale-free graphs.

surrounding Ds. Somewhat intriguingly, however, conformity can also hamper cooperation when evolution is carried on top of random and scale-free networks. Indeed, for $\bar{k} = 8$ and fair amounts of conformity, these two types of heterogeneous graphs no longer enhance cooperation as compared to regular structures. In particular, the highly degree-heterogeneous scale-free networks lose their abilities to sustain considerable amounts of cooperation in the PD for the whole spectrum of T.

To provide an explanation of this phenomenon, let us first review the mechanism responsible for promoting and sustaining cooperation in degree-heterogeneous graphs in the absence of conformity. For this, we make use of the notion of temperature of players as defined in [11]. Hot players are those who play more since they have a large number of neighbors, whereas cold players are those who have few neighbors and, consequently, play less games. By playing more often, and provided that pay-offs are positively biased (e.g. $S > 0$), hot players get higher accumulated payoffs than cold players. Under pay-off biased transmission, this implies that hot players are also more successful in being imitated

and in disseminating their strategies. The flow of information is thus biased in degree-heterogeneous graphs, with strategies generally spreading from hot to cold players [36].

Let us now consider the PD game. Both Cs and Ds do better when they are surrounded by Cs. Hot players easily spread their strategies. By spreading defective behavior, hot Ds become less and less successful, since the number of their cooperator neighbors decreases. Hot Cs, on the contrary, see their pay-off increased by spreading their own cooperative strategy. The more hot cooperators are imitated the more they earn and the more difficult for a surrounding D to invade. A typical example of such "hub dynamics" is illustrated in Fig. 7(a) for the most connected hub of a Barabasi-Albert graph. The hub is D at the beginning of the simulation, while the rest of the population is initialized to around 50% Cs and 50% Ds. Many cooperator neighbors imitate the defective hub (or other surrounding Ds) during the first steps of simulation, so that the proportion of cooperator neighbors is reduced to approximately 30%. As a consequence, the total pay-off of the hub is importantly reduced, and the hub becomes vulnerable to invasion from a neighboring C. When the hub becomes
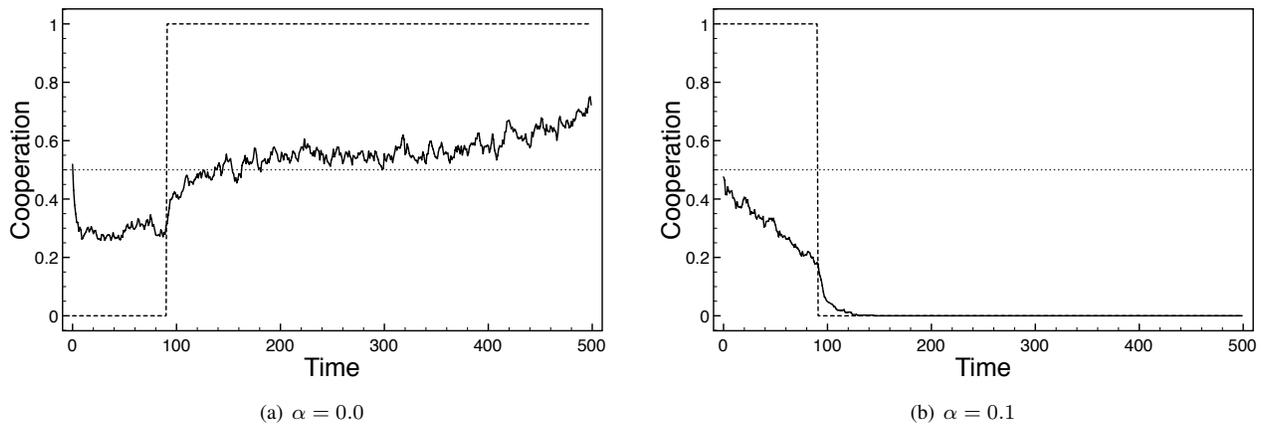
(a) $\alpha = 0.0$ (b) $\alpha = 0.1$

Fig. 7. Evolution of cooperation around the most connected hub of a Barabasi-Albert scale-free network with $N = 10^4$, $\bar{k} = 8$ and $T = 1.35$. The fraction of cooperator neighbors is shown in *solid* lines and the strategy of the hub is in *dashed* lines, so that a value of 0 represents a defective strategy and a value of 1 a cooperative strategy. As a reference, the level of 50% cooperation among the hub's neighbors is depicted in *dotted* lines. The most connected hub is initially set to D (a) or to C (b). The rest of the population is initialized to around 50% Cs and 50% Ds. (a) $\alpha = 0.0$. (b) $\alpha = 0.1$.

