



OPEN

Swarm intelligence inspired skills and the evolution of cooperation

SUBJECT AREAS:

EVOLVABILITY

EXPERIMENTAL EVOLUTION

Haibin Duan^{1,2} & Changhao Sun^{1,2}

¹State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing 100191, P. R. China, ²Science and Technology on Aircraft Control Laboratory, School of Automation Science and Electronic Engineering, Beihang University, Beijing 100191, P. R. China.

Received
24 February 2014

Accepted
2 May 2014

Published
9 June 2014

Correspondence and requests for materials should be addressed to H.B.D. (hbduan@buaa.edu.cn)

Many hostile scenarios exist in real-life situations, where cooperation is disfavored and the collective behavior needs intervention for system efficiency improvement. Towards this end, the framework of soft control provides a powerful tool by introducing controllable agents called skills, who are allowed to follow well-designed updating rules for varying missions. Inspired by swarm intelligence emerging from flocks of birds, we explore here the dependence of the evolution of cooperation on soft control by an evolutionary iterated prisoner's dilemma (IPD) game staged on square lattices, where the skills adopt a particle swarm optimization (PSO) mechanism for strategy updating. We demonstrate that not only can cooperation be promoted by skills effectively seeking for potentially better strategies and spreading them to others, but also the frequency of cooperation could be arbitrarily controlled by choosing appropriate parameter settings. Moreover, we show that adding more skills does not contribute to further cooperation promotion, while assigning higher weights to the collective knowledge for strategy updating proves a efficient way to induce cooperative behavior. Our research provides insights into cooperation evolution in the presence of PSO-inspired skills and we hope it will be inspirational for future studies focusing on swarm intelligence based soft control.

Cooperation is omnipresent in real-world scenarios and plays a fundamental role for complex organization structures, ranging from biological systems to economic activities of human beings. However, it seems to be at variance with natural selection of Darwin's evolutionary theory¹, which implies fierce competition for survival among selfish and unrelated individuals². Hence, understanding the emergence and sustainability of wide spread cooperative behavior is one of the central issues in both biology and social science. This problem is often tackled within the framework of evolutionary game theory^{3–5}. As the most stringent situation of reciprocal behavior through pairwise interactions, the prisoner's dilemma (PD) has long been considered as a paradigmatic example for studying the dilemmas between individual interests and collective welfare. In its original form, the PD is a two-player non-zero-sum game, where each player decides simultaneously whether to cooperate (C) or defect (D) without knowing *a priori* how its opponent will act. There are four possible outcomes for this game: (1) mutual cooperation (C, C) yields the largest collective payoff by offering each a reward R , (2) mutual defection (D, D) pays each a punishment P , and (3) the mixed choices (C, D) or (D, C) give the defector a temptation T and the cooperator the sucker's payoff S ^{6,7}, with the payoff ranking satisfying $T > R > P > S$ and $2R > T + S$ ⁸. The dilemma is given by the fact that although the collective payoff would be maximized if both cooperated, it is best for a rational player to defect no matter what strategy its opponent chooses in a single round, making defection the only equilibrium. To allow the evolution of cooperation in the PD, suitable extensions have been proposed accordingly. One possible way out is direct reciprocity⁹ in the iterated prisoner's dilemma (IPD) game¹⁰, where players meet more than once, keep in mind the results of previous encounters and play repeatedly. Under this circumstance, reciprocity is regarded as a crucial property for winning strategies, including, but not limited to, *Tit for Tat* (TFT)^{11–14}, *generous TFT* (GTFT)¹⁵, *win-stay lose-shift*^{16,17}, to name but a few. Secondly, placing players on spatial networked structures, e.g., the square lattice, the small-world network, the scale-free network as well as diluted and interdependent networks^{18–23}, has been acknowledged to be a new route to promotion and maintenance of cooperative behavior without any other assumptions or strategy complexity²⁴. Though this is not universally true^{25,26}, spatial interactions do provide cooperators an opportunity to agglomerate and grow, by which cooperators can finally resist exploitation by defectors and survive extinction in most cases^{27–33}. Finally, other mechanisms favoring cooperation include co-evolution of the network structure along with playing rules^{34,35}, the ability to move and avoid nasty encounters³⁶, the freedom to withdraw from the game^{37,38}, and punishment and reward^{39–42}.



