Win-Stay-Lose-Learn Promotes Cooperation in the Spatial Prisoner's Dilemma Game

Yongkui Liu[1,2,3]*, Xiaojie Chen[4]*, Lin Zhang[1], Long Wang[5], Matjaž Perc[6]*

1 School of Automation Science and Electrical Engineering, Beihang University, Beijing, China, 2 School of Electronic and Control Engineering, Chang'an University, Xi'an, China, 3 Center for Road Traffic Intelligent Detection and Equipment Engineering, Chang'an University, Xi'an, China, 4 Evolution and Ecology Program, International Institute for Applied Systems Analysis, Laxenburg, Austria, 5 State Key Laboratory for Turbulence and Complex Systems, College of Engineering, Peking University, Beijing, China, 6 Faculty of Natural Sciences and Mathematics, Department of Physics, University of Maribor, Maribor, Slovenia

Abstract

Holding on to one's strategy is natural and common if the later warrants success and satisfaction. This goes against widespread simulation practices of evolutionary games, where players frequently consider changing their strategy even though their payoffs may be marginally different than those of the other players. Inspired by this observation, we introduce an aspiration-based win-stay-lose-learn strategy updating rule into the spatial prisoner's dilemma game. The rule is simple and intuitive, foreseeing strategy changes only by dissatisfied players, who then attempt to adopt the strategy of one of their nearest neighbors, while the strategies of satisfied players are not subject to change. We find that the proposed win-stay-lose-learn rule promotes the evolution of cooperation, and it does so very robustly and independently of the initial conditions. In fact, we show that even a minute initial fraction of cooperators may be sufficient to eventually secure a highly cooperative final state. In addition to extensive simulation results that support our conclusions, we also present results obtained by means of the pair approximation of the studied game. Our findings continue the success story of related win-stay strategy updating rules, and by doing so reveal new ways of resolving the prisoner's dilemma.

Introduction

Evolutionary game theory provides a powerful mathematical framework for studying the emergence and stability of cooperation in social, economic and biological systems [1–5]. The prisoner's dilemma game, in particular, is frequently considered as a paradigm for studying the emergence of cooperation among selfish and unrelated individuals [6]. The outcome of the prisoner's dilemma game is governed by pairwise interactions, such that at any instance of the game two individuals, who can either cooperate or defect, play the game against each other by selecting their strategy simultaneously and without knowing what the other player has chosen. Both players receive the reward R upon mutual cooperation, but the punishment P upon mutual defection. If one player defects while the other cooperates, however, the cooperator receives the sucker's payoff S while the defector receives the temptation $T = b$. Since $T > R > P > S$, there is an innate tension between individual interests (the rational strategy, yielding an optimal outcome for the player regardless of what the other player chooses, is defection) and social welfare (for the society as a whole the optimal strategy is cooperation) that may result in the "tragedy of the commons" [7]. Five prominent rules for the successful evolution of cooperation, which may help avert an impeding social decline, are kin selection, direct and indirect reciprocity, network reciprocity as well as group selection, as comprehensively reviewed in [8].

Since the pioneering work of Nowak and May [9] spatial games have received ample attention, and they have become inspirational for generations of researchers trying to reveal new ways by means of which cooperation can prevail over defection [10–12]. In the context of spatial games, network topology and hierarchies have been identified as a crucial determinant for the success of cooperative behavior [13–28], where in particular the scale-free topology has proven very beneficial for the evolution of cooperation. In fact, payoff normalization [29–31] and conformity [32] belong to the select and very small class of mechanisms that can upset the success of cooperators on such highly heterogeneous networks. Other approaches facilitating the evolution of cooperation include the introduction of noise to payoffs and updating rules [33–38], asymmetry between interaction and replacement graphs [39,40], diversity [41–44], differences between time scales of game dynamics [44–47], as well as adoption of simultaneous different strategies against different opponents [48]. Somewhat more personally-inspired features supporting the evolution of cooperation involve memory effects [49], heterogeneous teaching activity [50–52], preferential learning [53,54], mobility [55–59], myopically selective interactions [60], and coevolutionary partner choice [61–63], to name but a few examples studied in recent years.

Regardless of the details of mechanisms that may promote the evolution of cooperation in the spatial prisoner's dilemma game, most frequently, it was assumed that individual players learn from

their neighbors to update their strategies, and that they do so more or less in every round of game. But in reality, we are much less prone to changing our strategy (see [64] and references therein). Withstanding ample trial and error, only when we feel to a sufficiently high degree unsuccessful and dissatisfied may we be tempted into altering it. Enter aspirations, which play a pivotal role in determining our satisfaction and the notion of personal success. Indeed, the subtle role of aspirations in evolutionary games has recently received a lot of attention [65–73], and amongst others, it was discovered that too high aspirations may act detrimental on the evolution of cooperation. This result is quite intuitive, as very high aspirations will inevitably lead players to choose defection over cooperation in order to achieve their high goals. Regardless of the level, however, aspiration provides an elegant means to define when a player is prone to changing its strategy and when not. In particular, if the performance trails behind the aspiration, then the player will likely attempt to change its strategy. On the other hand, if the performance agrees or is even better than the aspiration, then the player will not attempt to change its strategy. Here we make use of this simple and intuitive rationale to build on the seminal works that introduced and studied the win-stay-like (win-stay-lose-shift being the most prominent example) strategy updating rules [74–76].

