Effect of memory on the prisoner's dilemma game in a square lattice

Shao-Meng Qin,[1] Yong Chen,[1,2,*] Xiao-Ying Zhao,[1] and Jian Shi[1]

[1]*Institute of Theoretical Physics, Lanzhou University, Lanzhou 730000, China*

[2]*Key Laboratory for Magnetism and Magnetic Materials of the Ministry of Education, Lanzhou University, Lanzhou 730000, China*

We have studied the effect of memory on the evolution of the prisoner's dilemma game in square lattice networks. Based on extensive simulations, we found that the density of cooperators was enhanced by an increasing memory effect for most parameters. However, we also observed that the density of cooperators decreased with an increased memory effect in the case of a large memory and moderate temptation. It is interesting to note that memory makes cooperators immune from temptation. The strength of protection reaches its maximal value only for a moderate memory effect.

I. INTRODUCTION

The evolutionary prisoner's dilemma game (PDG) has attracted substantial attention over the past few decades [1]. In this game, two agents must simultaneously select one of two strategies: cooperation or defection. The prisoners receive payoffs that are dependent on their choices. Selfish agents will adapt their strategy to maximize their payoff. Game theory involves the construction of many types of models and analysis of these models using varied parameters. Therefore, game theory serves as a powerful metaphor for simulation of the interactions between individuals in many domains, including biology, economy, and ecology.

In the PDG, mutual cooperation generates the highest return for the community. However, the Nash equilibrium state is mutual defection because defection is a better choice for the prisoner, regardless of the strategy of the other prisoner. Importantly, in the real world, mutual cooperation is the most commonly utilized strategy. Systems such as the PDG are considered to be an important tool for studying the emergence of cooperative behavior between selfish individuals [2–4]. Nowak and May [5] introduced a spatial prisoner's dilemma game (SPDG) consisting of a two-state cellular automaton. In the general SPDG, the agents in the game play the PDG with their network neighbors and get payoffs according to a payoff matrix. The total payoff of each agent is the sum of all payoffs in this step. An agent may then mimic his neighbor's strategy by comparing his payoffs in this step with his neighbor's payoffs. An important conclusion is that spatial structure can promote the persistence of cooperation. Because the interactions of an agent are limited to its local neighbors, PDG models have been extensively explored in the past few years [1,6–8]. In addition to spatial structure, there are several mechanisms that may facilitate the emergence and persistence of cooperation among populations. Hamilton found that kin selection can favor cooperation [9]. Axelrod's model demonstrated that the *tit for tat* strategy could sustain cooperation in systems of all players playing the game together. The simulation performed by Szabó, Vukov, and Szolnoki provided evidence that noise and irra-

tional choices affect the maintenance of cooperative behavior [10].

In the traditional SPDG model, the changing probability of strategy is determined by the agents' performance on one step. In other words, people assume that the agents are shortsighted and forgetful. In fact, when people make an important decision, they generally consider the current situation and their experiences. Therefore, the effect of memory should be taken into account. Historical memory plays a key role in the evolutionary game [11]. The purpose of this paper is to evaluate whether memory enhances the density of cooperators and protects the cooperators from the temptation. We observed the maximum value of critical points from a homogeneous cooperator to a mixed state of cooperator and defector.

In this paper, we consider an evolutionary SPDG with the memory effect in a square lattice, in which players update their strategy by considering previous payoffs. The rules of the game are explained in Sec. II. The simulations, which are detailed in Sec. III, show that the evolution of the SPDG depends on the magnitude of the memory effect and payoff-matrix elements. Conclusions are drawn in the last section.

II. MODEL

In the traditional PDG, there are two players. Each player chooses one of two strategies: *cooperator* (*C*) or *defector* (*D*). There are four combinations for the two players, (C,C), (C,D), (D,C), and (D,D), which corresponded to payoffs (R,R), (S,T), (T,S), and (P,P). The rewards or punishments for each player can be tabulated as 2×2 payoff matrices (see Table I).