While the various mechanisms have achieved significant success in interpreting the evolution of cooperation, they are largely based on the assumption of particular networks of contacts and rules of local interaction. In other words, previous works mainly focus on which scenario favors cooperation. In the meanwhile, however, there exist many hostile scenarios in real life situations, where cooperation is disfavored and hence the undesirable outcomes need to be controlled for system efficiency improvement or disaster avoidance. For example, cooperative communications have been considered to be a promising transmit paradigm for future wireless networks, in which the interactions between neighboring nodes can be modeled via the evolution of cooperation. However, when the cost to cooperate exceeds a threshold, individual nodes tend to refuse to offer help to others, thus posing a great challenge to cooperation emergence as well as the performance of the whole system⁴³. Generally, this problem becomes especially difficult when it is hard or even impossible to change the underlying playing rules of the original individuals, such as the behavior rules of crowds in a panic and the flying strategies of a flock of birds⁴⁴. It is therefore of great interest and practical significance to design schemes that can effectively intervene in the collective behavior of a particular system, such that a proper equilibrium is guaranteed given the local rules of individual players and the network of contacts. Though many efforts have been devoted to pinning control law design and theoretical analysis of differential equations based control systems^{45,46}, few literature have studied its application to complex systems of networked evolutionary games. To address this particular issue, Han *et al.* have proposed a new framework termed as soft control, which aims to induce desired collective behavior out of a multi-agent system via introducing a few controllable individuals, called shills, to the original population consisting of normal individuals^{47,48}. Within the framework, shills are treated equally as normal agents by conforming to the basic playing rules, but are allowed to adopt elaborated strategies and updating rules for different mission objectives. Different from conventional distributed control that focuses on designing local rules for each agent in the network, soft control treats all the original individuals as one system and the control law is only for the shills. In real-world applications, there are two ways to account for the existence of shills: (i) under some certain circumstance, particular individuals of the original population turn into shills either voluntarily or forced by external factors, e.g., men of high principle who stand out in chaotic scenes to maintain order; (ii) additional individuals that are added to the original population to intervene in the macroscopic feature of the system, e.g., a few detectors who sneak into a band of gangsters with the purpose of making them confess⁴⁹. So far, several preliminary proposals have been advanced to control the macroscopic behavior of multi-agent systems. For example, it has been proposed that a controllable powerful robot bird might be used as a shill to drive flocking of birds in airports, assuming the underlying dynamics of real birds follows Vicsek model^{47,50}. Another work by Wang studied soft control in the well-mixed population case, where frequency-based tit for tat (F-TFT) was utilized as a shill's strategy⁴⁸.

This paper goes beyond the updating rules of conventional evolutionary game theory and considers instead a swarm intelligence inspired strategy updating mechanism for shills. Generally, swarm intelligence refers to the collective behavior of a decentralized, self-organized natural system, such as a flock of birds trying to reach an unknown destination, a colony of ants searching for the best path to the food source, and a group of people brainstorming to come up with a solution to a specific problem⁵¹. Due to its decentralized distribution and self-adaptation ability, swarm intelligence has provided a good source of intuition for solutions to realistic problems and several models have been proposed accordingly, among which are particle swarm optimization (PSO) proposed by Eberhart in collaboration with Kennedy^{52,53}, ant colony optimization (ACO) by Dorigo^{54,55}, and brain storm optimization (BSO) developed by

Shi^{56–58}. The fundamental concept of PSO is derived from the underlying rules that enable large numbers of birds to flock synchronously, where each individual in the swarm is seen as a particle, moving in a multi-dimensional space and searching for an unknown food source. Following random initialization of position and velocity vectors, each individual updates its state by combining one's past experience with some aspects of collective knowledge of the whole swarm in the following generations. As the searching process carries on, the population of particles, like a flock of birds foraging for food sources, move closer to and can eventually find the optimum of the utility function. Since miraculous swarm intelligence emerges from the simple and self-organized particles in PSO, we introduce the strategy updating mechanism to the shills and study whether a more desirable collective behavior could be induced in hostile environments. Furthermore, other than assuming a well-mixed population which may not be always true in real world systems, we extend soft control to a structured population, with each individual located on the vertex of a square lattice and its interaction restricted to the immediate neighbors along the social ties.

We address the above issue by introducing a model in which normal individuals interplay with a small fraction of shills and study what occurs depending on the frequency of each type of individuals and relative parameters tuning the updating procedure. To be specific, we choose a finite IPD game played on a square lattice as our strategic problem, where each individual adopts a stochastic strategy and plays according to Markov process during each evolution time step. All the details can be found in the Methods section. In what follows, we will focus on the influence of PSO inspired soft control on the evolution of cooperation under different settings by numerical simulations. Firstly, we start our study by varying the game strength from 1.1 to 2, in order to demonstrate that the proposed methodology takes effect upon cooperation promotion throughout a wide range of conflict intensity. Following this, we proceed to study the impact of strategy diversity resulting from soft control and find that PSO mechanisms enable evolutionary outcomes to be less dependent on random factors. Finally, we focus on the most unfavorable scenario of cooperation and study how the two crucial parameters, the shill fraction p_s and the weighting coefficient ω , influence the evolution. We show that the frequency of cooperation can not only be promoted by shills, but can also be controlled to an expected value by choosing appropriate parameter settings.

Results

Numerical simulations are carried out on an $L \times L$ square lattice with $N=L^2$ vertices and equal connectivity $\langle z \rangle = 4$. The results shown were obtained for communities of $N=10000$ individuals. The equilibrium results, including frequencies of cooperation f_c and average strategies $\langle x \rangle$, are averaged over the last 10% time steps after a transient period of 300 time steps. This procedure is repeated 100 times for 100 independent random realizations of the game considered. The frequency of cooperation (or cooperation level) f_c used to evaluate the system performance is defined as the percentage of Cs of all the actions taken by all individuals during one time step t . As there are $8TN$ actions in $4N$ T -IPD games, f_c is calculated via

$$f_c = \frac{1}{8TN} \sum_{i=1}^N w(i) \quad (1)$$

where $w(i)$ denotes the number of Cs during the interactions between agent i and its four neighbors.

