**THE EUROPEAN
PHYSICAL JOURNAL B**

Regular Article

Statistical physics of the spatial Prisoner's Dilemma with memory-aware agents

Marco Alberto Javarone[1,2,a]

[1] Department of Mathematics and Computer Science, University of Cagliari, 09123 Cagliari, Italy
[2] DUMAS – Department of Humanities and Social Sciences, University of Sassari, 07100 Sassari, Italy

Abstract. We introduce an analytical model to study the evolution towards equilibrium in spatial games, with 'memory-aware' agents, i.e., agents that accumulate their payoff over time. In particular, we focus our attention on the spatial Prisoner's Dilemma, as it constitutes an emblematic example of a game whose Nash equilibrium is defection. Previous investigations showed that, under opportune conditions, it is possible to reach, in the evolutionary Prisoner's Dilemma, an equilibrium of cooperation. Notably, it seems that mechanisms like motion may lead a population to become cooperative. In the proposed model, we map agents to particles of a gas so that, on varying the system temperature, they randomly move. In doing so, we are able to identify a relation between the temperature and the final equilibrium of the population, explaining how it is possible to break the classical Nash equilibrium in the spatial Prisoner's Dilemma when considering agents able to increase their payoff over time. Moreover, we introduce a formalism to study order-disorder phase transitions in these dynamics. As result, we highlight that the proposed model allows to explain analytically how a population, whose interactions are based on the Prisoner's Dilemma, can reach an equilibrium far from the expected one; opening also the way to define a direct link between evolutionary game theory and statistical physics.

1 Introduction

Evolutionary games [1–3] represent the attempt to study the evolution of populations [4–6] by the framework of game theory [7]. Notably, these games allow to analyze simplified scenarios in different domains, spanning from socio-economic dynamics to biological systems [1,8–17]. In general, evolutionary games consider a population of agents whose interactions are based on games like the Prisoner's Dilemma (hereinafter PD) or the Hawk-Dove game [4], where there are two possible strategies: cooperation and defection. As in classical game theory, the concept of equilibrium represents a core aspect [18]. Therefore, we aim to evaluate if a population reaches an equilibrium equal or different from the expected one, i.e., the Nash equilibrium of the considered game. At each interaction, agents gain a payoff according to the adopted strategy and to a payoff matrix. The payoff represents a form of reward in the considered domain (e.g., money in an economic system or food in an ecosystem). Remarkably, as agents are allowed to change their strategy over time, we can map them to spins with states $\sigma = \pm 1$, representing cooperation and defection, respectively. In doing so, we can analyze order-disorder transitions in the spatial PD. Previous studies [19–25] have shown that, under particular condi-

tions, it is possible that a population playing a game like the PD, i.e., a game characterized by defection as Nash equilibrium, can be able to reach a final state of full cooperation. For instance, it seems that both motion [19–22] and competitiveness [24] can lead an agent population to cooperate [26] and, more in general, spatial structure plays a key role in the evolution of cooperation [27,28]. Usually, adding properties to agents, as motion, conformity and competitiveness, entails to increase the complexity of the resulting model. Thus, most investigations on evolutionary games are based on computational approaches. Therefore, in this work we try to provide an analytical description of the spatial PD, in order to explain how a population can become cooperative and to strengthen the link between evolutionary game theory [29] and statistical physics [30]. It is worth highlighting that we consider 'memory-aware' agents, i.e., agents that accumulate their payoff over time. Remarkably, this last condition represents the major difference with most of the evolutionary game models studied by computational approaches (see for instance [31,32]). On the other hand, considering 'memory-aware' agents makes the problem more tractable from an analytical perspective. The remainder of the paper is organized as follows: Section 2 introduces the proposed model and its analytical formulation. Section 3 shows analytical results. Eventually, Section 4 ends the paper.