Four elements in the payoff matrix satisfy the ranking order $T > R > P > S$ and the additional constraint $T + S < 2R$

TABLE I. The payoff matrix of the prisoner's dilemma game.

Player 1 (player 2)	C	D
C	R (R)	S (T)
D	T (S)	P (P)

*Corresponding author. ychen@lzu.edu.cn

for repeated interactions. As suggested by Nowak and May [5], the parameters in this paper are $R=1$, $T=b$, $S=0$, and $P=0$. Our model preserves the essentials of the PDG and b is the only tunable parameter.

Our study is based on systematic Monte Carlo (MC) simulations on a square lattice network with periodic boundary conditions. When we applied the PDG on the network, the players were located on the nodes. In every MC step, the players simultaneously play the PDG with their network neighbors (only the first neighbors) and themselves. The sum payoff of each player is the sum over all games. The evolutionary process is governed by strategy imitation. In every MC step, all agents may mimic their neighbors' strategy. Player i adopts a (randomly chosen) neighbor's strategy (at site j) with a probability that depends upon the payoff difference:

$$W = \frac{1}{1 + \exp\{[E_m(i) - E_m(j)]/\kappa\}}, \quad (1)$$

where κ indicates the noise generated by the players, allowing irrational choices [12,13]. In this work, we use $\kappa=0.1$ for all simulations. $E_m(i)$ and $E_m(j)$ are the total payoffs, which contain the sum of payoffs at this MC step U and the cumulative historical payoff. For each node i, there are two memories $M_c(i,t)$ and $M_d(i,t)$ at step t. When node i is associated with the strategy C and the sum payoff at this MC step is U,

$$E_m(i) = U + M_c(i,t),$$

$$M_c(i,t+1) = (M_c(i,t) + U)\tau,$$

$$M_d(t+1) = M_d(t)\tau, \quad (2)$$

for this time step. When the node i is associated with strategy D,

$$E_m(i) = U + M_d(i,t),$$

$$M_c(i,t+1) = M_c(i,t)\tau,$$

$$M_d(i,t+1) = [U + M_d(i,t)]\tau. \quad (3)$$

Here, τ is the memory factor and $0 \leq \tau < 1$. $M_c(i,t)$ and $M_d(i,t)$ represent the historical payoffs of C and D, respectively. The memory effect for each MC step declines with time. In other words, the memories of the payoffs, $M_c(i,t)$ and $M_d(i,t)$, will be lost as time passes. $\tau=0$ indicates that there is no memory effect. As τ nears 1, there exists an almost perfect memory effect in the model. Starting from a random initial state with an equal fraction of C and D and $M_c(i,0)=M_d(i,0)=0$, we iterate the model with a synchronized update.

III. SIMULATION RESULTS

Our simulations are carried out by varying b and τ. The results described in this paper are obtained from MC simulations with a system size of 200×200, with the exception of the results shown in Fig. 4. It is true that a network with
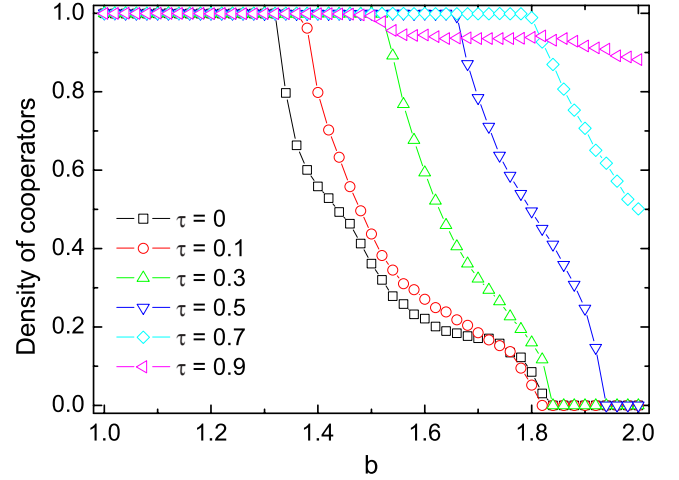

