

Payoff components and their effects in a spatial three-strategy evolutionary social dilemmaJeromos Vukov,¹ Levente Varga,² Benjamin Allen,³ Martin A. Nowak,^{3,4} and György Szabó⁵¹*Research Center for Educational and Network Studies, Centre for Social Sciences, Hungarian Academy of Sciences, P. O. Box 20, H-1250 Budapest, Hungary*²*Babeş-Bolyai University, Faculty of Physics, RO-400084 Cluj-Napoca, Romania*³*Program for Evolutionary Dynamics, Harvard University, One Brattle Square, Cambridge, Massachusetts 02138, USA*⁴*Department of Mathematics, Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138, USA*⁵*Institute of Technical Physics and Materials Science, Centre for Energy Research, Hungarian Academy of Sciences, P. O. Box 49, H-1525 Budapest, Hungary*

(Received 3 April 2015; published 17 July 2015)

We study a three-strategy spatial evolutionary prisoner's dilemma game with imitation and logit update rules. Players can follow the always-cooperating, always-defecting or the win-stay-lose-shift (WSLS) strategies and gain their payoff from games with their direct neighbors on a square lattice. The friendliness parameter of the WSLS strategy—characterizing its cooperation probability in the first round—tunes the cyclic component of the game determining whether the game can be characterized by a potential. We measured and calculated the phase diagrams of the system for a wide range of parameters. When the game is a potential game and the logit rule is applied, the theoretically predicted phase diagram agrees very well with the simulation results. Surprisingly, this phase diagram can be accurate even in the nonpotential case if there are only two surviving strategies in the stationary state; this result harmonizes with the fact that all 2×2 games are potential games. For the imitation dynamics, we found that the effects of spatiality combined with the presence of two cooperative strategies are so strong that they suppress even substantial changes in the payoff matrix, thus the phase diagrams are independent of the cyclic component's intensity. At the same time, this type of strategy update mechanism supports the formation of cooperative clusters that results in a cooperative society in a wider parameter range compared to the logit dynamics.

DOI: [10.1103/PhysRevE.92.012813](https://doi.org/10.1103/PhysRevE.92.012813)

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

I. INTRODUCTION

The mathematical explanation of the ubiquitous cooperation is one of the greatest challenges of our days. Evolutionary game theory [1–6] offers a suitable framework to study this conundrum. This interdisciplinary research tool incorporates the assets of mathematics, biology, physics, economy, sociology, and other disciplines and creates an ideal environment to investigate the evolutionary processes taking place in populations.

In this paper, we examine the evolution of cooperation using a prisoner's dilemma (PD) [6–8] model characterizing the sharpest social dilemma situation that can emerge in a scenario of conflicting interests. In a PD encounter, the two participating individuals have the option to choose between cooperation (C) and defection (D). Depending on their decisions, they earn different payoffs. Mutual cooperation (defection) results in the reward R (punishment P), a defector encountering a cooperator earns the temptation T , and the exploited cooperator gets the sucker's payoff S . Due to the $T > R > P > S$ payoff ranking that characterizes the PD, the choice to defect results in higher earning than the choice to cooperate; however, if both players think along this argumentation, they end up with the second-worst payoff P instead of the second-best R that could be achieved by mutual cooperation. This outlines the dilemma.

This reasoning shows that mutual defection is the Nash equilibrium in the one-shot PD. Cooperation needs additional mechanisms to emerge: One of these is to turn the interaction between the two players to a repeated game. In the iterated

PD (IPD) setup, players play more rounds of one-shot PD games with each other and their success is measured by the total (or the average) payoff they acquire during this longer interaction sequence. Along this change, a plethora of possible strategies become accessible that can promote cooperation as, during the longer interaction, players can utilize the information gathered in the previous rounds. In our model, players can adopt three types of strategies: in addition to the simple always-cooperating AllC and always-defecting AllD, the win-stay lose-shift (WSLS) strategy is available for them. WSLS evaluates the outcome of the last round and if the earned payoff was below a given threshold, then the player changes his or her action for the next round. Players with this strategy possess a friendliness parameter that characterizes how cooperative they are in the first round. WSLS proved to be very effective in promoting cooperation in many cases due to its resilience against errors [9,10]. It is already clarified that the effect of memory [11], the competition between the aspiration levels [12], and the stochastic effects in systems with structured populations [13,14] can improve the efficiency of the WSLS strategy in the maintenance of cooperation.

We have chosen these three strategies because the resulting 3×3 payoff matrix has peculiar properties. Recently, it was discovered [15] that all symmetric 3×3 games can be decomposed into the linear combination of a potential game matrix and a cyclic component. Potential games are convenient in the sense that they behave as many well-known physical systems: The random sequential application of the

logit rule [16–20] drives potential games to a state where the stationary strategy distribution is determined by the Boltzmann distribution. The remaining cyclic component is responsible for the probability currents between the microscopic states that drive the system out of equilibrium [21,22]. In our model, the friendliness parameter can be tuned so the payoff matrix describes a pure potential game. Changing the parameter introduces the cyclic component and thus the effect of cyclicity can be investigated in a controlled environment.

We study the model in a spatial setting [5,23–29] where the interaction network is defined by a square lattice. Players play IPD games with their four nearest neighbors and acquire their accumulated payoff from them. To have a deeper understanding of the emerging phenomena, we study two types of update mechanisms. Imitation [30,31] is an update method widely used in evolutionary game theory on the analogy of biological processes: Players compare their payoff with those around them and try to adopt more successful strategies played by their peers. Spatiality proved to be beneficial for cooperation when this rule is applied as it fosters the formation of cooperative clusters. Clustered cooperators can support each other and can protect the cluster from defector attacks. The other update rule is the so-called logit rule originating from economy and physics models. Here players are able to calculate their payoff, taking into account every available strategy assuming their coplayers do not change their strategies. They exponentially prefer to choose the strategy that bids the highest prospects for them. In the low-noise limit, the logit rule coincides to a best response dynamics [32,33]. This rule is an important link between evolutionary game theory and equilibrium statistical physics because, for multiagent potential games, the random sequential logit rule drives the system into a Boltzmann distribution [16].