[a] e-mail: marcojavarone@gmail.com

2 Model

In the proposed model, we are interested in studying the spatial prisoner's dilemma by an analytical approach. Let us start by introducing the general form of a payoff matrix

$$
\begin{array}{cc}
 & \begin{array}{cc} C & D \end{array} \\
\begin{array}{c} C \\ D \end{array} & \begin{pmatrix} R & S \\ T & P \end{pmatrix}
\end{array} \tag{1}
$$

where the set of strategies is $\Sigma = \{C, D\}$: C stands for 'Cooperator' and D for 'Defector'. In the matrix (1), R is the gain obtained by two interacting cooperators, T represents the *Temptation*, i.e., the payoff that an agent gains if it defects while its opponent cooperates, S the *Sucker's payoff*, i.e., the gain achieved by a cooperator while the opponent defects, eventually P the payoff of two interacting defectors. In the case of the PD, matrix elements of (1) are: $R = 1$, $0 \leq S \leq -1$, $1 \leq T \leq 2$ and $P = 0$. As stated before, during the evolution of the system agents can change their strategy from C to D, and vice versa, following an updating rule, as for instance the one named 'imitation of the best' (see [4,19]), where agents imitate the strategy of their richest neighbor.

Mean field approach

Now, we consider a mixed population of N agents with, at the beginning, an equal density of cooperators and defectors. Under the hypothesis that all agents interact together, at each time step the payoffs gained by cooperators and defectors are computed as follows

$$
\begin{cases}
\pi_c = (\rho_c N - 1) + (\rho_d N)S \\
\pi_d = (\rho_c N)T
\end{cases} \tag{2}
$$

with $\rho_c + \rho_d = 1$, ρ_c density of cooperators and ρ_d density of defectors. We recall that defection is the dominant strategy in the PD and, even if we set $S = 0$ and $T = 1$, it corresponds to the final equilibrium because π_d is always greater than π_c. At this point, it is important to highlight that previous investigations [19–21] have been performed by 'memoryless' agents (i.e., agents that do not accumulate the payoff over time) whose interactions were defined only with their neighbors, and focusing only on one agent (and on its neighbors) at a time. These conditions are fundamental. For instance, if at each time step we randomly select one agent interacting only with its neighbors, there exists the probability to select consecutively a number of close cooperators; thus, in this occurrence, very rich cooperators may emerge and then prevail on defectors, even without introducing mechanisms like motion. It is also worth observing that as $P = 0$, a homogeneous population of defectors does not increase its overall payoff. Instead, according to the matrix (1), a cooperative population continuously increases its payoff over time.

Now, we consider a population divided into two groups by a wall: a group G^a composed of cooperators, and a mixed group G^b, i.e., composed of cooperators and defectors in equal amount. Agents interact only with members of the same group, then the group G^a never changes and, in addition, it strongly increases its payoff over time. The opposite occurs in the group G^b, as it converges to an ordered phase of defection, limiting its final payoff. Remarkably, in this scenario, we can introduce a strategy to modify the equilibria of the two groups. In particular, we can both change to cooperation the equilibrium of G^b, and to defection that of G^a. In the first case, we have to wait a while, before moving one or few cooperators to G^b, so that defectors increase their payoff, but during the revision phase they change strategy to cooperation as the newcomers are richer than them. In the second case, if we move after few time steps a small group of defectors from G^b to G^a, the latter converges to a final defection phase. These preliminary and theoretical observations let emerge an important property of the 'memory-aware' PD: considering the two different groups, cooperators may succeed when act after a long time and individually. Instead, defectors may succeed acting fast and in group. Notably, rich cooperators have to move individually since otherwise many rich cooperators risk to increase too much the payoff of defectors that, in this case, will not change strategy. The opposite holds for defectors that, acting in group, may strongly reduce the payoff of a community of cooperators (for $S < 0$).