As shown in Figure 1(a), with a small fraction ($p_s=0.05$) of PSO-inspired shills introduced, the average cooperation level of the original population is greatly promoted, despite the fact that much fiercer conflict arises between individual interests and collective welfare when the temptation to defect b increases from 1 to 2. This should be largely attributed to the global search ability of PSO, by



which the skills can find more successful strategies and spread them to the neighbors. In such a case, the temptation to defect b has a negligible effect on the value of f_c in the equilibrium, which, as a matter of fact, only diminishes from 0.944 for $b=1$ to 0.933 for $b=2$. On the other hand, although spatial reciprocity and iterated interactions favor the emergence and maintenance of cooperation to some extent, they cannot guarantee a desirable level of cooperation in more defection-prone environments as the game intensity b grows. Expectedly and as shown by red circles in Figure 1(a), we observe a gradual decline of f_c , which finally drops below the initial level ($f_{c0}=0.5$) of cooperation as b approaches its maximum limit of 2. It is also important to note that direct reciprocity emerging from iterated interactions does effect the evolutionary outcome, as we find that similar results cannot be obtained for one-shot IPD with $b=1$.

As each individual adopts a stationary Markov strategy (p_0, p_c, p_d) (see Methods), we examine the average state $\langle x \rangle = (\langle p_0 \rangle, \langle p_c \rangle, \langle p_d \rangle)$ reached by the population in the equilibrium in Figure 1(b),

each element of which denotes the average probability for a random individual to cooperate in the first stage of the IPD, the conditional probability to cooperate when the last move of the opponent is C or D respectively. For $p_s=0.05$ and $1 < b < 2$, the individuals come to a consensus of Markov strategy approximate to (1,1,0), which can be viewed as TFT. Actually, TFT has been proven the simplest and the most successful strategy in the IPD game, via which one offers to cooperate initially, punishes defectors and reward cooperators in the successive rounds with “an eye for an eye”. Hence, desirable social welfare can always be achieved in such a case, as shown by blue squares in Figure 1(c). On the contrary, the original population without skills results in a worse strategy, in which one tends to cooperate with smaller probabilities as b grows. As a consequence, the average payoff an individual receives for each action diminishes from 0.9 to 0.4, as shown by the red circles in Figure 1(c). To have a better understanding of the evolution process, we plot the cooperation frequency as a function of time steps t in Figure 1(d) for both