FIG. 1. (Color online) Density of cooperators, f_c, as a function of the payoff parameter b with various memory factors τ.

larger size will decrease the ensemble error, which is caused by the finite scale of the network. We have simulated our model with 100×100 and 400×400 systems. There is no conspicuous difference between these networks. The results in this paper are the average of 20 trials with various random seeds. Repeating simulations with different random seeds can also reduce the error. Therefore, the 200×200 system is large enough. The transient time is varied from 20 000 to 80 000 MC steps. After the transient state, the system reached a stable state, and the amplitudes of population fluctuations were considerably smaller than the corresponding average value.

To characterize the macroscopic behavior of the system, we measure the density of C, f_c first. Figure 1 shows the f_c on a square lattice as a function of b for several values of τ. We find that there are two thresholds of temptation, b. When $b < b_{c1}$, the networks in which only C can survive are in the stable state. The density of C decreases monotonically with increasing values of b for $b > b_{c1}$. Movies that present how the systems with network size of 240×240 evolve in 300 MC steps after a transient time for different b and τ are available online [14]. In these movies the C's are presented as black boxes and the D's as red boxes. It is observed that the agents who utilize the same strategies join together to form complex patterns that continuously move and change shape. These patterns develop because agents change their strategies by learning from their neighbors. Furthermore, the C's who join together are more stable because they support each other by earning payoffs from their C neighbors. For $b > b_{c2}$, the C strategies die out. Both the memory factors τ and κ affect the critical point [10]. Recently, in Ref. [15] Szabó, Vukov, and Szolnoki showed the κ-b plane of Newman-Watts networks. In contrast to τ, κ does not conspicuously affect b_{c1} or b_{c2} in this model. The main focus of this paper is to evaluate how the memory effect τ affects the density of C and b_{c1}. Determination of the κ-b_{c1} and κ-b_{c2} planes is beyond the scope of this paper. b_{c2} increases with τ monotonically; however, b_{c1} reaches its maximum value near $\tau=0.72$, and b_{c1} tends toward $4/3$ as τ approaches 0 or 1 (see black squares in Fig. 2). From Fig. 1, we find that the
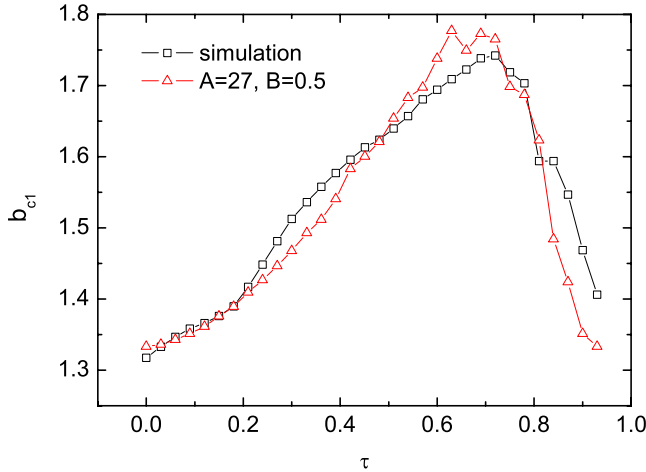
FIG. 2. (Color online) b_{c1} as a function of memory factor τ on the square lattice. The data points depicted by squares (black) are the result of MC simulations, and the data points depicted by triangles (red) were derived from Eq. (6).

memory effect enhances the density of C in most cases; however, the density of C decreases with increasing τ only for $\tau > 0.72$ and $1.75 < b < 1.8$. It should be noted that our simulations are consistent with those presented in Fig. 1 of Ref. [12] for $\tau = 0$ despite the fact that Szabó and Töke used the asynchronized update law in their model. The mean-field results for six-point approximations [16] agree with the simulation in [12] and our model in the case of $\tau = 0$. We assume that the six-point approximation includes the main features of the two models. Importantly, the six-point approximation does not contain a restriction of the update law. Therefore, it is conjectured that the synchronized update does not play an important role in the two models.