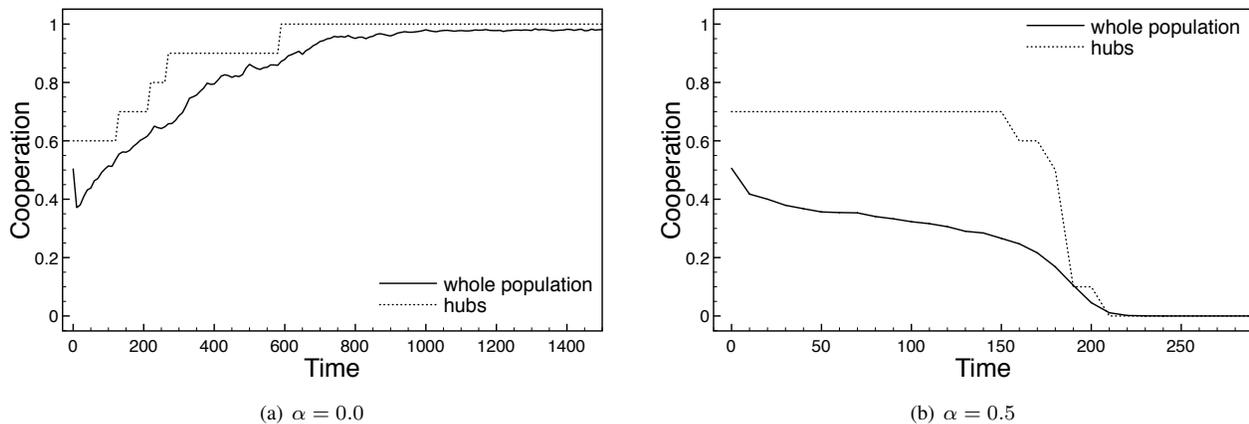(a) $\alpha = 0.0$ (b) $\alpha = 0.5$

Fig. 8. Evolution of cooperation on a Barabasi-Albert scale-free network with $N = 10^4$, $\bar{k} = 8$ and $T = 1.35$. The fraction of Cs in the whole population is shown in *solid* lines while the fraction of Cs among the 10 most connected hubs is shown in *dotted* lines. (a) $\alpha = 0.0$. (b) $\alpha = 0.5$.

a C, more and more of its defective neighbors also switch their strategies. Consequently, the proportion of cooperator neighbors (and the total pay-off to the hub) increases and is maintained at a high level afterwards. The presence of such positive feedback mechanism, and the fact that it only works for Cs, greatly enhances cooperation in heterogeneous graphs and, particularly, in scale-free networks [28]. Under strict pay-off biased imitation, cooperator hubs 'guide' the evolution of cooperation in the rest of the population, as it is shown in Fig. 8(a).

The introduction of conformity decreases the bias in the flow of information in heterogeneous graphs, making hubs vulnerable to invasion from their cold neighbors. While hubs are unlikely to imitate their low connected neighbors when using a pay-off biased rule, nothing prevents them from imitating a cold surrounding player if it holds the strategy of the local majority. Since the fraction of Cs generally decreases at the outset of the simulation (see the first time steps of the curves shown in Fig. 5), conformity further favors Ds, which become predominant in the population. An example

of this dynamics is shown in Fig. 7(b). Initially, the hub is a C. Many of the hub's neighbors turn to defection during the first time steps, making cooperation the less common strategy in the hub's neighborhood. Around the 100th time step, the hub imitates by conformity one of its defector neighbors, leading to a quicker decrease in the proportion of cooperation in its neighborhood. Shortly after, Cs completely vanish around the most connected hub. During those first time steps, hubs imitating according to a conformist bias will have high chances of becoming defectors. Such initial asymmetry in the strategies of the hubs can account for the negative effects of conformity in the evolution of cooperation on scale-free networks. Indeed, as it has been discussed elsewhere [37], Cs go easily extinct when the initial distribution of strategies is biased so that Ds have higher chances of occupying highly connected nodes. Conformity partly reverses the flow of information on degree-heterogeneous networks so that hubs no longer take the lead of the dynamics and instead conform to the general trend of the whole population (see Fig. 8(b)).