Mapping agents to gas particles

We hypothesize that the spatial PD, with moving agents, can be successfully studied by the framework of kinetic theory [30]. Therefore, in the proposed model, we map agents to particles of a gas. In doing so, the average speed of particles is computed as $\langle v \rangle = \sqrt{\frac{3 T_s k_b}{m_p}}$, with T_s system temperature, k_b Boltzmann constant, and m_p particle mass. Particles are divided into two groups by a permeable wall, so that it can be crossed by particles, but it avoids interactions among particles belonging to different groups. Now, it is worth emphasizing that we can provide a dual description of our system: one in the 'physical' domain of particles, the other in the 'information' domain of agents. Notably, to analyze the system in the 'information' domain we will introduce, as above discussed, the mapping of agents to a spin system (see [33]). Summarizing, we map agents to gas particles in order to represent their 'physical' property of motion, and we map agents to spins for representing their 'information' property (i.e., their strategy). Remarkably, these two mappings can be viewed as two different layers for studying how the agent population evolves over time. Although the physical property (i.e., the motion) affects the agent strategy (i.e., its spin), the equilibrium can be reached in both layers/domains independently. This last observation is important since we are interested in evaluating only the final equilibrium reached in the 'information' domain. Then, as stated before, agents interact only with those belonging to the same group, so the evolution of the mixed group G^b can be described by
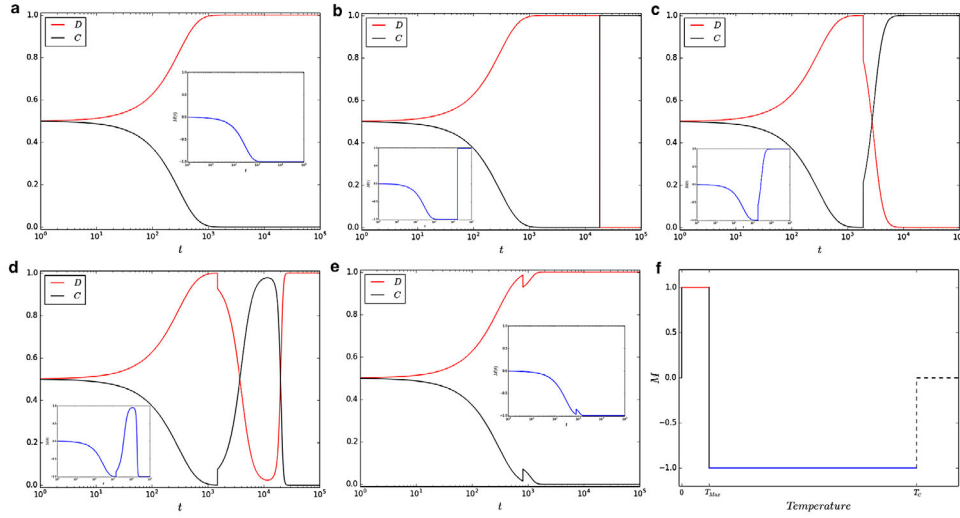
Fig. 1. From (a) to (e): evolution of the group G^b, with $N = 100$ and $\epsilon = 1$, on varying the temperature: (a) $T_s = 0$; (b) $T_s = 0.1$; (c) $T_s = 9$; (d) $T_s = 15$; (e) $T_s = 50$. Insets show the system magnetization over time. The istant $t = t_c$, can be detected in plots (c)–(e) as a discontinuity of the two lines (i.e., red and black). (f) Final magnetization M, of G^b, for different temperatures (T_c indicates the 'critical temperature').