In comparison to the case of $\tau = 0$, we know that enhancement of the density of C is caused by M_c and M_d. From the above-mentioned definitions, the M_c and M_d of one node are determined by two factors: (1) the payoff income U of every MC step and (2) whether the node maintains one strategy. M_c or M_d is aggravated if the node persists in C or D, respectively. Figure 3 plots the average M_c and M_d of all nodes as a function of b. It should be noted that M_c is always larger than M_d. Therefore the memory effect almost always enhances f_c in this model. For $b < b_{c1}$, the networks include only C. Every node can receive five payoffs at every MC step and M_c is $5 * \tau/(1 - \tau)$. Then, with an increase of b, the emergence of D reduces the value of C's payoff for every MC step and decreases the continuous accumulation of M_c. As a result, M_c gradually decreases with b until C dies out and $M_c = 0$. In contrast to M_c, M_d has a peak in the C-D coexisting states. When D is outside the mixed region, M_d is equal to 0. D earns a payoff only by playing the game with C. Therefore, M_d is not equal to 0 in the C-D coexistence region $b_{c1} < b < b_{c2}$. When b is a little bit larger than b_{c1} and $1 - f_c \ll 1$, D forms small isolated gangs. As discussed in [12], the behaviors of D gangs are considered as branching and annihilating random walkers [17,18]. The D gangs undergo four basic processes: random walk; an annihilation reaction (two D gangs can unite); death (one gang of D will die due
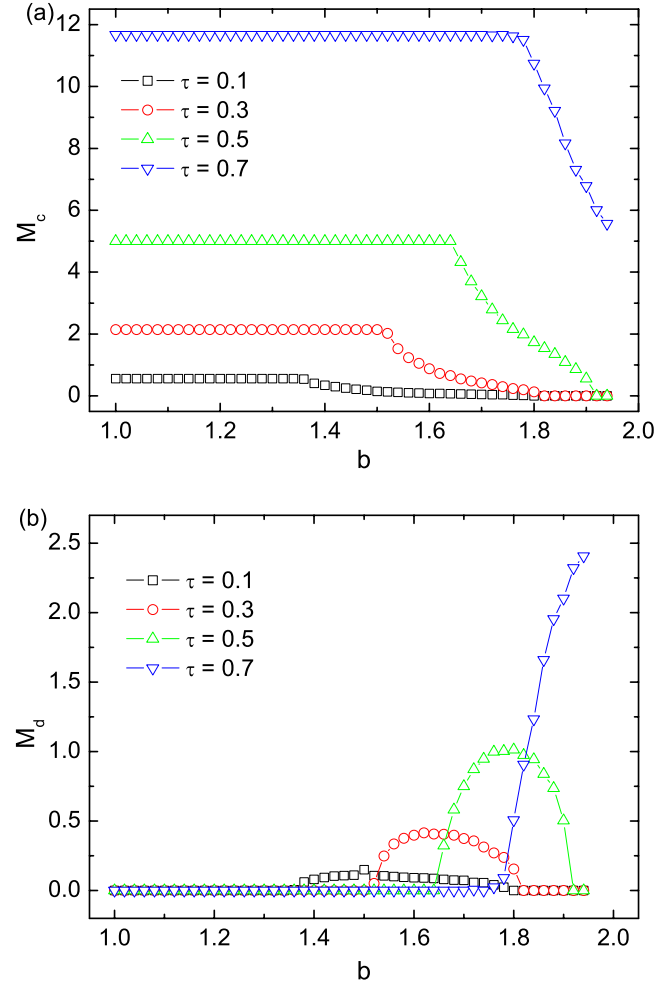

FIG. 3. (Color online) Average payoffs for strategies C and D as a function of the payoff parameter b for several values of the memory factor τ.

to the irrational choice); and branching (one gang of D can divided into two gangs). Every D gang that obtains the highest payoff at every MC step is surrounded by cooperators. However, the density of D is low, and random walking breaks the continual accumulation of M_d. Therefore, M_d is small. When D is dominant, the random walking of C gangs does not deplete the accumulation of M_d but the average payoff of D decreases at each MC step. Thus M_d is maximized when there is a compromise between the average payoff at each MC step and continual accumulation of M_d.