In this paper, we thus introduce a so-called win-stay-lose-learn strategy updating rule as follows. When satisfied, players maintain their strategies and do not attempt to change them. When dissatisfied, however, players proceed as it is traditionally assumed, i.e., by attempting to imitate the strategy of one of their neighbors [77]. It is worth pointing out that since in evolutionary games on structured populations individuals need to interact with their neighbors for collecting payoffs, it can be assumed that such an interaction mode provides enough opportunities for players to observe the information of their neighbors (including payoffs). Nevertheless, it is difficult to pinpoint whether our model can be held accountable only for human behavior or also for animal behavior. Certainly, some level of intelligence is needed by the players to accommodate our assumptions. In the proposed model the strategy updating is thus conditional on whether the players are satisfied or not, which we determine by means of the aspiration level $A > 0$ that is considered as a free parameter. We note again that the majority of previous studies assumed that players will always try and adopt the strategy of one of their neighbors, even if the neighbor is performing worse, and regardless of the individual

level of satisfaction. Here we depart from this somewhat simplifying assumption, and by doing so discover that the aspiration-based conditional strategy updating, termed win-stay-lose-learn, strongly promotes the evolution of cooperation in the spatial prisoner's dilemma game, even under very unfavorable initial conditions [78,79] and by high temptations to defect b. All the details of the proposed win-stay-lose-learn strategy updating rule and the setup of the spatial prisoner's dilemma game are described in the Methods section, while here we proceed with presenting the main results.

Results

We start by presenting the results as obtained when the cooperators and defectors are distributed uniformly at random, each thus initially occupying half of the square lattice. As the main parameters, we consider the aspiration level A and the temptation to defect b. Figure 1 shows the fraction of cooperators ρ_C as a function of the temptation to defect b for different aspiration levels A. We find that the aspiration level has a significant influence on the density of cooperators. In particular, for small values of A, intermediate levels of cooperation are maintained, and the temptation to defect has no effect on cooperation, e.g., for $A = 0$ and $A = 0.2$, $\rho_C = 0.5$ and $\rho_C = 0.47$, respectively, irrespective of the value of b. When A is within an intermediate range, the density of cooperators increases with increasing A, however, the maximal value of b still warranting high cooperation levels becomes smaller, e.g., when $A = 0.4$, $\rho_C = 0.7$ for $b \in [0,1.6]$ and when $A = 0.6$, $\rho_C \to 1.0$ for $b \in [0,1.2]$. In addition, as b increases, transitions to different stationary states can be observed for certain values of A. Notable examples occur at $b = 1.6$ for $A = 0.4$ and at $b = 1.2$ for $A = 0.6$, which will be explained below. When $A = 0.8$, cooperation cannot evolve even if the value of b is only slightly larger than 1.0. When $A = 2.0$, the result coincides with that for $A = 0.8$. It is worth pointing out that, in fact, when A is large, e.g., $A = 2.0$, individuals are always dissatisfied (as shown below), and that then our model recovers the traditional version of the prisoner's dilemma game. By comparing the results for $A = 2.0$ and those for other values of A, as shown in Fig. 1, we find that the present updating rule can effectively facilitate the evolution of cooperation. In particular, when $A \geq 2.0$, cooperators can survive only if $b < 1.05$. By contrast, with win-stay-lose-learn updating, cooperators can not only survive but also thrive even for much larger values of b.
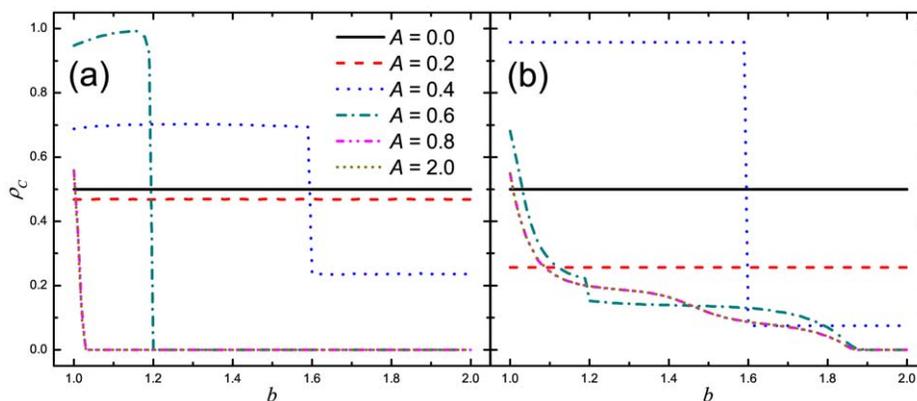
Figure 1. Win-stay-lose-learn promotes the evolution of cooperation, especially if intermediate aspirations determine the satisfiability of players. Presented is the stationary fraction of cooperators ρ_C in dependence on the temptation to defect b for different values of the aspiration A, as obtained by means of simulations (panel a) and the pair approximation (panel b). By comparing the results presented in the two panels, it can be observed that the pair approximation is to a large degree successful in reproducing the qualitative features of the simulations.
doi:10.1371/journal.pone.0030689.g001

The results obtained by means of pair approximation are presented in Fig. 1(b). It can be observed that the pair approximation can qualitatively correctly predict the cooperation level, especially for small values of A. For example, the result for $A = 0.0$ is exactly the same as the simulation result. The transition at $A = 0.4$ can be observed in Fig. 1(b). On the other hand, when $A > 0.5$, the satisfaction of individuals is increasingly difficult to achieve such that individuals tend to learn their neighbors when updating strategies, and the current model approaches the model of continuous updating [53]. Hence, there exist some differences between the results obtained by means of simulations and the pair approximation. Despite this, in general the results of the pair approximation support the main conclusions at which we arrive at by means of simulations.

In order to obtain a more complete picture about the joint effects of the aspiration level and the temptation to defect, we show the simulation results as a function of both A and b, as shown in Fig. 2. The results are consistent with those presented in Fig. 1(a), e.g., when $A \leq 0.25$, an intermediate level of cooperation ($\rho_C \approx 0.5$) is maintained, irrespective of the value of b. Within the interval of $0.25 < A < 0.5$, the cooperation level is higher than that for $A \leq 0.25$ and the transition can be observed at a fixed value of b for each value of A. It is interesting that the highest levels of cooperation occur within the interval of $0.5 \leq A < 0.75$. Moreover, it can be observed that as A increases, discontinuous transitions occur at $A = 0.0, 0.25, 0.5$ and 0.75.