V. Conclusions

The presented results and analyses confirm that conformity and network reciprocity, when acting independently on the framework of evolutionary game theory, are able to promote cooperation in the one-shot Prisoner's Dilemma beyond the limits of the standard replicator dynamics. However, when they are simultaneously present, they do not always reinforce their cooperation-promoting abilities, as it could be naively expected. Indeed, whether conformity strengthens or weakens the evolution of cooperation depends on the intrinsic characteristics of the underlying graph. Conformity favors cooperation when evolution is carried out on degree-homogeneous graphs for the whole range of the temptation to defect T. Conversely, it can hinder cooperation in degree-heterogeneous graphs for medium to large values of T. In particular, for fair amounts of conformity, scale-free networks with relatively small mean degrees do not show the great improvement over regular structures that has been previously reported in the literature.

Acknowledgment

References

[1] J. Maynard Smith, *Evolution and the Theory of Games*. Cambridge Univ. Press, Cambridge, 1982.

[2] H. Gintis, *Game theory evolving*. Princeton: Princeton University Press, 2000.

[3] C. Hauert and M. Doebeli, "Spatial structure often inhibits the evolution of cooperation in the snowdrift game," *Nature*, vol. 428, no. 6983, pp. 643–646, Apr. 2004.

[4] B. Skyrms, *The Stag-Hunt Game and the Evolution of Social Structure*. Cambridge: Cambridge University Press, 2003.

[5] M. A. Nowak, "Five Rules for the Evolution of Cooperation," *Science*, vol. 314, no. 5805, pp. 1560–1563, 2006.

[6] W. D. Hamilton, "The genetical evolution of social behaviour," *J. Theor. Biol.*, vol. 7, pp. 1–16, 1964.

[7] D. S. Wilson, *Proc Natl Acad Sci USA*, vol. 72, no. 1, pp. 143–146, 1975.

[8] R. Trivers, "The evolution of reciprocal altruism," *Quart. Rev. Biol.*, vol. 46, p. 3557, 1971.

[9] R. Axelrod and W. Hamilton, "The evolution of cooperation," *Science*, vol. 211, pp. 1390–1396, 1981.

[10] M. A. Nowak and K. Sigmund, "Evolution of indirect reciprocity by image scoring," *Nature*, vol. 393, no. 6685, pp. 573–577, June 1998.

[11] E. Lieberman, C. Hauert, and M. A. Nowak, "Evolutionary dynamics on graphs," *Nature*, vol. 433, no. 7023, pp. 312–316, Jan. 2005.

[12] H. Ohtsuki, C. Hauert, E. Lieberman, and M. A. Nowak, "A simple rule for the evolution of cooperation on graphs and social networks," *Nature*, vol. 441, no. 7092, pp. 502–505, May 2006.

[13] G. Szabo and G. Fath, "Evolutionary games on graphs," *Physics Reports*, vol. 446, no. 4-6, pp. 97–216, July 2007.

[14] M. E. J. Newman, "The structure and function of complex networks," *SIAM Review*, vol. 45, pp. 167–256, 2003.

[15] M. A. Nowak and R. M. May, "Evolutionary games and spatial chaos," *Nature*, vol. 359, no. 6398, pp. 826–829, 1992.

[16] R. Albert and A.-L. Barabasi, "Statistical mechanics of complex networks," *Reviews of Modern Physics*, vol. 74, pp. 47–97, 2002.

[17] F. C. Santos and J. M. Pacheco, "Scale-free networks provide a unifying framework for the emergence of cooperation," *Phys. Rev. Lett.*, vol. 95, no. 9, pp. 098 104–4, Aug. 2005.

[18] R. Boyd and P. J. Richerson, *Culture and the evolutionary process*. Chicago: University of Chicago Press, 1985.