In our analysis we found that the analytical phase diagram agrees very well with the results provided by the logit rule. For most parameter values, it is accurate even in the case when cyclicity is present in the system. Further examination of the results reveals that they coincide at the interfaces where only two types of strategies are present. This surprising result can be explained by the fact that all 2×2 games are potential games, thus the potential analysis can result in a good prediction at these places. In the case of imitation dynamics, unconditional cooperators were eliminated due to the joint effect of spatiality and the presence of a more competitive cooperative strategy. These collective effects suppress the differences in the payoff matrix and make the phase diagram independent of the cyclic component's intensity. Comparing the two strategy update mechanisms, we can point out that imitation fosters the spread of cooperative behavior in a wider parameter range due to the strong cluster formation mechanism.

The paper is structured in the following way: In the next section we give a description about the model and the technical details of the simulations. In Sec. III, we study analytically the case when the game is a pure potential game and calculate the related phase diagram. In Secs. IV and V, we specify the simulation results for the logit and the imitation update mechanisms, respectively, and compare them with the analytical predictions along with a detailed discussion of the results. Finally, in Sec. VI we summarize our findings.

II. MODEL

We consider a spatial evolutionary model where players play IPD games with their neighbors. In a one-shot PD encounter, the players can choose between cooperation (C) and defection (D) and earn different payoffs depending on their simultaneous decisions. Mutual cooperation (defection) results in the reward R (punishment P), a successful defector exploiting a cooperator gets the temptation T while the victim of the exploitation receives the sucker's payoff S . The payoffs in the PD satisfy the $T > R > P > S$ ranking. Without loss of generality, we can fix two values of the payoff matrix thus we set the reward to 1 ($R = 1$) and the punishment to 0 ($P = 0$), leaving us with two payoff parameters. In an IPD, players play PD games repeatedly with each other, gathering a long-term payoff. Their success is best measured by the averaged payoff they acquire during this longer interaction sequence. Players can adopt three types of strategies: AllC, AllD, or win-stay-lose-shift with different friendliness parameter w . AllC (AllD) players always cooperate (defect), while WSLS players choose an action depending on the outcome of the last round of the IPD. They have a payoff expectation level; if their payoff in the last round was below this level, then they choose a different action in the next round (they switch from C to D or from D to C). If their payoff is above the expectation level, they stick to the action they used in the previous round. In the case of the PD, this level is set somewhere between $R(=1)$ and $P(=0)$; for easier reference, we can set it to 0.5. Translating this threshold value to actual actions means that a WSLS player will stick to his or her action after a successful exploitation or after mutual cooperation and he or she will change his or her action after mutual defection (from D to C) or after being cheated (from C to D). This rule determines the long-term behavior of WSLS; however, the first round action is not decided. The friendliness parameter w ($0 \leq w \leq 1$) defines the probability of cooperation in the first round; consequently, high w can be associated with a friendly behavior. To facilitate the analysis (by getting rid of an additional model parameter) we study the limit of infinitely iterated PD in this paper.

Players are located on the nodes of a square lattice of the size $N = L \times L$. Using the matrix formalism, the strategy of the player in the spatial position x , \mathbf{s}_x takes the form of one of the three-dimensional unit vectors:

$$\mathbf{AllD} \equiv \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{AllC} \equiv \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \text{or} \quad \mathbf{WSLS} \equiv \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \quad (1)$$

In this framework, the accumulated payoff gained from the neighbors on the square lattice can be calculated as

$$u_x(\mathbf{s}_x, \mathbf{s}_{\Omega_x}) = \sum_{y \in \Omega_x} \mathbf{s}_x^T \cdot \mathbf{M} \mathbf{s}_y, \quad (2)$$

where \mathbf{s}_x^T is the transpose of the player's strategy vector and \mathbf{s}_{Ω_x} denotes the strategy profile of players in the interaction neighborhood Ω_x of player x . In the case of the square lattice, the summation runs over the four nearest neighbors of player x . The matrix \mathbf{M} is of the following form, taking into account the

long-term payoffs between strategy pairs in the IPD encounters

$$\mathbf{M} = \begin{bmatrix} 0 & T & T/2 \\ S & 1 & w + (1-w)S \\ S/2 & w + (1-w)T & 1 \end{bmatrix}. \quad (3)$$

It can be seen that in the limit of the infinitely repeated PD, the result of the first round does not always influence the long-term average payoff of the opposing strategies; w does not show up in every WLSL-related entry of the payoff matrix.

After the games with their neighbors, players have the option to revise their strategy and adopt a more successful one in accord with the evolutionary aspect of the model. As we already mentioned in the Introduction, we study two types of strategy update methods.

In the case of the pairwise imitation rule, player x compares his or her payoff with one of his or her randomly chosen neighbor's (y) payoff and adopts the neighbor's strategy with a probability $W(s_x \leftarrow s_y)$ depending on the payoff difference as

$$W(s_x \leftarrow s_y) = \frac{1}{1 + \exp[(u_x - u_y)/K]}, \quad (4)$$

where K is associated with errors in the decision making that can originate from various sources (emotions, free will, fluctuation in the payoffs, external effects, noise) and, as a consequence, less-successful strategies can also be imitated with a low probability [5,34]. In the low- K limit, however, the more successful strategies are almost always imitated. The inverse of this parameter is also known as the intensity of selection [35].

The other studied strategy update method is the logit rule. In this case, players check the payoff they could gather by adopting any of the available strategies, provided their neighbors stick to their present strategies. Players adopt a strategy (or keep their strategies) depending on these expected payoff values:

$$\Pr(s'_x) = \frac{\exp[u_x(s'_x, s_{\Omega_x})/K]}{\sum_{s''_x \in \Gamma} \exp[u_x(s''_x, s_{\Omega_x})/K]}. \quad (5)$$

$\Pr(s'_x)$ defines the probability for player x to select strategy $s'_x \in \{\mathbf{AIIID}, \mathbf{AIIIC}, \mathbf{WSLS}\}$, where $u_x(s'_x, s_{\Omega_x})$ is her expected payoff when playing strategy s'_x while her neighbors keep playing their present strategies. Γ denotes the full strategy set in the model, i.e., $\Gamma = \{\mathbf{AIIID}, \mathbf{AIIIC}, \mathbf{WSLS}\}$. K indicates the measure of noise here as well that makes possible the appearance of less successful strategies. The probability of choosing a strategy with smaller expected payoff grows with increasing K .