following equations

$$\begin{cases} \frac{d\rho_c^b(t)}{dt} = p_c^b(t)\rho_c^b(t)\rho_d^b(t) - p_d^b(t)\rho_d^b(t)\rho_c^b(t) \\ \frac{d\rho_d^b(t)}{dt} = p_d^b(t)\rho_d^b(t)\rho_c^b(t) - p_c^b(t)\rho_c^b(t)\rho_d^b(t) \\ \rho_c^b(t) + \rho_d^b(t) = 1 \end{cases} \quad (3)$$

with $p_c^b(t)$ probability that cooperators prevail on defectors (at time t), and $p_d^b(t)$ probability that defectors prevail on cooperators (at time t). These probabilities are computed according to the payoffs obtained, at each time step, by cooperators and defectors

$$\begin{cases} p_c^b(t) = \frac{\pi_c^b(t)}{\pi_c^b(t)+\pi_d^b(t)} \\ p_d^b(t) = 1 - p_c^b(t). \end{cases} \quad (4)$$

The system (3) can be analytically solved provided that, at each time step, values of $p_c^b(t)$ and $p_d^b(t)$ be updated. So, the density of cooperators reads

$$\rho_c^b(t) = \frac{\rho_c^b(0)}{\rho_c^b(0) - [(\rho_c^b(0) - 1)e^{\frac{\tau t}{N^b}}]} \quad (5)$$

with $\rho_c^b(0)$ initial density of cooperators in G^b, $\tau = p_d^b(t) - p_c^b(t)$, and N^b number of agents in G^b. Recall that setting $T_s = 0$, not allowed in a thermodynamic system, corresponds to a motionless case, leading to the Nash equilibrium in G^b. Instead, for $T_s > 0$ we can find more interesting scenarios. Now we suppose that, at time $t = 0$, particles of G^a are much closer to the wall than those of G^b (later we will relax this constraint); for instance, let us consider a particle of G^a that, during its random motion, it is following a trajectory of length d (in the n-dimensional physical space) towards the wall. Assuming this particle is moving with speed equal to $\langle v \rangle$, we can compute the instant of crossing $t_c = \frac{d}{\langle v \rangle}$, i.e., the instant when it moves

from G^a to G^b. Thus, on varying the temperature T_s, we can vary t_c.

Let us consider the payoff of cooperators in the two groups. Each cooperator in G^a gains

$$\pi_c^a = (\rho_c^a N^a - 1)t. \quad (6)$$

On the other hand, the situation for cooperators in G^b is much more different as, according to the Nash equilibrium, their amount decreases over time. Therefore, we can consider how changes the payoff of the last cooperator survived in G^b

$$\pi_c^b = \sum_{i=0}^{t} [(\rho_c^b N^b - 1) + (\rho_d^b N^b)S]_i \quad (7)$$

moreover, $\pi_c^b \to 0$ as $\rho_c^b \to 0$. At $t = t_c$, a new cooperator reaches G^b, with a payoff computed with equation (6).

3 Results

The analytical solution (5) allows to analyze the evolution of the system and to evaluate how initial conditions affects the outcomes of the model. Let us observe that, if $\pi_c^a(t_c)$ is enough big, the new cooperator may modify the equilibrium of G^b, turning defectors to cooperators. Notably, the payoff considered to compute p_c^b, after t_c, corresponds to $\pi_c^a(t_c)$, as the newcomer is the richest cooperator in G^b. Furthermore, we note that $\pi_c^a(t_c)$ depends on N^a, hence we study the evolution of the system on varying the parameter $\epsilon = \frac{N^a}{N^b}$, i.e., the ratio between particles in the two groups. Eventually, for numerical convenience, we set $k_b = 1 \times 10^{-8}$, $m_p = 1$, and $d = 1$.