[19] J. C. Coultas, "When in Rome... an evolutionary perspective on conformity," *Group Processes Intergroup Relations*, vol. 7, no. 4, pp. 317–331, Oct. 2004.

[20] C. Efferson, R. Lalive, P. J. Richerson, R. McElreath, and M. Lubell, "Conformists and mavericks: the empirics of frequency-dependent cultural transmission," *Evolution and Human Behavior*, vol. 29, no. 1, pp. 56–64, Jan. 2008.

[21] J. P. Carpenter, "When in Rome: conformity and the provision of public goods," *Journal of Socio-Economics*, vol. 33, no. 4, pp. 395–408, Sept. 2004.

[22] N. Bardsley and R. Sausgruber, "Conformity and reciprocity in public good provision," *Journal of Economic Psychology*, vol. 26, no. 5, pp. 664–681, Oct. 2005.

[23] J. Henrich, "Cultural transmission and the diffusion of innovations: Adoption dynamics indicate that biased cultural transmission is the predominate force in behavioral change and much of sociocultural evolution," *American Anthropologist*, vol. 103, pp. 992–1013, 2001.

[24] B. Skyrms, "Dynamics of conformist bias," *The Monist*, vol. 88, no. 2, pp. 259–269, April 2005.

[25] J. Peña, "Conformist transmission and the evolution of cooperation," in *Artificial Life XI: Proceedings of the Eleventh International Conference on the Simulation and Synthesis of Living Systems*, S. Bullock, J. Noble, R. Watson, and M. A. Bedau, Eds. MIT Press, Cambridge, MA, 2008, pp. 458–465.

[26] G. Bravo, "Imitation and cooperation in different helping games," *Journal of Artificial Societies and Social Simulation*, vol. 11, no. 1, p. 8, 2008. [Online]. Available: http://jasss.soc.surrey.ac.uk/11/1/8.html

[27] F. Mengel, "Conformism and cooperation in a local interaction model," *Journal of Evolutionary Economics*, pp. –, -. [Online]. Available: http://dx.doi.org/10.1007/s00191-008-0131-7

[28] F. Santos, J. Rodrigues, and J. Pacheco, "Graph topology plays a determinant role in the evolution of cooperation," *Proceedings of the Royal Society B: Biological Sciences*, vol. 273, no. 1582, pp. 51–55, Jan. 2006.

[29] J. Henrich and R. Boyd, "Why people punish defectors: conformist transmission stabilizes costly enforcement of norms in cooperative dilemmas," *Journal of Theoretical Biology*, vol. 208, pp. 79–89, 2001.

[30] G. Abramson and M. Kuperman, "Social games in a social network," *Phys. Rev. E*, vol. 63, no. 3, pp. 030 901–, Feb. 2001.

[31] L. Luthi, E. Pestelacci, and M. Tomassini, "Cooperation and community structure in social networks," *Physica A*, vol. 387, pp. 955–966, February 2008.

[32] B. Bollobás, *Random Graphs*. New York: Academic Press, 2001, 2nd ed.

[33] F. C. Santos, J. F. Rodrigues, and J. M. Pacheco, "Epidemic spreading and cooperation dynamics on homogeneous small-world networks," *Phys. Rev. E*, vol. 72, no. 5, pp. 056 128–5, Nov. 2005.

[34] M. Doebeli and C. Hauert, "Models of cooperation based on the prisoner's dilemma and the snowdrift game," *Ecology Letters*, vol. 8, no. 7, pp. 748–766, 2005.

[35] J. Gómez-Gardeñes, M. Campillo, L. M. Floría, and Y. Moreno, "Dynamical organization of cooperation in complex topologies," *Phys. Rev. Lett.*, vol. 98, no. 10, pp. 108 103–4, Mar. 2007.

[36] N. Masuda, "Participation costs dismiss the advantage of heterogeneous networks in evolution of cooperation," *Proceedings of the Royal Society B: Biological Sciences*, vol. 274, no. 1620, pp. 1815–1821, Aug. 2007.

[37] X. Chen, F. Fu, and L. Wang, "Influence of initial distributions on robust cooperation in evolutionary Prisoner's Dilemma," *ArXiv Physics e-prints*, Jan. 2007.