During the Monte Carlo (MC) simulations, we study the model on a square lattice with periodic boundary conditions. The system size—when not stated otherwise—is $N = 200 \times 200 = 40\,000$. In most cases, simulations are started with random initial conditions where, initially, players can adopt one of the three available strategies with equal probability. However, in some cases, we used prepared initial states to make the cluster-growing process faster and to avoid the unwanted finite-size effects: we started the system from a state where each 10×10 block could be occupied either homogeneously by one of the strategies or by the AIIIC-AIIID chessboard

structure with equal probability. For the phase diagrams, stationary strategy concentrations were obtained by averaging the concentrations after a transient time of $t_{tr} = 50\,000$ Monte Carlo steps (MCS) over a $t_s = 10\,000$ MCS sampling time. The averaging parameters are adjusted to the system behavior. In one MCS, each player has the option to modify his or her strategy once on average in a random sequential order.

III. ANALYTICAL ANALYSIS FOR THE POTENTIAL GAME CASE ($w = 0.5$)

In a recently published paper [15], it was shown that all symmetric two-player three-strategy games can be decomposed into the linear combination of nine elementary 3×3 game matrices $[\mathbf{M} = \sum_k \beta(k)\mathbf{g}(k), k = 1, \dots, 9]$ where the matrices $\mathbf{g}(k)$ are considered as orthonormal basis vectors in the corresponding nine-dimensional vector space, i.e., $\sum_{ij} g_{ij}(k)g_{ij}(l) = \delta_{kl}$, with δ_{kl} being Kronecker's δ . For the decomposition, the coefficients $\beta(k)$ can be calculated as

$$\beta(k) = \sum_{i,j=1}^3 g_{ij}(k)M_{ij}. \quad (6)$$

The main advantage of the suggested two-dimensional Fourier decomposition is that different games can be analyzed in a general framework using the basis matrices. The first three basis matrices form an important group for the analysis. They contain uniform columns; for such payoff matrices, the focal player's income depends solely on the opponent's strategy and as such we can call these matrices *cross-dependent* components. The sum of these three terms obey the following form:

$$\mathbf{M}^{(\text{cross})} = \sum_{k=1}^3 \beta(k)\mathbf{g}(k) = \begin{pmatrix} c_1 & c_2 & c_3 \\ c_1 & c_2 & c_3 \\ c_1 & c_2 & c_3 \end{pmatrix}, \quad (7)$$

where $c_j = \frac{1}{3} \sum_{i=1}^3 M_{ij}$ ($j = 1, 2, \text{ and } 3$). These payoffs do not motivate a selfish individual to choose another strategy; moreover, if the logit rule is applied, then these payoffs incline the players to adopt a random strategy.

On the contrary, for the second important group of matrices, the so-called *self-dependent* payoff components, the sum of three basis games ($k = 1, 4, \text{ and } 5$ in the notation of paper [15]) can be described by a matrix composed of uniform rows as

$$\mathbf{M}^{(\text{self})} = \begin{pmatrix} s_1 & s_1 & s_1 \\ s_2 & s_2 & s_2 \\ s_3 & s_3 & s_3 \end{pmatrix}, \quad (8)$$

where $s_i = \frac{1}{3} \sum_{j=1}^3 M_{ij}$ ($i = 1, 2, \text{ and } 3$). For the self-dependent payoff the income of a selfish player depends only her own strategy, that is, direct interaction between the two players is missing here as well. In this case, the logit rule dictates the same choice for each player. As a result, in the zero-noise limit ($K \rightarrow 0$) players choose the strategy providing the maximum income $[\max(s_1, s_2, s_3)]$. Notice that

the first component $\mathbf{g}(1)$ is involved in both the cross- and self-dependent payoffs as $g_{ij}(1) = 1/3$.

The direct player-player interaction is described by the last four terms. The ninth component [the so-called rock-scissors-paper (RSP) component] is particularly important as it represents the cyclic dominance among the three strategies and prevents the existence of the potential. It is found that the cyclic component is missing [$\beta(9) = 0$] when

$$M_{12} + M_{23} + M_{31} - M_{21} - M_{32} - M_{13} = 0. \quad (9)$$

In the absence of the RSP component the game is a potential game and a symmetric potential matrix \mathbf{V} ($V_{ij} = V_{ji}$ for $i, j = 1, 2, 3$) can be derived from the payoff matrix \mathbf{M} that satisfies the following conditions:

$$M_{ij} - M_{i'j} = V_{ij} - V_{i'j} \quad (10)$$

for each possible value of $i, i', j = 1, 2, 3$. Evidently, similar conditions must be satisfied from the point of view of the second player that ensures the mentioned symmetry of the potential matrix. In fact, the elements of the potential matrix \mathbf{V} summarize the individual incentives for the active player to modify her strategy unilaterally [36]. In a spatial multiagent evolutionary game that is built from uniform pair interactions described by two-player potential games, the potential for the whole system can be given as

$$U(\mathbf{S}) = \sum_{\langle x, y \rangle} \mathbf{s}_x^T \cdot \mathbf{V} \mathbf{s}_y, \quad (11)$$

where the summation runs over the nearest neighbors on the (square) lattice and \mathbf{S} denotes the strategy profile of the system. The role of the scalar function $U(\mathbf{S})$ is similar to the negative Hamiltonian of the Ising- or Potts-type models in physics. Blume [16,17] has shown that if the evolution is controlled by random sequential logit rule, then the system evolves into the Boltzmann distribution where the strategy profile \mathbf{S} is present with a probability $p(\mathbf{S}) = e^{U(\mathbf{S})/K} / Z$ with Z as a normalization factor. In short, for the quantitative analysis of evolutionary potential games with logit rule we can apply the tools and methods of equilibrium statistical physics [37,38].