These transitions can be understood as follows. On a square lattice with nearest neighbor interactions, the payoffs of a cooperator and a defector are given by $n_1 R + n_2 S$ and $n_3 T + n_4 P$, respectively, where $n_k \in \{0,1,2,3,4\}$, and $k \in \{1,2,3,4\}$. Given that $T = b$, $R = 1$, and $P = S = 0$, the above payoffs can be simplified as n_1 and $n_3 b$, respectively. In our model, when an individual is dissatisfied, it will learn a randomly chosen neighbor, which may lead to the change of ρ_C. For a cooperator, when $n_1 < 4A$, it is dissatisfied. While for a defector, the condition for its dissatisfaction is $n_3 b < 4A$. The phase transition points can be

obtained by letting $n_1 = 4A$ and $n_3 b = 4A$. Thus, the value of A at which phase transition occurs is given by $A = n_1/4$ and that for b is given by $b = (4A)/n_3 > 1$. Considering all the possible values of n_1, that is $n_1 = 0, 1, 2, 3$ and 4, we can obtain the phase transition points of A, which are $A = 0.0, 0.25, 0.5, 0.75$ and 1.0, as shown in Fig. 2. (As a matter of fact, $A = 1.0$ is also a phase transition point, however, because when $A = 1.0$, the density of cooperators is very low such that the phase transition phenomenon cannot be observed). The phase transition points of b can be calculated similarly, e.g., when $n_3 = 1$ and $A = 0.4$, the phase transition point of b is $b = 1.6$, which is verified by our simulation (see Fig. 2).

Since the strategy changes of individual players are determined by their satisfaction, we proceed with the results on the satisfaction rates in the population as a function of b for different values of A, as shown in Fig. 3. A highly cooperative society where each member is satisfied can be declared as the ultimate goal. If all members cooperate, then the social welfare will peak. Moreover, if then every member is satisfied, the society will be stable. We find from Fig. 3 that, if we regard the present system as a social prototype, then the optimal situation occurs within the interval of $b \in [1.0, 1.2]$ for $A = 0.6$, since it leads to a highly cooperative society with a high satisfaction rate. For the extreme case of $A = 0.0$, all individuals are satisfied. On the contrary, at the other extreme, i.e., at $A = 2.0$, no individual is ever satisfied. For $A = 0.2$, even though that more than 90% of individuals are satisfied, the cooperation level is not high ($\rho_C = 0.47$). This indicates that a large number of defectors are satisfied by exploiting cooperators. The obtained result for $A = 0.2$ reveals that a society where each member has a low aspiration level cannot be cooperative due to an inherent lack of incentives. When $A = 0.4$, the fraction of satisfied individuals drops suddenly at $b = 1.6$. When $b \leq 1.6$, nearly 70% of individuals cooperate and the satisfaction rate is high, which is more or less a better situation. Whereas $b > 1.6$ results in the low cooperation level as well as the low satisfaction rate, which is a society that should be avoided. When $A = 0.8$, few individuals in the population are satisfied such that almost all individuals seek for higher payoffs by imitation, and ultimately defection becomes the
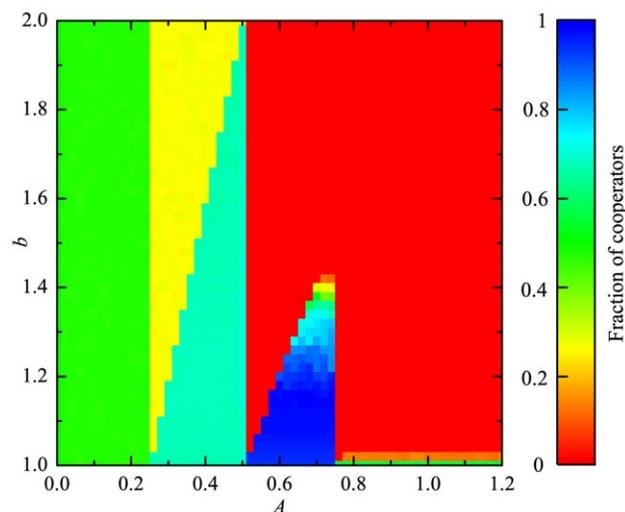

Figure 2. Transitions from predominantly defective to predominantly cooperative states in dependence on the aspiration level A and the temptation to defect b. Presented is the color-coded (see bar on the right) fraction of cooperators. The multitude of transitions in the color map points towards a high complexity of the underlying mechanisms warranting highly cooperative states (see main text for details).
doi:10.1371/journal.pone.0030689.g002

Figure 3. The fraction of satisfied players decreases with increasing aspirations. Presented is the fraction of satisfied players in the population, for which it holds that $P_i \geq A_i$, in dependence on the temptation to defect b for different values of the aspiration A. It is interesting to observe that for low values of A the fraction of satisfied players is independent of b, while for intermediate and large values of A it decreases with increasing b. Also note that for $A = 0.0$ ($A = 2.0$) all (no) players are satisfied.
doi:10.1371/journal.pone.0030689.g003

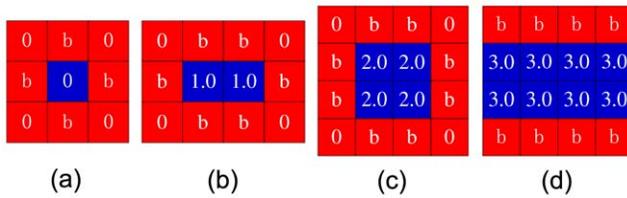**Figure 4. Special initial configurations of cooperators reveal their potential to expand into the territory of defectors.** In all panels the cooperators are depicted blue while defectors are depicted red. Each small square corresponds to a single player. Denoted values correspond to the payoffs of individual players, as obtained for the presented configurations. See also Fig. 5 for related results.
doi:10.1371/journal.pone.0030689.g004

unique choice. This confirms the standard game theoretical result that, in a society where individuals imitate each other, individual greediness (characterized by too high aspiration) may hinder the emergence of cooperation and eventually harm the benefit of each member of the society.