Figure 1 shows the evolution of G^b, for $\epsilon = 1$ on varying T_s and, depicted in the inner insets, the variation of
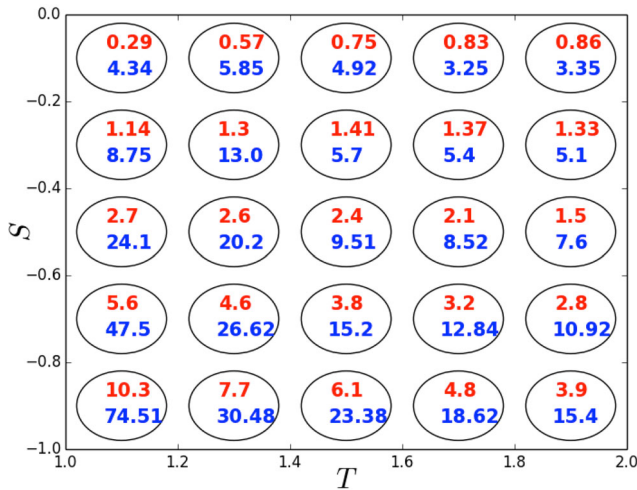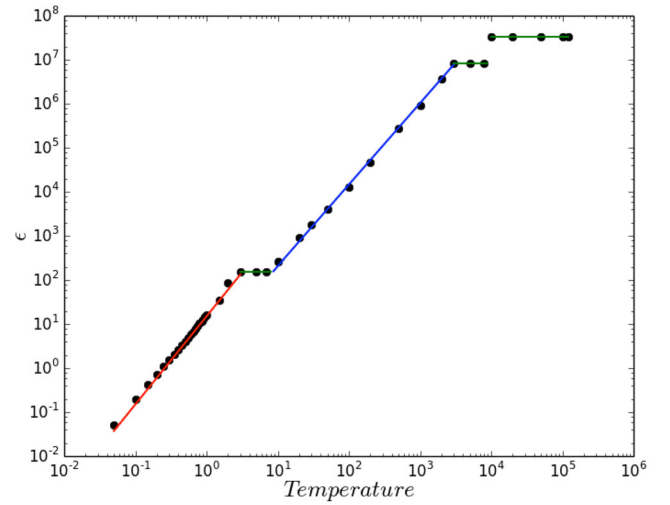
Fig. 2. Maximum values of temperature T_s that allow the group G^b to converge to cooperation. Red values correspond to results computed with $\epsilon = 0.5$, while blue values to those computed with $\epsilon = 1$. Circles are placed in the TS diagram indicating values of T and S, of the payoff matrix, used for each case. Even for high values of T, and small values of S, it is possible to achieve cooperation.

system magnetization over time (always inside G^b) computed as [34]

$$M = \frac{\sum_{i=1}^{N^b} \sigma_i}{N^b} \qquad (8)$$

with σ_i strategy of the i-agent. As discussed before, in the physical domain of particles, heating the system entails the average speed of particles increases. Thus, under the assumption that two agents play together if they stay close (i.e., in the same group) for a long enough time, we hypothesize that exists a maximum speed such that for greater values interactions do not occur (in terms of game). This hypothesis requires a critical temperature T_c, above which no interactions, in the 'information' domain, are possible. As shown in Figure 1f, for temperatures in range $0 < T_s < T_{\max}$ the system converges to a cooperation phase (i.e., $M = +1$), for $T_{\max} < T_s < T_c$ the system follows the Nash equilibrium (i.e., $M = -1$), and for $T > T_c$ a disordered phase emerges at equilibrium. Remarkably, results of our model suggest that it is always possible to compute a range of temperatures to obtain an equilibrium of full cooperation (see Fig. 2). Moreover, we study the variation of $T_{\max}$ on varying ϵ (see Fig. 3) showing that, even for low ϵ, it is possible to obtain a time t_c that allows the system to converge towards cooperation. Eventually, we investigate the relation between the maximum value of T_s that allows a population to become cooperative and its size N (i.e., the number of agents). Remarkably, as shown in Figure 4, the maximum T_s scales with N following a power-law function characterized by a scaling parameter (i.e., an exponent) $\gamma \sim 2$. The value of γ has been computed by considering values of T_s shown in Figure 2 for the case $\epsilon = 2$. Eventually, it is worth to highlight that all analytical results let emerge a link