For $w = 1/2$, condition (9) is satisfied in our model, thus the potential exists and a possible \mathbf{V} can be given for this particular case as

$$\mathbf{V} = \begin{pmatrix} 0 & S & S/2 \\ S & 1 + S - T & 1/2 + S - T/2 \\ S/2 & 1/2 + S - T/2 & 1 + S/2 - T/2 \end{pmatrix}. \quad (12)$$

In the knowledge of the pair potential \mathbf{V} , one can evaluate its maximal component that coincides with the strategy pair preferred by players in the low-noise (K) limit. For example, in the present spatial model all players choose defection unconditionally (AllD strategy) if V_{11} is the largest component of \mathbf{V} . Thus we can determine the payoff region (on the T - S plane) where this condition is satisfied. Straightforward calculation gives that unconditional defection occurs in the system if the payoff parameters simultaneously satisfy the following two conditions: $S < 0$ and $T > S + 2$. It is worth mentioning that the direct comparison of the V_{ij} elements gives two weaker conditions, namely $T > 2S + 1$ and $T > S + 1$, whereas the condition $S < 0$ appears twice. A similar

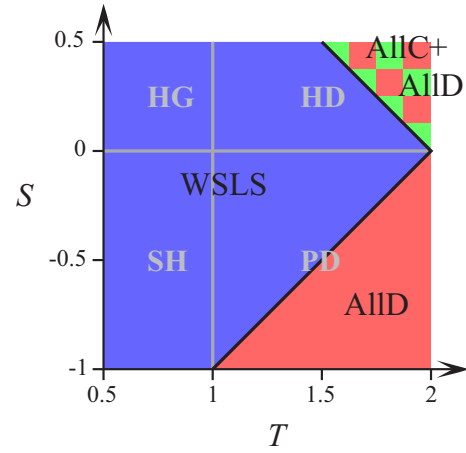


FIG. 1. (Color online) Phase diagram determined by the maximal value of the potential matrix (that exists for $w = 1/2$). Black labels in the regions refer to the preferred strategy. In the AllC+AllD phase, AllC and AllD players occupy the square lattice according to a chessboard or anti-chessboard-like arrangement. Gray labels and lines show the different social dilemma parameter regions, the prisoner's dilemma (PD), the hawk-dove game (HD), the stag hunt game (SH), and the harmony game (HG).

calculation can be performed for all other strategy profiles. The result of this analysis is summarized in the phase diagram of Fig. 1. To have a clearer picture and a deeper understanding of the model, we have plotted the results beyond the PD parameter limits, in the $0.5 < T < 2$, $-1 < S < 0.5$ domain. In this region the WSLs strategy behaves exactly the same way as in the PD area.

If a nondiagonal element of the potential matrix (for example, V_{12} or, equivalently, $V_{21} = V_{12}$) is maximal, then a sublattice ordered arrangement of the two strategies (in this case AllC and AllD) is predicted to rule the system in a chessboard or anti-chessboard-like structure. According to the quantitative analysis, two equivalent (chessboard and antichessboard) arrangements of AllC and AllD strategies exist if $S > 0$, $T > 1$, and $T + S > 2$ simultaneously. In this area, the matrix game has two symmetric strict, pure Nash equilibria characteristic to the two-strategy anticoordination games. When increasing the noise level K this system exhibits an order-disorder phase transition resembling those described by the antiferromagnetic Ising models [39,40]. More precisely, the suppressed third strategy (WSLs) is present with a low frequency, whereas a critical phase transition (sublattice ordering) can be observed in the K dependence of the AllC and AllD strategy frequencies. At low noise values, one of the sublattices is occupied by AllD players while AllC players reside in the other sublattice; above a critical noise level ($K > K_c$), AllC and AllD strategies are present in both sublattices with the same concentration.

The phase boundary separating the homogeneous WSLs and AllD states in the low-noise limit is defined by the equation $V_{11} = V_{33}$ ($S = T - 2$ and $S < 0$). The competition between the mentioned phases, however, depends on how V_{11} and V_{33} are composed of the contributions of the self-dependent payoffs and coordination-type interactions. These latter contributions can be constructed as the linear combination of

the matrices belonging to the third important basis matrix group [$\mathbf{g}(6)$, $\mathbf{g}(7)$, and $\mathbf{g}(8)$] defined in Ref. [15]. Within this three-dimensional subspace of games we can distinguish three other basis games, namely

$$\mathbf{g}'(6) = \frac{1}{2} \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (13)$$

$$\mathbf{g}'(7) = \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix}, \quad (14)$$

$$\mathbf{g}'(8) = \frac{1}{2} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix} \quad (15)$$

which represent the coordination- (or Ising) type interactions between two of three strategies. The latter games are not orthogonal to each other; their sum is equivalent to one of the Fourier components corresponding to the three-state Potts model. The games defined by Eqs. (13)–(15) span the three-dimensional subspace of coordination-type interactions and their strength in a game \mathbf{M} can be characterized by the coefficients:

$$\beta'(k) = \sum_{i,j=1}^3 g'_{ij}(k) M_{ij}, \quad (16)$$

defined on the analogy of Eq. (6) for $k = 6, 7$, and 8 . Straightforward calculation justifies that at the $T = 1$, $S = -1$ point, the three coefficients are equal [$\beta'(6) = \beta'(7) = \beta'(8)$] and the self-dependent components support the strategies AllD and WSLs equally ($s_1 = s_3 > s_2$). In other words, at this particular point on the T - S plane, the game is equivalent to a three-state Potts model in which an external field favors the first and third states identically [41]. The resultant K dependence in the strategy frequencies and the Ising-type critical phase transition is illustrated numerically in the following section. Above this point along the WSLs-AllD phase boundary, the stronger attractive strength of the WSLs-WSLs interaction [that can be quantified as $\beta'(7) + \beta'(8)$] is accompanied by a lower self-dependent component (s_3). On the contrary, for $T < 1$ along the phase boundary the weaker attractive WSLs-WSLs interaction is compensated by a stronger self-dependent contribution to achieve $V_{11} = V_{33}$. Due to the mentioned differences along the phase boundary, the equivalence between the AllD and WSLs phases is violated and the resultant symmetry breaking prohibits the critical phase transition when the noise level K is increased. This is demonstrated in Sec. IV as well.