Because initial conditions are relevant for the evolutionary success of cooperators in spatial games [78,79], it is also of interest to test the robustness of the proposed updating rule. We thus investigate how cooperation evolves under different (adverse) initial conditions, which are shown in Fig. 4. We first focus on the initial configuration of cooperators and consider the case of Fig. 4(a), where only one cooperator exists in the population initially. For $A > 0$, the cooperator surely resorts to defection by imitation because of his dissatisfaction. Hence, a single cooperator surrounded by defectors cannot survive $A > 0$. When initially there

are two neighboring cooperators [see Fig. 4(b)], for $A \leq 0.25$, both cooperators and defectors at the boundary are satisfied such that the pattern is stable. However, when $0.25 < A \leq b/4$, defectors are satisfied but cooperators are not. Thus, cooperators will become defectors by imitation. When $A > b/4$, all individuals are dissatisfied such that all of them imitate neighbors' strategies. Since defectors have a higher payoff, cooperators are more likely to become defectors. Therefore, $A \leq 0.25$ is needed to make cooperators survive. When there exist four cooperators in the population initially, as shown in Fig. 4(c), the pattern is frozen if $A \leq b/4$. However, when $b/4 < A \leq 0.5$, cooperators can have advantages over defectors such that they can invade defectors and dominate the population ultimately, as shown in Fig. 5. This indicates the relevance of our model since it allows cooperators to thrive in harsh conditions where there only exist several cooperators initially. When $A > 0.5$, cooperators cannot expand their territories and, at the same time, they confront the intense invasion by defectors. Eventually cooperators are wiped out from the population. In this scenario, $A \leq 0.5$ is needed to maintain the pattern. This indicates that greediness may be detrimental to the emergence of cooperation. The more favorable case emerges when each cooperator has three cooperative neighbors, *i.e.* the population is initialized with two neighboring straight lines of cooperators, as shown in Fig. 4(d). Under these circumstances, $A \leq 0.75$ is required for cooperators to maintain strategies, which, however, does not warrant cooperators to expand territories. In order to realize the expansion, we need $A \leq 0.5$. When $A > 0.5$, the expansion of areas of cooperators is significantly restrained, which is demonstrated in Fig. 5. One can find that in order for the boom of cooperation, cooperators must, first, form clusters, which
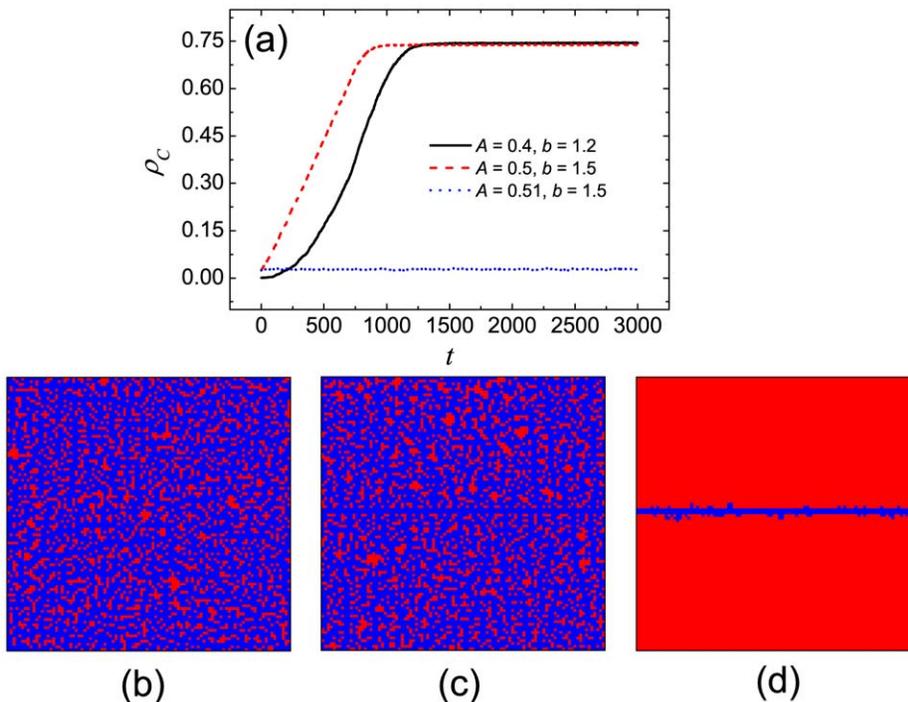

Figure 5. Robustness of the evolution of cooperation under adverse initial conditions. Panel (a) features the time evolution of the fraction of cooperators for different combinations of A and b, as obtained when using the initial conditions presented in Figs. 4(c) (black solid line) and (d) (dashed red and dotted blue line). Bottom row features the characteristic snapshots of the spatial grid (cooperators are blue, defectors are red), corresponding to the black solid line (panel b), the dashed red line (panel c), and the dotted blue line (panel d). It can be observed that cooperators may significantly outnumber defectors in the stationary state, even if starting from highly unfavorable conditions, as long as the aspirations are appropriately adjusted.
doi:10.1371/journal.pone.0030689.g005

ensures they have advantages over defectors in terms of payoffs. Second, cooperators cannot set their aspirations too high such that their satisfactions can easily be achieved [80]. Thus, they will hold their strategies, which is the precondition for the spreading of strategies. The fulfillment of the above two conditions as well as the dissatisfaction of defectors leads to the dissemination of the cooperative strategy in the population.

Lastly, we elaborate on how cooperators can resist the invasion by defectors. For this purpose, we consider the special initial conditions depicted in Fig. 6. Focusing first on Fig. 6(a), we find that when $A > 0.75$, cooperators are dissatisfied, which can lead to the extinction of cooperators. The case in Fig. 6(b) is qualitatively the same, the only difference being that the payoff of each defector is $3b$. In the situation where there exists a square block of four defectors [Fig. 6(c)], when $A \leq 0.75$, cooperators can at least survive. If $b/2 < A \leq 0.75$ ($b < 1.5$), cooperators can even invade the dissatisfied defectors. If $b > 1.5$, cooperators and defectors can coexist. On the contrary, when $A > 0.75$, cooperators are doomed to extinction. The above analysis explains why when $A > 0.75$, cooperators cannot flourish. That is, as long as cooperators do not set their goals too high (too greedy), cooperators can resist the invasion of defectors [80]. If they also have higher payoffs than neighboring defectors, as shown in Figs. 6(c) and (d), then defectors can be defeated.