Fig. 3. Maximum value of system temperature that allows to achieve cooperation at equilibrium versus ϵ (i.e., the ratio between particles in the two groups). Different colors identify different trends, fitted by power-law functions. After the final green plateau, temperatures are too high to play the spatial PD.
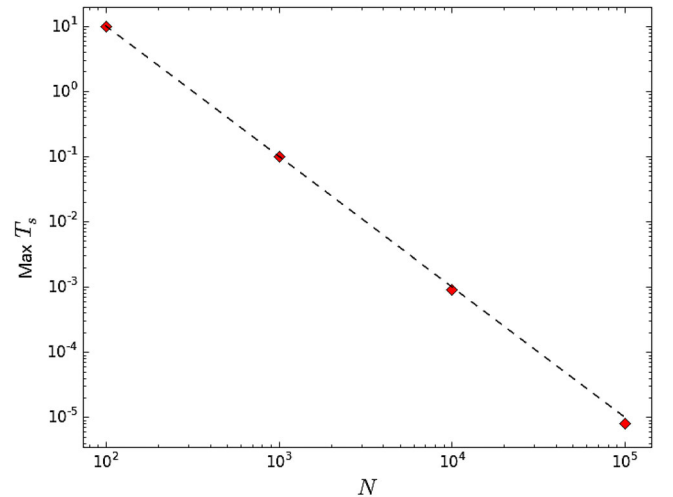

Fig. 4. Maximum value of T_s to achieve full cooperation at equilibrium in function of N, i.e., the size of the population. The fitting function (dotted line) is a power-law characterized by a scaling parameter equal to 2.

between the system temperature and its final equilibrium. Recalling that we are not considering the equilibrium of the gas, i.e., it does not thermalize in the proposed model, we emphasize that the equilibrium is considered only in the information domain.

Phase transitions in the spatial PD

As discussed before, in the information domain we can study the system by mapping agents to spins, whose value represents their strategy. In addition, we can map the difference between winning probabilities, of cooperators and
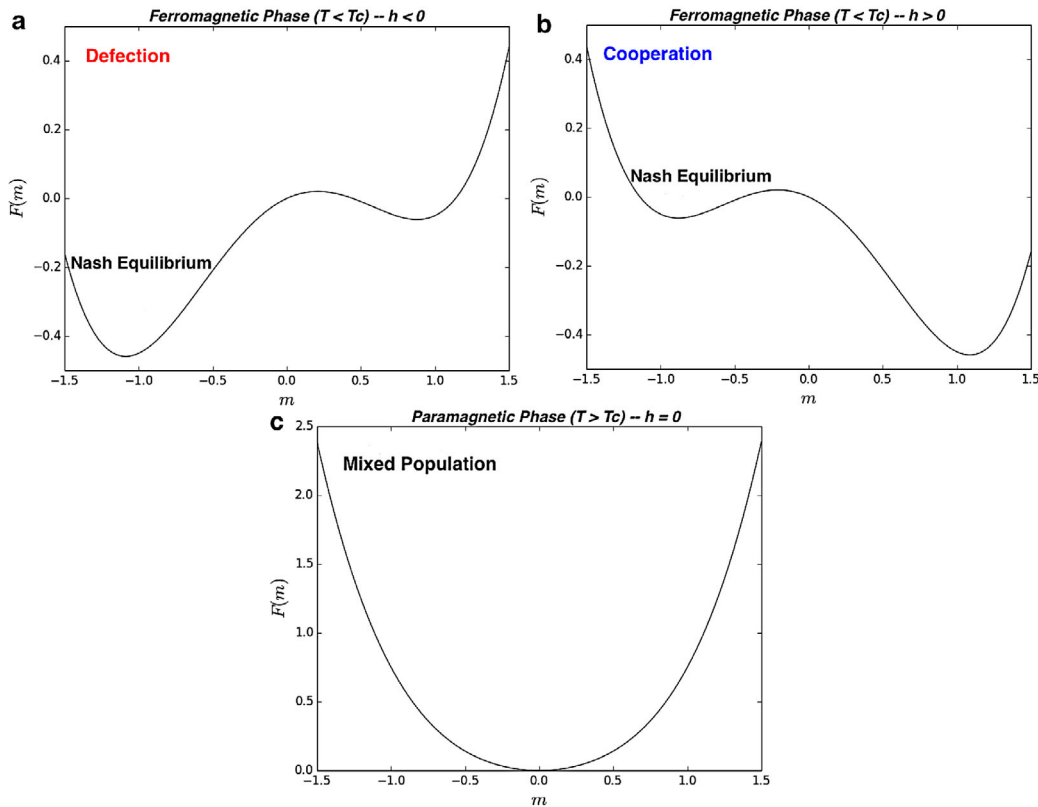
Fig. 5. Order-disorder phase transitions in the population. For $T_s < T_c$, the population is in a ferromagnetic phase: (a) applying an external negative field, the system converges to the Nash equilibrium, corresponding to $m = -1$ (as $\sigma = -1$ represents defection); (b) applying an external positive field, the population converges to cooperation ($\sigma = +1$), corresponding to $m = +1$. (c) For temperatures higher than T_c, a disordered paramagnetic phase emerges.