In [12,15], the authors discussed the critical exponents of b_{c1} and b_{c2}. Their MC simulations indicated a power-law behavior, namely, $f_c \propto (b_{c2} - b)^\beta$ and $1 - f_c \propto (b - b_{c1})^\beta$, and the values of β agreed with the directed percolation (DP) exponent. Grassberger and Janssen conjectured that all one-component models with a single absorbing state belong to the universality class of DP [19]. The value of critical exponents should be independent of the details of dynamical rules and dependent on the spatial dimension. In this paper, we investigated these exponents in the context of different values of τ. Figure 4 shows that β, which ranged from 0.47 ($\tau = 0$) to 1.108 25 ($\tau = 0.9$), monotonically increases with τ. Therefore, the value of the critical exponent is not universal
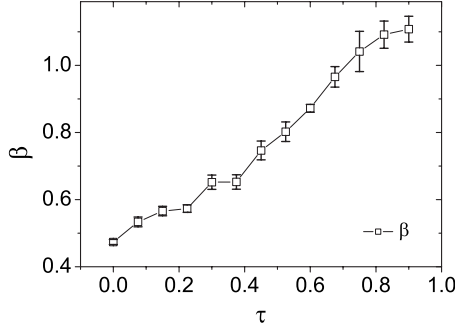
FIG. 4. Critical exponent β as a function of τ of b_{c1}. The error bar in the figure presents the standard deviation. In order to suppress the statistical error in the critical regions, we use the system size 600×600 for $\tau \leq 0.6$, 800×800 for $0.75 \geq \tau > 0.6$, and 1000×1000 for $\tau > 0.75$.

but depends on the memory factor τ in this model.

Considering that a persistent unchanged strategy at one site leads to the accumulation history payoff, we investigated the mobility of spatial patterns of τ. Population mobility is a central feature of real ecosystems: animals migrate, bacteria run and tumble. Similar phenomena can be observed in a rock-paper-scissors game [20]. Reichenbach, Mobilia, and Frey observed that mobility has a critical influence on species diversity. In this model, we find that the behavior of b_{c1} is caused by the decrease in mobility strategy. This means that C resists temptation b by decreasing mobility. Therefore, we introduce the time autocorrelation function of the strategy:

$$g(\tau, b, t) = \langle s_i(0) s_i(t) \rangle, \quad (4)$$

where $s_i(t)$ is the strategy of player i at MC step t. When player i chose C, $s_i(t) = 1$. In contrast, $s_i(t) = -1$ for D. $\langle \rangle$ denotes an average over all nodes in the network. Considering that $g(\tau, b, t)$ can be affected by the density of C and, in order to ensure that $g(\tau, b, t)$ ranges from 0 to 1, we chose b such that $f_c = 0.5$. This definition describes whether the node's recent strategy correlates with its strategy at later t MC steps.

Figure 5(a) displays the attenuation of $g(\tau, t)_{f_c=0.5}$ with time. It was found that $g(\tau, t)_{f_c=0.5}$ fits the form $g(\tau, t)_{f_c=0.5} = \exp[-t/\rho(\tau)]$. One can regard ρ as the characteristic residence time of the unaltered strategy. We define t_h as the number of MC steps for which one strategy was maintained and assume that the characteristic residence time ρ and t_h have a similar ratio:

$$t_h = \rho/A + B. \quad (5)$$

Figure 5(b) shows ρ as a function of the parameter τ. There is a critical behavior at $\rho \propto (1-\tau)^{-z}$, where the exponent is $z = 2.22$ with standard deviation 0.043.