Discussion

In summary, we have studied the impact of the win-stay-lose-learn strategy updating rule on the evolution of cooperation in the spatial prisoner's dilemma game. Unlike in the majority of previous works, in our case the strategy updating is not unconditional, but rather it depends on the level of satisfaction of individual players. The latter is determined by the aspiration level, which we have considered as a free parameter. If the payoff of a player is equal or higher than its aspiration, it is assumed that this player is satisfied and that there is thus no immediate need of changing its strategy. Conversely, if the payoff is lower than the aspired amount, the player will attempt to adopt the strategy of one of its nearest neighbors in the hope that it will reach the desired success. With this setup, we have found that if all players retain their strategies when being satisfied then the evolution of cooperation is remarkably facilitated. Especially for intermediate values of the aspiration parameter, e.g., $A = 0.6$, virtually complete cooperation dominance can be achieved even for values of the temptation to defect that significantly exceed 1. This is in sharp contrast to the results obtained with (too) large aspiration levels, e.g., $A = 2.0$, where the traditional version of the spatial prisoner's dilemma game is essentially fully recovered. The presented results also indicate that as long as individuals are not too greedy, i.e.,

aspire to modest (honest) incomes, cooperation thrives best, which is also in agreement with recent results obtained by means of a different model [80]. Moreover, we have tested the impact of different initial configurations, in particular such where cooperators initially have an inherent disadvantage over defectors, and we have discovered that the studied win-stay-lose-learn rule ensures that cooperators are able to spread even from very small numbers. In this sense, the proposed rule is very effective in unleashing the spreading potential of cooperative behavior, which is to some extent already provided (seeded) by means of spatial reciprocity [9]. We have also employed the pair approximation method to support our simulation results with semi-analytical calculations and to explain the observed transitions to different levels of cooperation on the square lattice.

It is instructive to discuss the differences between this work and related previous works [66,67,70–73]. For example, Chen and Wang [66] investigated a stochastic win-stay-lose-shift (WSLS) rule, under which dissatisfied individuals switch their strategies to the opposite one. It was reported that for small values of the temptation to defect cooperation can be best promoted at intermediate values of the aspiration level. Moreover, in [67] a N-person prisoner's dilemma game in a continuous population with a time-dependent aspiration level was investigated, while in [70] each individual had an aspiration-based learning motivation (which actually can depend directly on the aspiration level according to the rule of WSLS). It was reported that the results produced in [70] are similar to those in [66]. In [72] a payoff-based preferential learning mechanism was investigated (where individuals with higher payoffs are more likely to be imitated), and in [73] an aspiration-based preferential learning mechanism was studied where an individual whose strategy can provide the desired payoff when being imitated will be imitated also in the next round. In our model, however, we incorporate individual aspirations into the traditional imitation rule, and investigate how cooperation evolves under the aspiration-based conditional learning in the spatial prisoner's dilemma game. The proposed rule is simple and reasonable, and moreover, we show that it is effective and robust in promoting cooperation. In particular, cooperation can be maintained (and can even thrive) even under unfavorable initial conditions.

Our work also has some parallels with other models not necessarily incorporating aspirations. Namely, the model introduced recently in [81], where inertia was considered as something that can disable players to actively change their strategies, or with the early win-stay-lose-shift models [74–76]. Furthermore, it is possible to relate our work to those considering the importance of time scales in evolutionary games [44,45,47]. Note that under the presently introduced win-stay-lose-learn rule, for small values of A, the strategy transfers are rare and far apart in time. This has

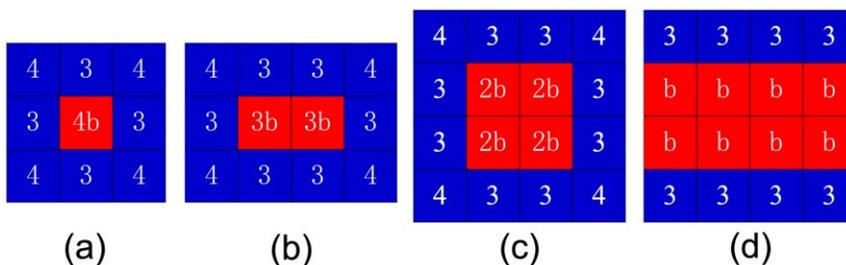**Figure 6. Special initial configurations of defectors reveal their potential to invade cooperators.** In all panels the cooperators are depicted blue while defectors are depicted red. Each small square corresponds to a single player. Denoted values correspond to the payoffs of individual players, as obtained for the presented configurations (see main text for details).
doi:10.1371/journal.pone.0030689.g006

similar consequences as when decreasing the strategy-selection rate [47]. For intermediate aspiration levels, however, we essentially have a segregated population if judging from individual satisfaction, i.e., some players are satisfied while others are not. This in turn implies that different players have different strategy-selection time scales. Even more importantly, these time scales are adaptive (they change over time), as players that eventually do change their strategies may go from being dissatisfied to becoming satisfied, or vice versa. In this sense, our model introduces different evolutionary time scales by means of aspiration, which is an endogenous property of individuals, and in so doing, it also relaxes the demand for their (the players') rationality. Note that in our model, there is frequently no need to compare payoffs with the neighbors, apart from when approaching the $A \rightarrow 2$ limit. Lastly, we hope that this study will enrich our knowledge on how to successfully resolve the prisoner's dilemma, and we hope it will inspire further work along this very interesting and vibrant avenue of research.