defectors, to an external magnetic field: $h = p_c^b - p_d^b$. In doing so, by the Landau theory [30], we can analytically identify an order-disorder phase transition. Notably, we analyze the free energy F of the spin system on varying the control parameter m [35] (corresponding to the magnetization M)

$$F(m) = -hm \pm \frac{m^2}{2} + \frac{m^4}{4} \qquad (9)$$

where the sign of the second term depends on the temperature, i.e., positive for $T_s > T_c$ and negative for $T_s < T_c$; recalling that T_c represents the temperature beyond which it is not possible to play the PD due to the high particles speed (according to the condition before discussed). For the sake of clarity, we want to emphasize that the free energy is introduced in order to evaluate the nature of the final equilibrium achieved by the system. In particular, looking for the minima of F allows to investigate if our population reaches the Nash equilibrium, or different configurations (e.g., full cooperation). Figure 5 shows a pictorial representation of the phase transitions that occur in our system, on varying T_s and the external field h. Finally, the constraints related to the average speed of particles, and to the distance between each group and the permeable wall, can in principle be relaxed as we can imagine to extend this description to a wider system with several groups (as done in previous investigations, e.g. [20]), where agents are uniformly spread in the whole space. It is worth to highlight that our results are completely in agreement with those achieved by authors who studied the role of motion in the PD (as [19,20]), explaining why clusters of cooperators emerge in their simulations [20]. We also recall that, in the proposed model, we are using memory-aware agents, while in previous computational investigations agents reset their payoff at each step, i.e., before to start new interactions.

4 Conclusions

To conclude, in this work we provide an analytical description of the spatial Prisoner's Dilemma, by using the framework of statistical physics, studying the particular case of agents provided with memory of their payoff (defined memory-aware agents). This condition entails that their payoff is not reset at each time step, so that they can increase it over time. In particular, we propose a model based on the kinetic theory of gases, showing how motion may lead a population towards an equilibrium far from the expected one (i.e., the Nash equilibrium). Remarkably, the final equilibrium depends on the system temperature, so that we have been able to identify a range of temperatures that triggers cooperation for all values of the payoff

matrix (related to the PD). In addition, we found an interesting relation between the maximum temperature that fosters cooperation and the size of the system. Notably, a scaling parameter in that relation has been computed by investigating different orders of magnitude of the size of the system. Furthermore, the dynamics of the resulting model have been also described in terms of order-disorder phase transitions. Finally, we deem that our results open the way to define a direct link between evolutionary game theory and statistical physics.