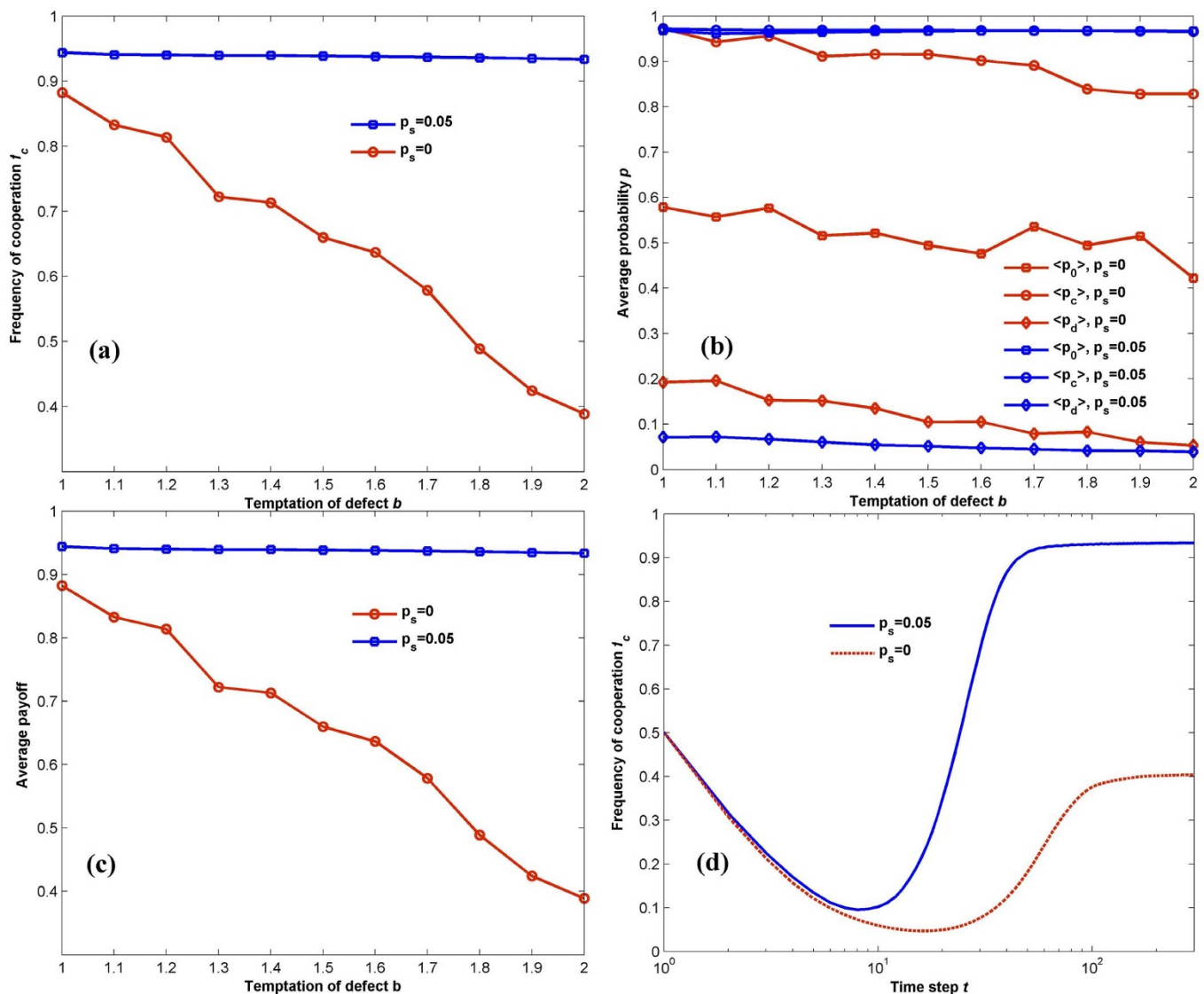


Figure 1 | Comparison of evolution characteristics between cases with ($p_s=0.05$) and without skills ($p_s=0$). (a): average level of cooperation in dependence of b . (b): average Markov strategy of the population as a function of b . (c): average payoffs of one action in dependence of b . (d): evolution of cooperative behavior concentration over time step t for $b=2$. We observe that while the spatial structure and iterated interactions contribute to the promotion of cooperation in PD games compared with one-shot games in a well-mixed population, soft control with a small proportion ($p_s=0.05$) of skills can further enhance cooperation levels even in strongly defection-prone environments, resulting in an average strategy of TFT. All trajectories are averaged over 100 independent realizations of the game considered, whose parameter setting is $L=100$, $T=50$, $\omega=0.95$, and the maximum evolution generation $G_{\max}=300$.



cases. Due to the random initialization for strategies, the evolutionary curves both start from $f_c=0.5$, an equal probability for each agent to cooperate or defect in the first generation. Similar to the cases in most literature^{15,59}, the evolution courses follow the pattern of endurance and expansion: defectors take advantage of the random arrangements at the early stage and can thus make the greatest profits by exploiting cooperators. It enables defectors to spread across the population such that only little clusters of cooperators exist and a decline of f_c is observed in the critical time step. Restarting from the lowest point, where an ordered distribution of cooperators and defectors is reached, clusters of cooperators begin to expand until a new equilibrium between cooperation and defection is achieved, so we see a rise of cooperation levels which gradually converge to a stable state.

It is also found that, without soft control, the outcome of one random run of the simulation is more dependent on the initial strategy profile and its distribution, as is shown in Figure 2. For $p_s=0$, the stationary outcomes of five runs diverge significantly from the average performance value in Figure 2(b), whereas the dramatic deviation is removed by the proposed soft control mechanism with $p_s=0.05$ in Figure 2(a). This finding is easy to understand if we analyze the evolutionary dynamics of the two cases. On the one hand, as the players in the original population adopt the updating rule of *unconditional imitation* (see the Methods section for the detailed definition), which means strategies are produced by copying the old ones and thus no new acting tactics are created, the strategy diversity without soft control is highly dependent on the initial strategy profile. As a consequence, the outcome of one random simulation makes little sense in representing the evolution dynamics, for it is more sensitive to randomness. Yet, as more runs of simulations are carried out, the result gradually converges to a value which can be used to approximate the average cooperation level within a specific parameter setting. On the other hand, with PSO inspired soft control, new strategies are adaptively generated to maximize the potential payoff by considering simultaneously the most profitable strategy in the past and the best strategy of the neighbors, enabling the strategy variety to be less dependent of randomness brought about by the initialization operation. In sum, while promoting cooperation in strongly hostile environments, the proposed approach can lower

the impact of random factors by adding to the strategy diversity in the population.

The effect of skills in the boost of cooperation can be partly comprehended by the comparison between the evolution courses of cooperation frequencies for the whole population and the individuals within the neighborhood of skills. Focusing on Figure 3, we have two important findings: firstly, while both cases experience a decreasing phase of cooperation due to invasion of defectors in the early stage, the individuals adjoining to skills do maintain a higher level of cooperation; secondly, however, as cooperative behavior spreads with the evolution carrying on, the existence of skills hampers its further expansion. These results demonstrate that, although skills facilitate propagation of cooperation by exploring the strategy space, especially in the endurance period when the individuals are mostly defectors, the current parameter setting ($p_s=0.05$, $\omega=0.95$) cannot guarantee a global best strategy for skills in the equilibrium. Since the skills imitate the updating mechanism in PSO, the results can be interpreted as follows: unlike the normal agents who greedily switch their strategies to the best one within the neighborhood or remain unchanged, skills conduct effective search for potentially better strategies in a continuous 3-dimension space, which helps to increase the fraction of cooperation in skills and their neighbors in the early stage. On the other hand, as $\omega=0.95$ is used to balance the weight between one's best strategy in the history as well as the most profitable strategy within its neighborhood, skills tend to take into account more history information when updating strategies, which plays a less important role for payoff improvement in the current situation. As a result, while the average strategy of the whole population converges approximately to TFT, the skills result in an reacting rule $\langle x \rangle = (0.58, 0.55, 0.26)$ as shown in Figure 3(b), which makes the average payoff of individuals within skills' neighborhood lower than the average level of the whole system. Hence, we argue that, while skills facilitate the propagation of cooperative behavior across the network, relying more on history information ($\omega=0.95$) does not contribute to maintenance of cooperation within their neighborhood.