Methods

The spatial prisoner's dilemma game is staged on a square lattice of size $L \times L$ with periodic boundary conditions. In accordance with common practice [9], the payoffs are as follows: $T = b$ is the temptation to defect, $R = 1$ is the reward for mutual cooperation, while $P = S = 0$ are the punishment for mutual defection and the sucker's payoff, respectively, where $1 < b < 2$. Although this formulation of the game has $P = S$ rather than $P > S$, it captures succinctly the essential social dilemma, and accordingly, the presented results can be considered fully relevant and without loss of generality with respect to more elaborated formulations of the payoffs. Moreover, each player i has an aspiration level $A_i = k_i A$, where k_i is the player's degree and A is a free parameter that determines the overall aspiration level of the population, which is typically constrained to the interval $0 < A < b$. Since we consider the square lattice as the interaction network, we have $k_i = k = 4$, which in turn postulates that each player in this study has an equal aspiration equal to kA.

Player i acquires its payoff P_i by playing the game with its four nearest neighbors. A randomly selected nearest neighbor j acquires its payoff P_j likewise by playing the game with its four nearest neighbors. If $P_i \geq A_i$, i.e., if the payoff of player i is equal or higher than its aspiration, then strategy adoption from player j is not attempted. If, however, $P_i < A_i$, then player i adopts the strategy of player j with the probability

$$W = \frac{1}{1 + \exp[(P_i - P_j)/\kappa]} \qquad (1)$$

where κ determines the amplitude of noise [33], accounting for imperfect information and errors in decision making. It is well-known that there exists an optimal intermediate value of κ at which the evolution of cooperation is most successful [34,35], yet in general the outcome of the prisoner's dilemma game is robust to variations of κ. Without much loss of generality, we use $\kappa = 0.1$, meaning that it is very likely that the better performing players will pass their strategy to other players, yet it is not impossible that players will occasionally learn also from the less successful neighbors. The simulations of this spatial prisoner's dilemma game were performed by means of a synchronous updating rule, using $L = 100$ to 400 system size and discarding the transient times prior to reaching the stationary states before recording the average fraction of cooperators ρ_C in the population. We have verified that the presented results are robust to variations of the system size, and to the variation of the simulation procedure (e.g., by using random rather than synchronous updating). It is also worth noting that because $A < b$ and $b < 2$, the present definition of the win-stay-lose-learn transforms to the traditional spatial prisoner's dilemma game when $A \geq 2.0$, given that then individual cannot be satisfied and thus attempt to change their strategy whenever they receive a chance to do so.

In addition to the simulation results of the proposed spatial game, we also present the results of pair approximation [53,82–85] that are obtained with the rate equations of cooperator-cooperator (c,c) and cooperator-defector (c,d) edges, which are as follows:

$$\dot{p}_{c,c} = \sum_{x,y,z} [n_c(x,y,z) + 1] p_{d,x} p_{d,y} p_{d,z} \times$$
$$\sum_{u,v,w} p_{c,u} p_{c,v} p_{c,w} f[P_d(x,y,z), P_c(u,v,w)]$$
$$- \sum_{x,y,z} n_c(x,y,z) p_{c,x} p_{c,y} p_{c,z} \times$$
$$\sum_{u,v,w} p_{d,u} p_{d,v} p_{d,w} f[P_c(x,y,z), P_d(u,v,w)], \qquad (2)$$

$$\dot{p}_{c,d} = \sum_{x,y,z} [1 - n_c(x,y,z)] p_{d,x} p_{d,y} p_{d,z} \times$$
$$\sum_{u,v,w} p_{c,u} p_{c,v} p_{c,w} f[P_d(x,y,z), P_c(u,v,w)]$$
$$- \sum_{x,y,z} [2 - n_c(x,y,z)] p_{c,x} p_{c,y} p_{c,z} \times$$
$$\sum_{u,v,w} p_{d,u} p_{d,v} p_{d,w} f[P_c(x,y,z), P_d(u,v,w)], \qquad (3)$$

where x,y,z are either cooperators or defectors and $n_c(x,y,z)$ denote the number of cooperators among x,y,z. Moreover,

$$f(P_i, P_j) = \begin{cases} \dfrac{1}{1 + exp[(P_i - P_j)/\kappa]}, & P_i < A_i \\ 0, & P_i \geq A_i \end{cases} \qquad (4)$$

where P_i and P_j are the payoffs of the two neighboring players i and j, respectively, and A_i is the payoff aspiration of player i (equal to Ak for all i). By performing the numerical integration for the above two differential equations (2,3), and by using $p_{c,d} = p_{d,c}$ and $p_{c,c} + p_{c,d} + p_{d,c} + p_{d,d} = 1$, we can obtain ρ_C from $p_{c,c} + p_{c,d}$.

Author Contributions

References

1. Maynard Smith J (1982) Evolution and the Theory of Games. Cambridge, UK: Cambridge University Press.

2. Sigmund K (1993) Games of Life: Exploration in Ecology, Evolution and Behavior. Oxford, UK: Oxford University Press.

3. Hofbauer J, Sigmund K (1998) Evolutionary Games and Population Dynamics. Cambridge, UK: Cambridge University Press.

4. Nowak MA (2006) Evolutionary Dynamics. Cambridge, MA: Harvard University Press.

5. Sigmund K (2010) The Calculus of Selfishness. Princeton, MA: Princeton University Press.

6. Axelrod R (1984) The Evolution of Cooperation. New York: Basic Books.

7. Hardin G (1968) The tragedy of the commons. Science 162: 1243–1248.

8. Nowak MA (2006) Five rules for the evolution of cooperation. Science 314: 1560–1563.

9. Nowak MA, May RM (1992) Evolutionary games and spatial chaos. Nature 359: 826–829.

10. Szabó G, Fáth G (2007) Evolutionary games on graphs. Phys Rep 446: 97–216.

11. Roca CP, Cuesta JA, Sánchez A (2009) Evolutionary game theory: Temporal and spatial effects beyond replicator dynamics. Phys Life Rev 6: 208–249.

12. Perc M, Szolnoki A (2010) Coevolutionary games – a mini review. BioSystems 99: 109–125.

13. Santos FC, Pacheco JM (2005) Scale-free networks provide a unifying framework for the emergence of cooperation. Phys Rev Lett 95: 098104.