Until now we have studied the case $w = 1/2$ when the existence of the potential could be exploited in the analytical treatment. For the other cases ($w \neq 1/2$), the RSP component appears in the payoff matrix with a strength of $\beta(9) = (w - 1/2)(T - S)/\sqrt{6}$. The contribution of the RSP game drives the system out of the Boltzmann distribution and breaks the detailed balance at finite noise levels ($K > 0$) even if the dynamics is controlled by the logit rule (5). On the other hand, in the low-noise limit, the interactions of the potential games

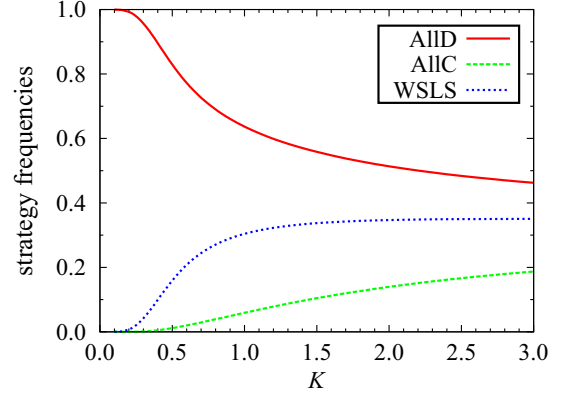


FIG. 2. (Color online) Strategy frequencies versus K with logit rule at $T = 2$ and $S = -0.5$ for $w = 1/2$. The accuracy of MC results is comparable to the line thickness.

can result in a spatially ordered strategy arrangement that blocks the effects induced by the RSP component if $\beta(9)$ does not exceed a threshold value. The existence of such situations is demonstrated numerically in the following section.

IV. SIMULATION RESULTS FOR THE LOGIT RULE

As a start, for the numerical analysis of the model with logit update rule, we study the noise dependence of the strategy frequencies for $w = 1/2$ in the parameter region where the theoretical analysis predicted a homogeneous AllD state in the low-noise limit. Figure 2 illustrates the MC results exhibiting a smooth and monotonous decrease in the frequency of AllD from 1 towards $1/3$ when K is increased. Simultaneously, the AllC and WSLs strategies appear as rare point defects at low noise levels and their frequencies go monotonously to $1/3$ in the large K limit when the players choose their strategy at random. Similar behavior can be observed when WSLs dominates the system at low noises. Naturally, in this case, the role of AllD and WSLs are exchanged.

Contrary to the above typical behavior, the MC analysis shows a critical transition in the strategy frequencies when varying the noise level at $S = -1$ and $T = 1$. The previous theoretical analysis has predicted the equivalence of our model to the three-state Potts model at this point with an external field supporting the AllD and WSLs states equally. We illustrate the critical transition by performing two series of simulations. The first series are started at $K = 0.2$ from a state where all players use the AllD strategy. After some thermalization time t_t , the average strategy frequencies are determined by averaging over a sampling time t_s . Afterward, K is increased and the simulation is continued from the last strategy distribution. This procedure is repeated many times; the results are indicated by lines in Fig. 3. We emphasize that the values of t_t , t_s , and the system size were increased gradually when we approached the critical point with smaller and smaller steps in K . Using this method, we could reduce the undesired effects of the divergency in fluctuations, relaxation time, and correlation length in agreement with the theoretical results of phase transitions [42]. In the close vicinity of the critical point $t_t = t_s = 10^6$ MCS are used on a square lattice of 10^6 sites.

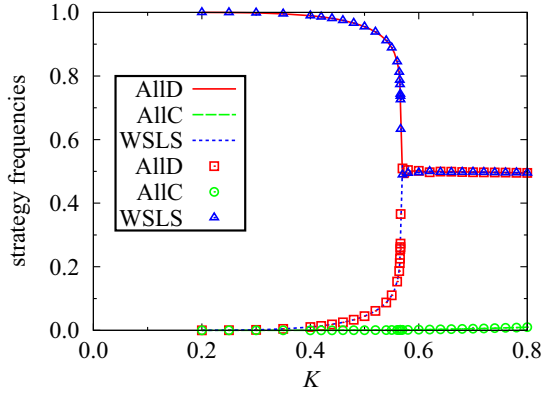


FIG. 3. (Color online) Coincidence between the strategy frequencies for two consecutive runs when K is increased gradually during the simulations at $S = -1$ and $T = 1$ for $w = 1/2$. In the first run, the system is started from homogeneous AllD state; the average strategy frequencies are indicated by lines. In the second run, all players use the WSLS strategy initially; strategy frequencies are illustrated by symbols.

Figure 3 compares the previous MC data with those we obtained when the simulation is started from a homogeneous WSLS phase. Notice the coincidence in the strategy frequencies for the two equivalent macroscopic states. Within the statistical error our MC data are consistent with the theoretical expectations, suggesting an Ising-type critical phase transition at $K_c = 0.56660(5)$.

The next plot on the strategy frequencies versus K is addressed to illustrate the typical behavior occurring along the WSLS-AllD phase boundary. In this case we have repeated the latter method for $T = 1.5$ and $S = -0.5$. The MC results are plotted in Fig. 4 in a similar manner. The most striking feature is the absence of a critical transition due to the above mentioned fact that here the AllD and WSLS phases are distinct. As a result, the WSLS phase becomes unstable when K is increased and this state is transformed into the noisy AllD phase for some noise level when a sufficiently large island of the preferred phase (AllD) appears via thermal fluctuations. A similar phenomenon is described by the Ising model when considering

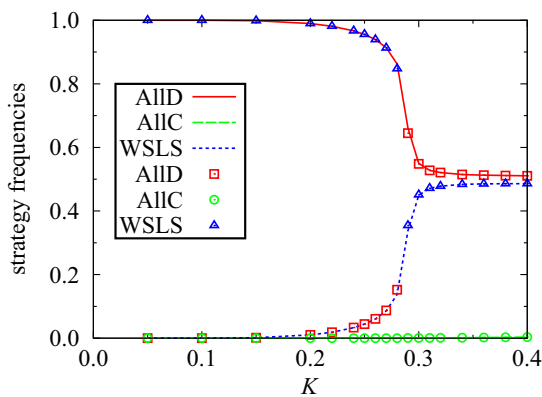


FIG. 4. (Color online) Strategy frequencies versus K for two consecutive series of runs at $S = -0.5$ and $T = 1.5$ for $w = 1/2$. The notations of strategy frequencies are the same as in Fig. 3.