Now we focus our attention on the behavior of b_{c1}. When $b > b_{c1}$, C cannot resist temptation b and D appears. Therefore, b_{c1} can be regarded as the ability of the model to protect C. As described in the discussion above, the D gangs undergo four basic processes. When $b = b_{c1}$ and $1 - f_c \ll 1$, the annihilation process is rare, while the death and branching
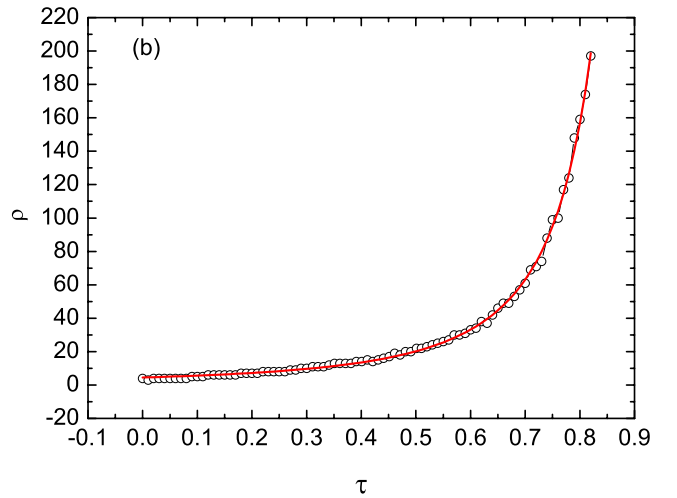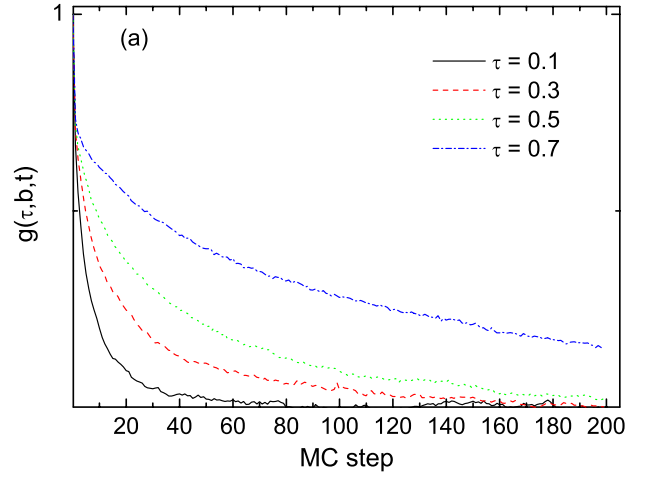

FIG. 5. (Color online) (a) Time autocorrelation function of strategy for several values of τ. (b) Characteristic time $\rho(\tau)$ as a function of τ. The red line is the fitting result of this figure $\rho = 4.29 * (1 - \tau)^{-2.22}$.

processes are major activities. Therefore, D gangs become stable if the branching rate is greater than the death rate. We found that the single D in the branching process will have an offspring and form D-D pairs (as shown in Fig. 6). The D-D pair plays an important role in the branching process of D
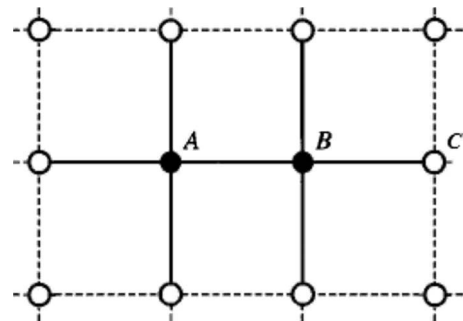

FIG. 6. Illustration of the D-D pair (nodes A and B) and neighbor C (node C). The black and white circles denote D and C, respectively. A D-D pair indicates that both nodes are connected in the networks by strategy D.

gangs. When we discard the effect of noise, the total payoff E_m of each player in D-D pairs (nodes A and B in Fig. 6) must be larger than the payoff of their neighbor C (node C in Fig. 6). Otherwise, the D gangs will eventually die. For example, in the case of $\tau=0$, the total payoff of each player in D-D pairs is $3b$, and the total payoff of their neighbor C is 4. Therefore, the threshold for a stable D is $b_{c1}=4/(3b)=4/3$. We suggest that the deviation of $4/3$ which was obtained in our simulations was caused by noise.