This reasoning is fully supported by the results shown in Figure 4, where we plot the dependence of the frequency of cooperation on both ω and p_s with $b=2$. For $p_s < 0.25$, all $0 < \omega < 1$ induces a higher level of cooperation compared to the average value $f_c=0.4556738$

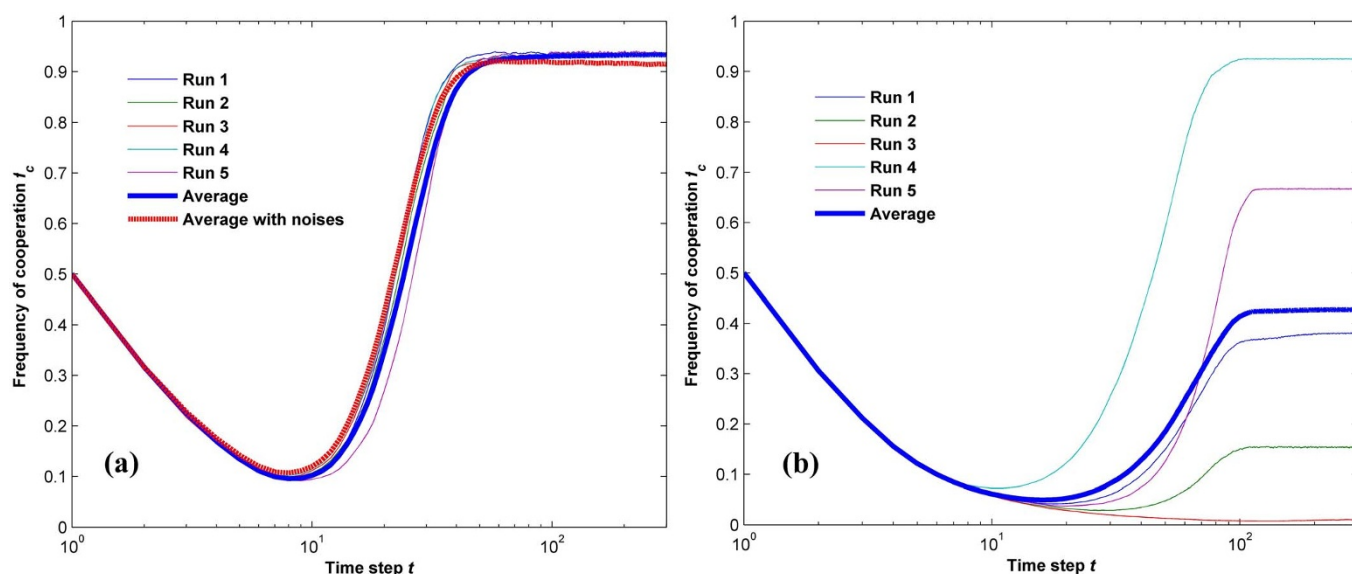


Figure 2 | Evolution of the cooperation level over time steps with and without soft control in 5 random runs for $b=2$. (a): $p_s=0.05$, $\omega=0.95$. (b) $p_s=0$. The blue, thicker trajectories indicate average performance of five simulation runs and the rest represent outcomes of 5 independent runs. The red dotted line in panel (a) shows the average cooperation frequency with noises included in the strategy updating process for skills. It can be observed that without soft control, the outcomes of one-shot run is dependent on randomness with a large value of the standard deviation. On the contrary, soft control not only makes the outcome undisturbed by randomness, but is also robust to noises in the updating process.

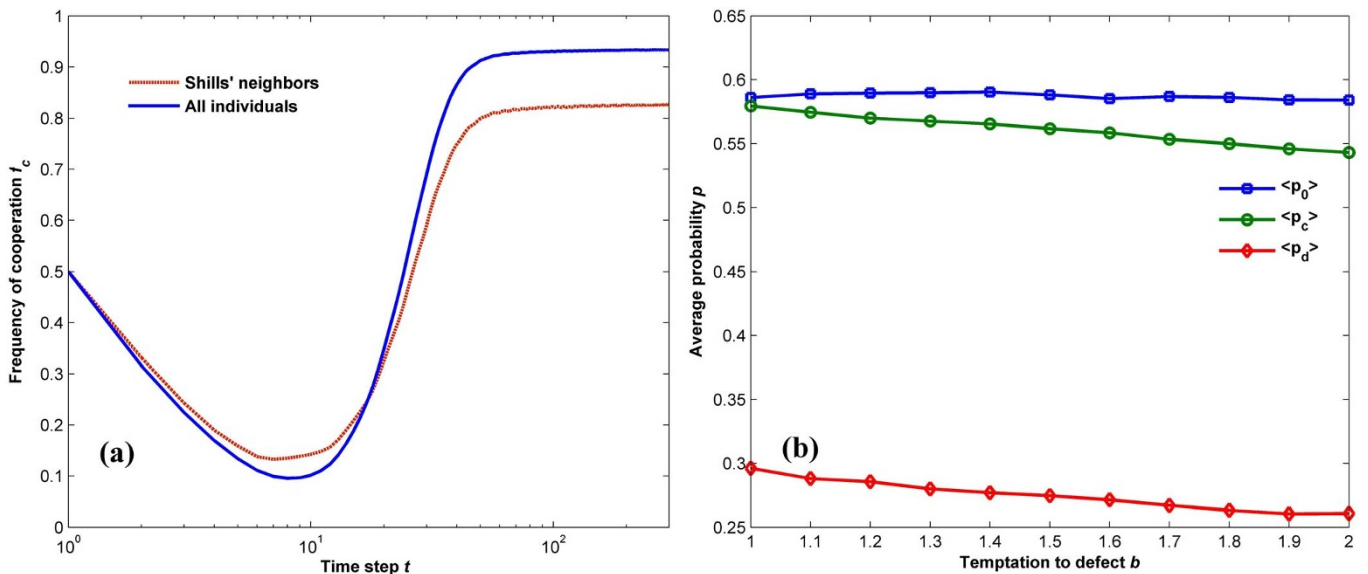


Figure 3 | Evolution characteristics of skills. (a): evolution of the frequency of cooperation for different individuals for $p_s=0.05$, $\omega=0.95$, and $b=2$. The blue line represents the frequency averaged over the whole population, while the red-dotted curve corresponds to the average cooperation of individuals within the neighborhood of skills. (b): average strategy of skills in the equilibrium as a function of the temptation to defect b . While skills promote the propagation of cooperative behavior across the network, relying more on history information ($\omega=0.95$) cannot maintain cooperation within their neighborhood.