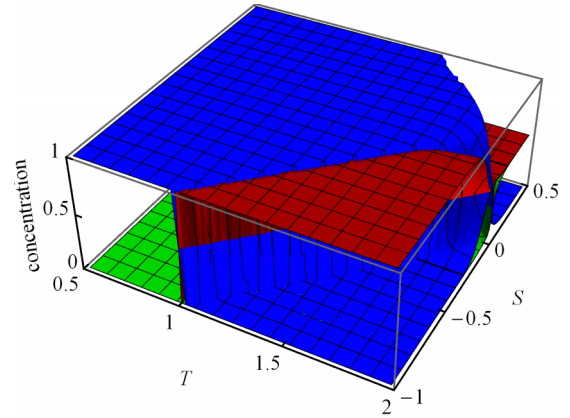


FIG. 5. (Color online) Strategy frequencies in the T - S plane when the logit strategy update rule is applied for $w = 1/2$ and $K = 0.1$. Blue (very dark gray) surface indicates the average frequency of WSLS players, red (dark gray) surface denotes the average frequency of the AllD strategy. AllC players [green (light gray)] appear only in the hawk-dove region, forming twofold degenerated sublattice ordered structures together with the AllD players.

the decay of the metastable state via nucleation and growth of the stable phase’s droplets [43]. A symmetric phenomenon can be observed along the WSLS-AllD phase boundary for $S < -1$ with the WSLS state becoming stable. We emphasize though that this is the low-noise behavior of the system, for high noise levels, the strategy adoption process becomes random and the strategies are present with equal frequencies.

We have determined the phase diagram in the T - S plane for the strategy frequencies at a low noise level with MC simulations as well; the results are illustrated in Fig. 5. Notice the agreement with the theoretical predictions in the case of a potential game ($w = 0.5$). In the prisoner’s dilemma parameter region, the system behavior is determined by the competition between the AllD and WSLS strategies; the dominated AllC strategy is present in the system with an extremely low frequency and practically does not disturb the competition of AllD and WSLS.

During the numerical evaluation of the phase diagram(s) in the low-noise limit we have faced technical difficulties related to the presence of metastable states and the relevant initial transient phenomena when the system was started from a random distribution of the three strategies. In agreement with the prediction of the mean-field analysis, for most of the social dilemmas the random initial state favors the spreading of the AllD strategy that may be so intensive that after a short time only isolated WSLS and AllC players remain in the system. In fact, the extinction of WSLS domains is supported by the early presence of AllC players who feed the defectors in their competition against the WSLS players. WSLS players can beat and invade the territories of AllD only if they survive by forming a sufficiently large compact domain in the spatial system. For small system sizes and at low noise levels this probability may be extremely low and the system can remain in the metastable AllD phase for a long time. For large sizes, however, the sufficiently large WSLS domain forms within a “short” time and afterwards the system evolves into the stable WSLS phase via a domain growing process. In the present

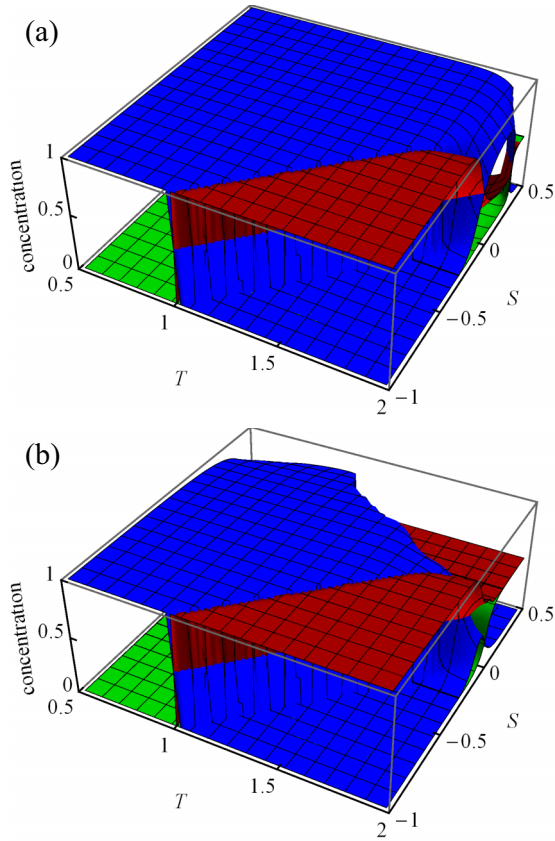


FIG. 6. (Color online) Comparison of strategy frequencies in the T - S plane when the logit strategy update rule is applied for (a) $w = 0.1$ and (b) $w = 0.9$ at $K = 0.1$. The color code for the strategies is the same as in Fig. 5.

work, our attention is focused on the determination of phase diagrams that are valid for the $N \rightarrow \infty$ limit in agreement with the results of statistical physics. In order to avoid the undesired finite-size effects, simulations are started from an artificial initial state instead of the random strategy distribution. This method proved to be successful for the evaluation of phase diagrams in other three-strategy evolutionary games as well [44]. In the present initial state, the lattice is divided into blocks of $l \times l$ sites in a way that each block is occupied by one of the possible ordered states selected at random.

At the end of this section we discuss the effect of the friendliness parameter when $w \neq 1/2$. For this goal we have evaluated and plotted the stable strategy frequencies in the T - S plane for $w = 0.1$ and 0.9 . Within the PD region, the MC data in Fig. 6 are very close to those we obtained for $w = 1/2$ (see Fig. 5). This behavior can be explained by recognizing that AllC remains a dominated strategy in this parameter region and the resultant two-strategy subgame (when the players are limited to choose either AllD or WSLS) is independent of w . Notice, furthermore, that this two-strategy subgame is equivalent to a PD game where the WSLS strategy represents the cooperative behavior with rescaled T and S payoffs. Evidently, w dependence in the stationary strategy frequencies appears at higher noise levels when AllC players appear in the system due to the fluctuations. At the same time, w dependence of the strategy frequencies becomes relevant in the low-noise

limit in the hawk-dove parameter region ($S > 0$ and $T > 1$) where AllC and AllD players form a sublattice ordered spatial structure. Readers can observe the relevant variations in the stationary states when comparing the plots of Fig. 5 and 6.

V. SIMULATION RESULTS FOR THE IMITATION RULE

In the general properties of the above-discussed model both the existence of the potential and the application of the logit rule have played a crucial role. Replacing the logit rule by pairwise imitation results in several very important consequences. As imitation is similar to the contact process [45–47], this model has three (homogeneous) absorbing states from which the system cannot escape. The existence of absorbing state(s) implies the appearance of critical phase transitions that is realized by the extinction of one or two strategies when tuning the parameters. These critical phase transitions exhibit general features and belong to the directed percolation universality class [34,48,49].

The visualization of the strategy distribution's evolution on the square lattice clearly illustrates that the AllC strategy dies out within a short transient period independently of the initial state. Subsequently, the model becomes a two-strategy (AllD and WSLS) PD game that was studied previously by many authors (for a survey see Ref. [5]), particularly in the case of $S = 0$. Figure 7 shows the very narrow coexistence region between the homogeneous WSLS and AllD domains. It is worth mentioning that the stationary state can be reached in the limit $N \rightarrow \infty$ more effectively by starting the system from an artificial initial state [44] as described above.