14. Vukov J, Szabó G (2005) Evolutionary prisoner's dilemma game on hierarchical lattices. Phys Rev E 71: 036133.

15. Santos FC, Pacheco JM, Lenaerts T (2006) Evolutionary dynamics of social dilemmas in structured heterogeneous populations. Proc Natl Acad Sci USA 103: 3490–3494.

16. Santos FC, Rodrigues JF, Pacheco JM (2006) Graph topology plays a determinant role in the evolution of cooperation. Proc R Soc B 273: 51–55.

17. Gómez-Gardeñes J, Campillo M, Moreno Y, Floría LM (2007) Dynamical organization of cooperation in complex networks. Phys Rev Lett 98: 108103.

18. Poncela J, Gómez-Gardeñes J, Floría LM, Moreno Y (2007) Robustness of cooperation in the evolu- tionary prisoner's dilemma on complex systems. New J Phys 9: 184.

19. Lozano S, Arenas A, Sánchez A (2008) Mezoscopic structure conditions the emergence of cooperation on social networks. PLoS ONE 3: e1892.

20. Poncela J, Gómez-Gardeñes J, Floría LM, Sánchez A, Moreno Y (2008) Complex cooperative networks from evolutionary preferential attachment. PLoS ONE 3: e2449.

21. Poncela J, Gómez-Gardeñes J, Floría LM, Moreno Y, Sánchez A (2009) Cooperative scale-free networks despite the presence of defector hubs. EPL 88: 38003.

22. Poncela J, Gómez-Gardeñes J, Traulsen A, Moreno Y (2009) Evolutionary game dynamics in a growing structured population. New J Phys 11: 083031.

23. Pacheco JM, Pinheiro FL, Santos FC (2009) Population structure induces a symmetry breaking favoring the emergence of cooperation. PLoS Comput Biol 5.

24. Grujić J, Fosco C, Araujo L, Cuesta JA, Sánchez A (2010) Social experiments in the mesoscale: Humans playing a spatial prisoner's dilemma. PLoS ONE 5: e13749.

25. Dai Q, Li H, Cheng H, Li Y, Yang J (2010) Double-dealing behavior potentially promotes cooperation in evolutionary prisoner's dilemma games. New J Phys 12: 113015.

26. Poncela J, Gómez-Gardeñes J, Moreno Y (2011) Cooperation in scale-free networks with limited associative capacities. Phys Rev E 83: 057101.

27. Lee S, Holme P, Wu ZX (2011) Emergent hierarchical structures in multiadaptive games. Phys Rev Lett 106: 028702.

28. Dai Q, Cheng H, Li H, Li Y, Zhang M, et al. (2011) Crossover between structured and well-mixed networks in an evolutionary prisoners dilemma game. Phys Rev E 84: 011103.

29. Tomassini M, Luthi L, Pestelacci E (2007) Social dilemmas and cooperation in complex networks. Int J Mod Phys C 18: 1173–1185.

30. Wu ZX, Guan JY, Xu XJ, Wang YH (2007) Evolutionary prisoner's dilemma game on Barabási-Albert scale-free networks. Physica A 379: 672–680.

31. Szolnoki A, Perc M, Danku Z (2008) Towards effective payoffs in the prisoner's dilemma game on scale-free networks. Physica A 387: 2075–2082.

32. Peña J, Volken H, Pestelacci E, Tomassini M (2009) Conformity hinders the evolution of cooperation on scale-free networks. Phys Rev E 80: 016110.

33. Szabó G, Tőke C (1998) Evolutionary prisoner's dilemma game on a square lattice. Phys Rev E 58: 69–73.

34. Szabó G, Vukov J, Szolnoki A (2005) Phase diagrams for an evolutionary prisoner's dilemma game on two-dimensional lattices. Phys Rev E 72: 047107.

35. Perc M (2006) Coherence resonance in spatial prisoner's dilemma game. New J Phys 8: 22.

36. Vukov J, Szabó G, Szolnoki A (2006) Cooperation in the noisy case: Prisoner's dilemma game on two types of regular random graphs. Phys Rev E 73: 067103.

37. Szolnoki A, Vukov J, Szabó G (2009) Selection of noise level in strategy adoption for spatial social dilemmas. Phys Rev E 80: 056112.

38. Szabó G, Szolnoki A, Vukov J (2009) Selection of dynamical rules in spatial prisoner's dilemma games. EPL 87: 18007.

39. Ohtsuki H, Nowak MA, Pacheco JM (2007) Breaking the symmetry between interaction and replacement in evolutionary dynamics on graphs. Phys Rev Lett 98: 108106.

40. Ohtsuki H, Pacheco JM, Nowak MA (2007) Evolutionary graph theory: Breaking the symmetry between interaction and replacement. J Theor Biol 246: 681–694.

41. Hu MB, Jiang R, Wu QS, Wu YH (2007) Simulating the wealth distribution with a richest-following strategy on scale-free networks. Physica A 381: 467–472.

42. Perc M, Szolnoki A (2008) Social diversity and promotion of cooperation in the spatial prisoner's dilemma game. Phys Rev E 77: 011904.

43. Santos FC, Santos MD, Pacheco JM (2008) Social diversity promotes the emergence of cooperation in public goods games. Nature 454: 213–216.

44. Wu ZX, Rong Z, Holme P (2009) Diversity of reproduction time scale promotes cooperation in spatial prisoner's dilemma games. Phys Rev E 80: 036103.

45. Roca CP, Cuesta JA, Sánchez A (2006) Time scales in evolutionary dynamics. Phys Rev Lett 97: 158701.

46. Szolnoki A, Perc M, Danku Z (2008) Making new connections towards cooperation in the prisoner's dilemma game. EPL 84: 50007.

47. Rong Z, Wu ZX, Wang WX (2010) Emergence of cooperation through coevolving time scale in spatial prisoner's dilemma. Phys Rev E 82: 026101.