achieved by the original population, while the promotion of cooperation is guaranteed only by ω not exceeding a threshold ω_c for $p_s > 0.25$. This should be attributed to the different impacts of ω and p_s on the evolutionary outcome. Firstly, for a fixed value of ω , adding more skills does not contribute to further enhancement of cooperation and there exists a moderate boundary value that warrants the best promotion of cooperation. When p_s is small (e.g., $p_s=0.05$), the cooperation level is enhanced jointly by the few skills and the majority of normal agents, among whom the skills act as pioneers by exploring potentially profitable strategies and the normal agents within the neighborhood of skills spread the successful ones by unconditional imitation. Although more skills being incorporated adds to the strategy diversity of the fundamental population, it however slows down the strategy propagation at the same time, for the proportion of normal agents is reduced. Therefore, it becomes more difficult for the population to reach a consensus without effective diffusion of strategies by enough normal agents. For this reason, as p_s grows from 0.05 to 0.4, it takes the network a much longer time to stabilize and the stationary state shows up as an equi-amplitude oscillation process as shown by the red line in Figure 5(a). However, the amplitude of the oscillation decreases as more neighborhood information is used when $\omega=0$ is chosen. It is also worthwhile to note that our results verify the judgment made by the authors in⁴⁸, as they believe there will be a critical value of skill numbers to achieve the best soft-control goal. Secondly, for a given fraction p_s of skills, smaller values of ω lead to stronger boosts of cooperation in contrast with larger ones as shown both in Figure 4 and Figure 5(b), suggesting that it benefits the population as a whole to learn from others in such a dynamical environment. In particular, $\omega=0$ can always result in the situation of global cooperation, in which $f_c=1$. This is different from the case in⁶, where all players follow the same updating rule of PSO and the most significant benefits are warranted by $\omega=0.99$ in scenarios strongly unfavorable for cooperation. The reason can be found when we compare our model of soft control with the optimization process using PSO. When searching for an optimal solution to a particular problem in the feasible space, each particle is faced with a static environment in terms of the fitness function, i.e., each strategy corresponds to a fixed payoff. However, in the evolutionary game scenario, the payoff each

individual receives depends on both its own strategy as well as the strategy of its opponent. In other words, the best strategy in the history is not necessarily a good choice for the current situation. In such a case, history information plays a less important role in updat-

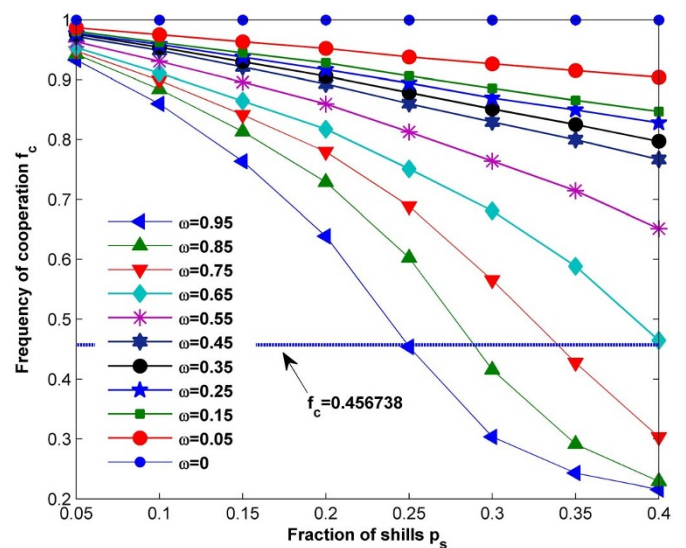


Figure 4 | Dependences of cooperation frequencies on p_s for different values of ω with $b=2$. Each curve with markers shows the frequency of cooperation in the equilibrium as a function of p_s for different ω , and the blue dotted line represents the average cooperation level $f_c=0.456738$, which is achieved by the original population without skills introduced. It can be observed that: (i) for all $0 < \omega < 1$, cooperation cannot be further promoted by adding more skills; (ii) for all $p_s < 0.4$, smaller ω facilitates the promotion of cooperation compared with larger values. We draw the conclusion that assigning higher weights to collective knowledge of the neighborhood swarm is a better choice in strategy updating for cooperation enhancement, whereas simply sticking to one's history memory results in low cooperation level. The results are averaged over 100 independent realizations and the parameter setting is: $L=100$, $T=50$, $G_{\max}=300$.