Until now, the systematic analysis of the effects of noise is limited to the weak PD games ($S = 0$) for several connectivity networks [49,50]. The present model behaves similarly. The coexistence regions are limited by two threshold values, that is, WSLS and AllD coexist if $T_{c1}(S, K) < T < T_{c2}(S, K)$. In the light of the latter results, one can expect similar behavior for other S and K parameters as well. To be more accurate, we expect that $[T_{c2}(S, K) - T_{c1}(S, K)] \rightarrow 0$ if K goes to ∞ .

At low noises the systematic analysis of the extinction of AllD strategies goes beyond the scope of the present work due to the extremely slow relaxation processes resembling a branching-unifying random walk of the isolated AllD players.

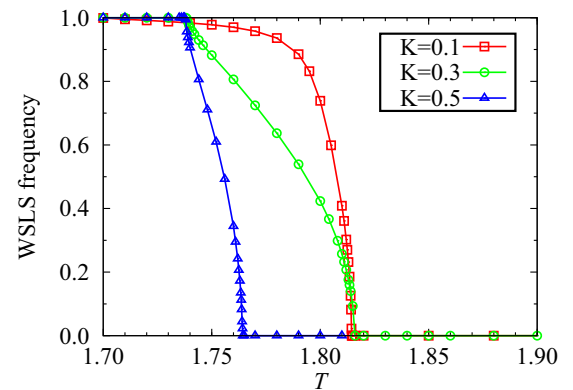


FIG. 7. (Color online) Frequency of the WSLS strategy as a function of T at $S = -0.5$ for different noise levels when imitation update rule is applied.

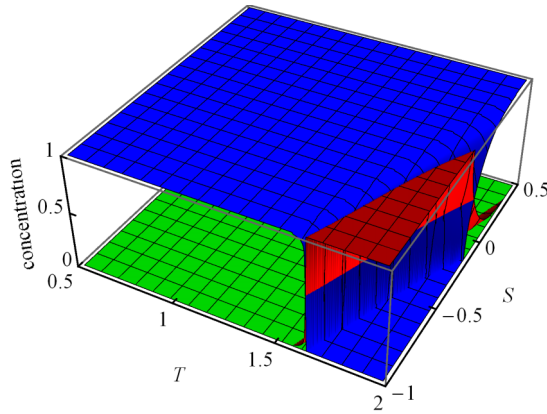


FIG. 8. (Color online) Stationary strategy frequencies in the T - S plane for any w parameter value when imitation rule is applied. The color code for the strategies is the same as in Fig. 5.

The MC data are illustrated in Fig. 8 when imitation dynamics is applied for the given region of the T - S plane for $K = 0.1$. As AllC players are out of the game, the parameter w no longer plays a role, thus we obtain the same phase diagram for all w values.

Figure 8 illustrates clearly that the homogeneous WSLs phase is expanded towards the hawk-dove region ($T > 1$ and $S > 0$) where the imitation of the nearest neighbor is not capable of building up the sublattice-ordered arrangements of AllD and AllC strategies that are preferred for the logit rule. In the absence of WSLs, a previous analysis [51] indicated a low level of cooperation depending on the actual values of T and S . It is more important, however, that the WSLs phase is extended towards the regions in the PD area where the logit rule predicted the dominance of AllD. This phenomenon is related to the cross-dependent components that do not affect the individual decisions in the case of the logit rule. On the contrary, for imitation, cooperators reward each other via the cross-dependent payoffs and the resultant extra profit helps them invade the AllD territories along the interfaces as it is demonstrated in the first spatial models [24,52]. On the other hand, the present stochastic rule makes the interfaces irregular that favors the spreading of defection. Despite the latter mechanism, the present simulations clearly indicate the positive impact of the imitation dynamics in the maintenance of cooperative (WSLS) behavior. In the light of previous results (surveyed in Ref. [5]), one can expect even more significant suppression of defection when other connectivity structures (e.g., square lattice with nearest and next-nearest neighbor interactions) are chosen.

VI. SUMMARY

We have studied a three-strategy spatial evolutionary prisoner's dilemma game with imitation and logit update rules where players could follow the always-cooperating, always-defecting or the win-stay-lose-shift strategy. Players were located on a square lattice and gained their payoff from games with their nearest neighbors. The payoff matrix of the game was built by introducing a friendliness parameter w of the WSLs strategy. It is found that for a suitable choice of the friendliness parameter ($w = 1/2$) the uniform 3×3 game becomes equivalent to a potential game. In this case, when the logit rule is applied, the behavior of the system in the low-noise limit is determined by the potential and can be evaluated analytically. Additionally, we have demonstrated the applicability of the concepts of decomposition [15], distinguishing four types of elementary games that explain the different noise dependence along the phase boundaries where the potential permits the existence of two ordered strategy arrangements.

The analytical results served as references when considering the effects of noise and w for the logit rule as well as when the evolution is controlled by stochastic pairwise imitation. During the systematic comparison of the effects we faced the negligibility of the cross-dependent payoffs for the logit rule, whereas these components played an important role for imitation by supporting the maintenance of cooperation.

Until now the systematic analysis of multiagent evolutionary games with many strategies is limited to models where the payoff matrix is defined by several parameters. The concept of matrix decomposition (particularly for potential games) presents us tools to analyze the general features of games. It can be useful even for imitation when several strategies die out and the system develops into a symmetric two-strategy subgame that is always a potential game [15]. In some sense, the systematic investigation of the effects of payoff components is similar to the study of the two-strategy subgames [53]. In the present model the latter approach has explained the absence of w dependence in the PD parameter region. These methods can serve as important precursors to understand and classify spatial games involving a large number of parameters.

ACKNOWLEDGMENTS

This work was supported by the John Templeton Foundation (FQEB Grant No. RFP-12-22), the Hungarian Scientific Research Fund (OTKA TK-101490 and OTKA K-112929), and the Lendület Program of the Hungarian Academy of Sciences.