References

1. M. Perc, P. Grigolini, Chaos Solitons Fractals **56**, 1 (2013)
2. M.A. Nowak, *Evolutionary Dynamics: Exploring the Equations of Life* (Harvard University Press, 2006)
3. M. Tomassini, Introduction to evolutionary game theory, in *Proc. Conf. on Genetic and evolutionary computation companion* (2014)
4. P.C. Julia, J. Gomez-Gardenes, A. Traulsen, Y. Moreno, New J. Phys. **11**, 083031 (2009)
5. L.M. Floria, C. Gracia-Lazaro, J. Gomez-Gardenes, Y. Moreno, Phys. Rev. E **79**, 026106 (2009)
6. J. Hofbauer, K. Sigmund, *The Theory of Evolution and Dynamical Systems* (Cambridge University Press, 1988)
7. A.M. Colman, *Game Theory and Its Applications* (Digital Printing, 2008)
8. M. Perc, A. Szolnoki, Phys. Rev. E **77**, 011904 (2008)
9. A. Szolnoki, M. Perc, J. R. Soc. Interface **12**, 20141299 (2015)
10. Z. Wang, A. Szolnoki, M. Perc, Sci. Rep. **3**, 1183 (2013)
11. A. Szolnoki, N.-G. Xie, C. Wang, M. Perc, Europhys. Lett. **96**, 38002 (2011)
12. M. Perc, A. Szolnoki, New J. Phys. **14**, 043013 (2012)
13. D. Friedman, J. Evol. Econ. **8**, 15 (1998)
14. S. Schuster, L. de Figueiredo, A. Schroeter, C. Kaleta, BioSystems **105**, 147 (2011)
15. E. Frey, Physica A **389**, 4265 (2010)
16. F. Fu, D.I. Rosenbloom, L. Wang, M.A. Nowak, Proc. R. Soc. B **278**, 42 (2011)
17. E. Lieberman, C. Hauert, M.A. Nowak, Nature **433**, 312 (2005)
18. S. Galam, B. Walliser, Physica A **389**, 481 (2010)
19. S. Meloni, A. Buscarino, L. Fortuna, M. Frasca, J. Gomez-Gardenes, V. Latora, Y. Moreno, Phys. Rev. E **79**, 067101 (2009)
20. A. Antonioni, M. Tomassini, P. Buesser, J. Theor. Biol. **344**, 40 (2014)
21. M. Tomassini, A. Antonioni, J. Theor. Biol. **364**, 154 (2015)
22. A. Antonioni, M. Tomassini, A. Sanchez, Sci. Rep. **5**, 10282 (2015)
23. M. Perc, J. Gomez-Gardenes, A. Szolnoki, L.M. Floria, Y. Moreno, J. R. Soc. Interface **10**, 20120997 (2013)
24. M.A. Javarone, A.E. Atzeni, Comput. Soc. Netw. **2**, 15 (2015)
25. M.A. Javarone, A.E. Atzeni, S. Galam, Lect. Notes Comput. Sci. **9028**, 155 (2015)
26. M.A. Nowak, Science **314**, 1560 (2006)
27. G. Szabo, G. Fath, Phys. Rep. **446**, 97 (2007)
28. M.A. Nowak, R.M. May, Nature **359**, 826 (1992)
29. C. Hauert, G. Szabo, Am. J. Phys. **73**, 405 (2005)
30. K. Huang, *Statistical Mechanics*, 2nd edn. (Wiley, 1987)
31. A. Szolnoki, G. Szabo, M. Perc, Phys. Rev. E **83**, 0361101 (2011)
32. A. Szolnoki, M. Perc, Europhys. Lett. **92**, 38003 (2010)
33. M.A. Javarone, Europhys. Lett. **110**, 58003 (2015)
34. M. Mobilia, S. Redner, Phys. Rev. E **68**, 046106 (2003)
35. A. Barra, J. Stat. Phys. **132**, 787 (2008)