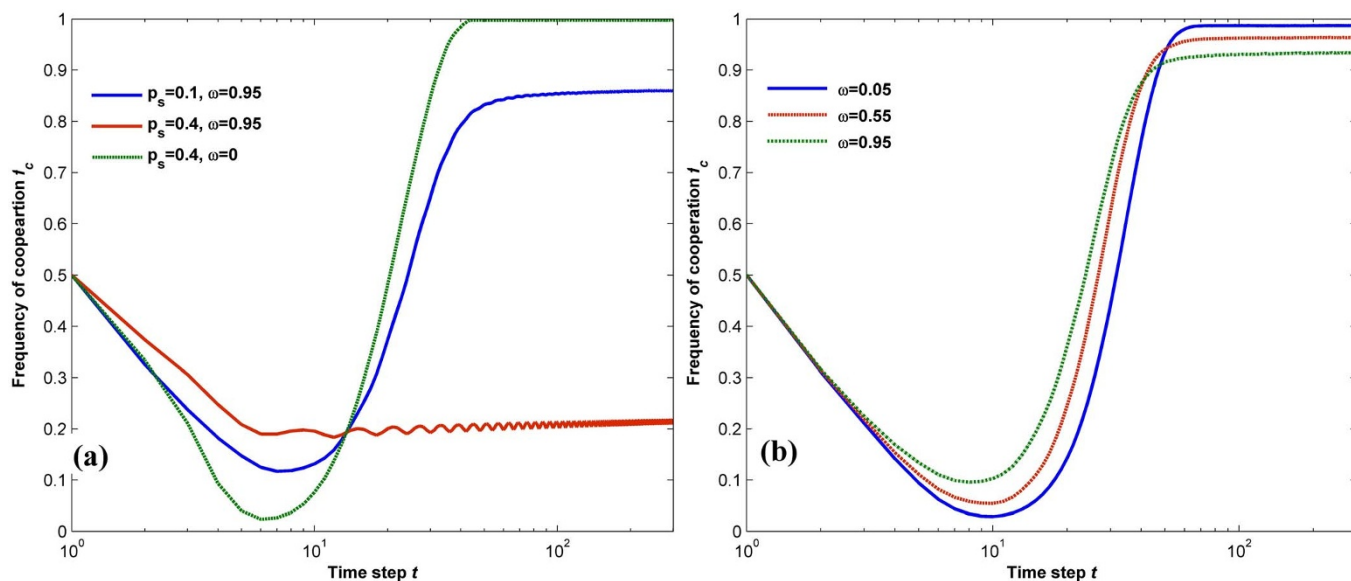


Figure 5 | Evolution of cooperation for different p_s and ω with $b=2$. (a): frequency of cooperation as a function of time steps t with $\omega=0.95$. (b): frequency of cooperation as a function of t for $p_s=0.05$. Results show that the parameter setting of large p_s and ω generates oscillation in the evolution course.

ing strategies of skills towards the highest cooperation promotion of the population. On the other hand, the strategies in the neighborhood provide more useful information. As a consequence, assigning higher weights to the collective knowledge of the neighbors via small ω proves an effective way of inducing cooperation in such defection-prone environment.

Discussion

In this paper, we have investigated the impact of PSO inspired soft control on the evolution of cooperation in a networked IPD game. Following the concept of soft control, we introduce skills into the original population without violating the underlying rules and guide the strategy updating of skills by PSO mechanisms, which make use of history information and the collective wisdom gained by the swarm to search for the most profitable strategy. Through intensive simulations, we demonstrate that the cooperation level can be controlled to particular values by selecting control parameters p_s and ω appropriately. Specifically, we have shown that, on the one hand, it does not contribute to further cooperation promotion to add more skills but suppresses the propagation of cooperative behavior instead. Hence, we draw the conclusion that there exists an optimal boundary value guaranteeing the best promotion of cooperation for such an evolutionary network. On the other hand, we have found that relying more on the collective knowledge of the neighbors during strategy updating of skills always results in a stronger promotion of cooperation, while blindly sticking to one's history information hinders the emergence and sustainability of cooperation in the population. Besides, it is shown that the incorporation of skills enables the evolutionary outcome to be less dependent of random factors of the evolution.

As the first step to introduce swarm intelligence to soft control in a spatial evolutionary game, our research sheds some light on the role of PSO inspired skills in evolution dynamics of cooperation and provides a useful tool to intervene in collective behavior of self organized individuals, with the purpose of promoting cooperative behavior to a desirable level. Furthermore, this study can also be used to interpret the emergence of cooperation in a structured population of unrelated individuals, since special individuals acting like skills in soft control have been widely observed in natural systems. However, much work related to soft control remains to be carried out and our future work will focus on the following three aspects. In the first

place, it is interesting to introduce other swarm intelligence mechanisms to soft control and study their impacts on cooperation evolution, such as ACO and BSO. Additionally, we will extend the networked structure to more realistic models and particularly study the dependence of the evolution of cooperation on the distribution of skills. Finally, we will conduct research on algebraic formulation and global dynamics analysis of evolutionary game with soft control.

Methods

For the calculations, we consider an evolutionary T -stage IPD game located on an $L \times L$ square lattice with periodic boundary conditions, where each player occupies a site of the graph and interacts with its four nearest neighbors. Following a large portion of literature, we parameterize the payoff matrix of the one-shot PD by

$$\begin{pmatrix} R & S \\ T & P \end{pmatrix} = \begin{pmatrix} 1-b & \\ & 1+b \end{pmatrix} \quad (2)$$

with b being the only free parameter to control the dilemma strength. Thus, the dilemma grows weaker as b approaches 0 and stronger conflict arises between individual interests and social welfare when b increases towards infinite. Note that b must be greater than 0 to conform to the definition of the PD game. Two types of individuals are involved in our model, namely normal individuals in the original population and skills added additionally for cooperation promotion, whose fractions are denoted by p_n and p_s respectively, satisfying $p_n + p_s = 1$. The initial distribution of skills is generated randomly and remains unchanged for the rest evolution duration.