48. Wardil L, Silva JKL (2009) Adoption of simultaneous different strategies against different opponents enhances cooperation. EPL 86: 38001.

49. Liu Y, Li Z, Chen X, Wang L (2010) Memory-based prisoner's dilemma on square lattices. Physica A 389: 2390–2396.

50. Szolnoki A, Szabó G (2007) Cooperation enhanced by inhomogeneous activity of teaching for evolutionary prisoner's dilemma games. EPL 77: 30004.

51. Szolnoki A, Perc M (2008) Coevolution of teaching activity promotes cooperation. New J Phys 10: 043036.

52. Szabó G, Szolnoki A (2009) Cooperation in spatial prisoner's dilemma with two types of players for increasing number of neighbors. Phys Rev E 79: 016106.

53. Guan JY, Wu ZX, Huang ZG, Xu XJ, Wang YH (2006) Promotion of cooperation induced by nonlinear attractive effect in spatial prisoner's dilemma game. Europhys Lett 76: 1214–1220.

54. Wu ZX, Xu XJ, Huang ZG, Wang SJ, Wang YH (2006) Evolutionary prisoner's dilemma game with dynamic preferential selection. Phys Rev E 74: 021107.

55. Vainstein MH, Silva ATC, Arenzon JJ (2007) Does mobility decrease cooperation? J Theor Biol 244: 722–728.

56. Helbing D, Yu W (2009) The outbreak of cooperation among success-driven individuals under noisy conditions. Proc Natl Acad Sci USA 106: 3680–3685.

57. Meloni S, Buscarino A, Fortuna L, Frasca M, Gomez-Gardeñes J, et al. (2009) Effects of mobility in a population of prisoner's dilemma players. Phys Rev E 79: 067101.

58. Droz M, Szwabinski J, Szabó G (2009) Motion of influential players can support cooperation in prisoner's dilemma. Eur Phys J B 71: 579–585.

59. Wu ZX, Holme P (2009) Effects of strategy-migration direction and noise in the evolutionary spatial prisoner's dilemma. Phys Rev E 80: 026108.

60. Chen X, Wang L (2009) Cooperation enhanced by moderate tolerance ranges in myopically selective interactions. Phys Rev E 80: 046109.

61. Fu F, Hauert C, Nowak MA, Wang L (2008) Reputation-based partner choice promotes cooperation in social networks. Phys Rev E 78: 026117.

62. Fu F, Wu T, Wang L (2009) Partner switching stabilizes cooperation in coevolutionary prisoner's dilemma. Phys Rev E 79: 036101.

63. Van Segbroeck S, Santos FC, Lenaerts T, Pacheco JM (2009) Reacting differently to adverse ties promotes cooperation in social networks. Phys Rev Lett 102: 058105.

64. Cho IK, Matsui A (2005) Learning aspiration in repeated games. J Econ Theor 124: 171–201.

65. Macy MW, Flache A (2002) Learning dynamics in social dilemmas. Proc Natl Acad Sci USA 99: 7229–7236.

66. Chen X, Wang L (2008) Promotion of cooperation induced by appropriate payoff aspirations in a small- world networked game. Phys Rev E 77: 017103.

67. Platkowski T, Bujnowski P (2009) Cooperation in aspiration-based n-person prisoner's dilemmas. Phys Rev E 79: 036103.

68. Platkowski T (2009) Enhanced cooperation in prisoner's dilemma with aspiration. Appl Math Lett 22: 1161–1165.

69. Yang HX, Wu ZX, Wang BH (2010) Role of aspiration-induced migration in cooperation. Phys Rev E 81: 065101(R).

70. Zhang H, Small M, Yang H, Wang B (2010) Adjusting learning motivation to promote cooperation. Physica A 389: 4734–4739.

71. Wang Z, Perc M (2010) Aspiring to the fittest and promotion of cooperation in the prisoner's dilemma game. Phys Rev E 82: 021115.

72. Perc M, Wang Z (2011) Heterogeneous aspirations promote cooperation in the prisoner's dilemma game. PLoS ONE 5: e15117.

73. Liu Y, Chen X, Wang L, Li B, Zhang W, et al. (2011) Aspiration-based learning promotes cooperation in spatial prisoner's dilemma games. EPL 94: 60002.

74. Nowak MA, Sigmund K (1993) A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. Nature 364: 56–58.

75. Posch M (1999) Win-stay, lose-shift strategies for repeated games–memory length, aspiration levels and noise. J Theor Biol 198: 183–195.

76. Imhof LA, Fudenberg D, Nowak MA (2007) Tit-for-tat or win-stay, lose-shift? J Theor Biol 247: 574–580.

77. Matsen FA, Nowak MA (2004) Win-stay, lose-shift in language learning from peers. Proc Natl Acad Sci USA 101: 18053–18057.

78. Langer P, Nowak MA, Hauert C (2008) Spatial invasion of cooperation. J Theor Biol 250: 634–641.

79. Chen X, Fu F, Wang L (2008) Influence of different initial distributions on robust cooperation in scale- free networks: A comparative study. Phys Lett A 372: 1161–1167.

80. Roca CP, Helbing D (2011) Emergence of social cohesion in a model society of greedy, mobile individuals. Proc Natl Acad Sci USA 108: 11370–11374.

81. Liu RR, Rong Z, Jia CX, Wang BH (2010) Effects of diverse inertia on scale-free-networked prisoner's dilemma games. EPL 91: 20002.

82. Hauert C, Doebeli M (2004) Spatial structure often inhibits the evolution of cooperation in the snowdrift game. Nature 428: 643–646.

83. Hauert C, Szabó G (2005) Game theory and physics. Am J Phys 73: 405–414.

84. Wu ZX, Wang YH (2007) Cooperation enhanced by the difference between interaction and learning neighborhoods for evolutionary spatial prisoner's dilemma games. Phys Rev E 75: 041114.

85. Chen X, Fu F, Wang L (2008) Interaction stochasticity supports cooperation in spatial prisoner's dilemma. Phys Rev E 78: 051120.