- [1] J. Maynard Smith, *Evolution and the Theory of Games* (Cambridge University Press, Cambridge, UK, 1982).
- [2] R. Axelrod, *The Evolution of Cooperation* (Basic Books, New York, 1984).
- [3] J. Hofbauer and K. Sigmund, *Evolutionary Games and Population Dynamics* (Cambridge University Press, Cambridge, UK, 1998).

- [4] H. Gintis, *Game Theory Evolving* (Princeton University Press, Princeton, NJ, 2000).
- [5] G. Szabó and G. Fáth, *Phys. Rep.* **446**, 97 (2007).
- [6] K. Sigmund, *The Calculus of Selfishness* (Princeton University Press, Princeton, NJ, 2010).
- [7] M. Dresher, *Games of Strategy: Theory and Applications* (Prentice Hall, Englewood Cliffs, NJ, 1961).

- [8] A. Rapoport and A. M. Chammah, *Prisoners Dilemma: A Study in Conflict and Cooperation* (University of Michigan Press, Ann Arbor, MI, 1965).
- [9] M. A. Nowak and K. Sigmund, *Nature* **364**, 56 (1993).
- [10] L. A. Imhof, D. Fudenberg, and M. A. Nowak, *J. Theor. Biol.* **247**, 574 (2007).
- [11] M. Posch, *J. Theor. Biol.* **198**, 183 (1999).
- [12] M. Posch, A. Pichler, and K. Sigmund, *Proc. R. Soc. Lond. B* **266**, 1427 (1999).
- [13] X.-J. Chen, F. Fu, and L. Wang, *Physica A* **387**, 5609 (2008).
- [14] X.-J. Chen and L. Wang, *Phys. Rev. E* **77**, 017103 (2008).
- [15] G. Szabó, K. S. Bodó, B. Allen, and M. A. Nowak, *Phys. Rev. E* **90**, 042811 (2014).
- [16] L. E. Blume, *Games Econ. Behav.* **5**, 387 (1993).
- [17] L. E. Blume, *Games Econ. Behav.* **11**, 111 (1995).
- [18] D. Fudenberg and D. Levine, *The Theory of Learning in Games* (MIT Press, Cambridge, MA, 1998).
- [19] J. Hofbauer and W. H. Sandholm, *J. Econ. Theor.* **132**, 47 (2007).
- [20] W. H. Sandholm, *Population Games and Evolutionary Dynamics* (MIT Press, Cambridge, MA, 2010).
- [21] J. Schnakenberg, *Rev. Mod. Phys.* **48**, 571 (1976).
- [22] B. Andrae, J. Cremer, T. Reichenbach, and E. Frey, *Phys. Rev. Lett.* **104**, 218102 (2010).
- [23] M. A. Nowak and R. M. May, *Nature* **359**, 826 (1992).
- [24] M. A. Nowak and R. M. May, *Int. J. Bifurcat. Chaos* **3**, 35 (1993).
- [25] H. Ohtsuki, C. Hauert, E. Lieberman, and M. A. Nowak, *Nature* **441**, 502 (2006).
- [26] H. Ohtsuki and M. A. Nowak, *J. Theor. Biol.* **247**, 462 (2007).
- [27] C. E. Tarnita, H. Ohtsuki, T. Antal, F. Fu, and M. A. Nowak, *J. Theor. Biol.* **259**, 570 (2009).
- [28] M. van Veelen, J. Garca, D. G. Rand, and M. A. Nowak, *Proc. Natl. Acad. Sci. USA* **109**, 9929 (2012).
- [29] B. Allen and M. A. Nowak, *EMS Surv. Math. Sci.* **1**, 113 (2014).
- [30] J. Maynard Smith and G. R. Price, *Nature* **246**, 15 (1973).
- [31] R. Axelrod and W. D. Hamilton, *Science* **211**, 1390 (1981).
- [32] D. Fudenberg and J. Tirole, *Game Theory* (MIT Press, Cambridge, MA, 1991).
- [33] R. Gibbons, *Game Theory for Applied Economists* (Princeton University Press, Princeton, NJ, 1992).
- [34] G. Szabó and C. Tóke, *Phys. Rev. E* **58**, 69 (1998).
- [35] A. Traulsen, M. A. Nowak, and J. M. Pacheco, *Phys. Rev. E* **74**, 011909 (2006).
- [36] D. Monderer and L. S. Shapley, *Games Econ. Behav.* **14**, 124 (1996).
- [37] L. D. Landau and E. M. Lifshitz, *Statistical Physics, Part 1*, 3rd ed. (Butterworth-Heinemann, Oxford, 1980).
- [38] K. Toda, R. Kubo, and N. Saito, *Statistical Physics I: Equilibrium Statistical Mechanics* (Springer, Berlin, 1991).
- [39] C. Domb, in *Phase Transitions and Critical Phenomena*, Vol. 3, edited by C. Domb and M. S. Green (Academic Press, London, 1974), pp. 357–484.
- [40] L. Onsager, *Phys. Rev.* **65**, 117 (1944).
- [41] F. Y. Wu, *Rev. Mod. Phys.* **54**, 235 (1982).
- [42] H. E. Stanley, *Introduction to Phase Transitions and Critical Phenomena* (Clarendon Press, Oxford, 1971).
- [43] K. Binder, *Rep. Prog. Phys.* **50**, 783 (1987).
- [44] A. Szolnoki, G. Szabó, and M. Perc, *Phys. Rev. E* **83**, 036101 (2011).
- [45] T. M. Liggett, *Interacting Particle Systems* (Springer, New York, 1985).
- [46] J. Marro and R. Dickman, *Nonequilibrium Phase Transitions in Lattice Models* (Cambridge University Press, Cambridge, UK, 1999).
- [47] H. Hinrichsen, *Adv. Phys.* **49**, 815 (2000).
- [48] J. R. N. Chiappin and M. J. de Oliveira, *Phys. Rev. E* **59**, 6419 (1999).
- [49] G. Szabó, J. Vukov, and A. Szolnoki, *Phys. Rev. E* **72**, 047107 (2005).
- [50] J. Vukov, G. Szabó, and A. Szolnoki, *Phys. Rev. E* **73**, 067103 (2006).
- [51] C. Hauert and M. Doebeli, *Nature* **428**, 643 (2004).
- [52] M. A. Nowak, S. Bonhoeffer, and R. M. May, *Int. J. Bifurcat. Chaos* **4**, 33 (1994).
- [53] P. Cui and Z. Wu, *J. Theor. Biol.* **361**, 111 (2014).