The behavior of a stable D is subtle for $\tau>0$. Based on the discussion above, the increase of t_h with τ and t_h determines the player's memory and total payoffs. Therefore, we can use t_h to approximate b_{c1}. When the t_h of the D-D pair is N and we neglect the remnants M_d, which accumulated many MC steps ago, and assume that neighbor C can remain as C indefinitely because of the dominance of C at $b=b_{c1}$, we find that

$$b_{c1} = \frac{4}{3(1 - t_h^{N+1})}. \tag{6}$$

In Fig. 2, we plot the results from Eq. (6) which are similar to the simulation results. We use $A=27$ and $B=0.63$ in Eq. (5).

IV. CONCLUSION

In this paper, we studied the ability of memory to protect C from D in an evolutionary PDG in square lattice networks. With an increase in the effect of memory, there is an increase in the density of C in most cases. In compution of the autocorrelation function, we used the characteristic residence time to measure the mobility of a spatial pattern. We also found that the mobility of a spatial pattern decreases with increasing memory effect. Decreasing mobility induces a maximum value of the critical coexistence point b_{c1} at $\tau=0.72$. It is obvious that mobility plays an important role in this model. The effect of memory on cooperative behaviors may draw some attention in evolutionary games.

We have also applied this model to the Newman-Watts small-world (NWSW) network [21]. The NWSW network is a two-dimensional small-world network. We found that moderate long-range links did not have an obvious qualitative influence on our model.

[1] R. Axelrod, *The Evolution of Cooperation* (Basic Books, New York, 1984).

[2] A. M. Colman, *Game Theory and Its Applications in the Social and Biological Sciences* (Butterworth-Heinemann, Oxford, 1995).

[3] J. M. Smith, *Evolution and the Theory of Games* (Cambridge University Press, Cambridge, UK, 1982).

[4] J. W. Weibull, *Evolutionary Game Theory* (MIT Press, Cambridge, MA, 1995).

[5] M. A. Nowak and R. M. May, Nature (London) **359**, 826 (1992); Int. J. Bifurcation Chaos Appl. Sci. Eng. **3**, 35 (1993).

[6] F. C. Santos and J. M. Pacheco, Phys. Rev. Lett. **95**, 098104 (2005).

[7] C. Hauert and M. Doebeli, Nature (London) **428**, 643 (2004).

[8] G. Szabó and G. Fáth, Phys. Rep. **446**, 97 (2007).

[9] W. D. Hamilton, J. Theor. Biol. **7**, 17 (1964).

[10] G. Szabó, J. Vukov, and A. Szolnoki, Phys. Rev. E **72**, 047107 (2005).

[11] W.-X. Wang, J. Ren, G. R. Chen, and B.-H. Wang, Phys. Rev. E **74**, 056113 (2006).

[12] G. Szabó and C. Töke, Phys. Rev. E **58**, 69 (1998).

[13] G. Szabó and C. Hauert, Phys. Rev. E **66**, 062903 (2002).

[14] http://www.youtube.com/qsminside

[15] J. Vukov, G. Szabó, and A. Szolnoki, Phys. Rev. E **77**, 026109 (2008).

[16] G. Szabó and A. Szolnoki, Phys. Rev. E **53**, 2196 (1996).

[17] T. M. Liggett, *Interacting Particle Systems* (Springer, New York, 1985).

[18] J. L. Cardy and U. C. Täuber, Phys. Rev. Lett. **77**, 4780 (1996).

[19] P. Grassberger, Z. Phys. B: Condens. Matter **47**, 365 (1982); H. K. Janssen, *ibid.* **42**, 151 (1981).

[20] T. Reichenbach, M. Mobilia, and E. Frey, Nature (London) **448**, 1046 (2007).

[21] M. E. J. Newman and D. J. Watts, Phys. Rev. E **60**, 7332 (1999).