Initially, each player is randomly assigned a stationary Markov decision strategy $x = (p_0, p_c, p_d)^{86,60}$, each element of which is drawn uniformly from the region between 0 and 1. Specifically, p_0 is the probability for an agent to offer cooperation at the first stage, p_c and p_d denote respectively the conditional probability of an agent to cooperate when the its opponent cooperates and defects in the previous encounter respectively. It is worth mentioning that for each individual, the strategy does not change with time during each time step. The evolution is carried out by implementing a synchronous update process: at each time step t , every individual i accumulates its payoff $\Pi_i(t)$ by means of pairwise interaction with all the neighbors in its neighborhood U_i

$$\Pi_i(t) = \sum_{j \in U_i} \pi_{i,j}(t) \quad (3)$$

where $\pi_{i,j}(t)$ is the instantaneous payoff of individual i playing a T -stage IPD game against its neighbor j ; following this, each individual will simultaneously attempts to adapt its strategy with the purpose of maximizing its potential probability of success in future generations. Within the framework of soft control, skills are treated equally as normal individuals by complying to the basic playing rules for consideration of avoiding deception and exploitation by normal agents⁴⁸. Only in one certain aspect are skills different from normal agents, i.e., skills are allowed to adopt elaborated strategies and updating rules, which is of the paramount importance for desirable collective behavior induction. Based on this, we assume that normal individuals follow the updating process of *unconditional imitation*²⁷, where the best



strategy in a normal individual's neighborhood is selected as the strategy at the next time step $t + 1$

$$x_i(t+1) = x_{j^*}(t)$$

$$j^* = \begin{cases} \arg \max_{j \in U^*(i)} \Pi_j, & \text{if } |\arg \max_{j \in U^*(i)} \Pi_j| = 1 \\ \min\{\mu | \mu \in \arg \max_{j \in U^*(i)} \Pi_j\}, & \text{else} \end{cases} \quad (4)$$

where set $U^*(i)$ includes individual i besides its neighbors and $|A|$ denotes the element number of set A . On the contrary, shills follow a more complex updating mechanism, which is inspired by swarm intelligence emerging from bird flocks searching for food sources: each shill is considered to be a particle moving in a multi-dimensional space $R[0,1]^3$, interacting with its neighbors, and adopting its strategy by combining some aspects of its memory as well as heuristic information from its neighbors. Specifically, each shill updates its strategy $x_i(t)$ following

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (5)$$

$$v_i(t+1) = v_i(t) + \omega(x_i^{best}(t) - x_i(t)) + (1-\omega)(x_{U_i}^{best}(t) - x_i(t)) \quad (6)$$

where $v_i(t+1)$ is the velocity vector of agent i at time step $t + 1$, $x_i^{best}(t)$ is the best-ever strategy of agent i throughout its history, and $x_{U_i}^{best}(t)$ denotes the best strategy in its neighborhood at time step t . In our model, the initial velocity vector of each shill is randomly generated in the continuous 3-dimensional space $R[-1,1]^3$. The weighting coefficient ω is utilized to keep a balance between one's own past information and swarm wisdom of the neighborhood. According to equations (5) and (6), a shill tends to stick to its own memory and make decisions based on its best ever strategy when ω is close to 1. Nonetheless, it takes into more consideration the swarm intelligence of its neighbors as ω approaches 0. In particular, a shill copies its best action in the past when $\omega = 1$ and imitates the current best strategy in the neighborhood with the highest performance Π when $\omega = 0$.

In the Results section, the proportion of shills p_s , the game strength b , and the weighting parameter ω are considered as crucial parameters in investigating the evolution of cooperation under the framework of soft control.

- Nowak, M. A. Evolving cooperation. *J. Theor. Biol.* **299**, 1–8 (2012).
- Zhang, C., Gao, J. X., Cai, Y. Z. & Xu, X. M. Evolutionary prisoner's dilemma game in flocks. *Physica A* **390**, 50–56 (2011).
- Smith, J. M. The importance of the nervous system in the evolution of animal flight. *Evolution* **6**, 127–129 (1952).
- Smith, J. M. & Price, G. R. The logic of animal conflict. *Nature* **246**, 15–18 (1973).
- Smith, J. M. The theory of games and the evolution of animal conflicts. *J. Theor. Biol.* **47**, 209–221 (1974).
- Zhang, J. L., Zhang, C. Y., Chu, T. G. & Perc, M. Resolution of the stochastic strategy spatial prisoner's dilemma by means of particle swarm optimization. *PLoS ONE* **6**, e21787 (2011).
- Stewart, A. J. & Plotkin, J. B. From extortion to generosity, evolution in the Iterated Prisoner's Dilemma. *Proc. Natl. Acad. Sci. USA* **110**, 15348–15353 (2013).
- Li, P. & Duan, H. Robustness of cooperation on scale-free networks in the evolutionary prisoner's dilemma game. *EPL* **105**, 48003 (2014).
- Nowak, M. A. Five rules for the evolution of cooperation. *Science* **314**, 1560–1563 (2006).
- Axelrod, R. & Hamilton, W. D. The evolution of cooperation. *Science* **211**, 1390–1396 (1981).
- Axelrod, R. (Chapter 1: Evolving New Strategies) [Axelrod R. (ed.)] [10–19]. *The Complexity of Cooperation*. (Princeton Univ. Press, 1997).
- Nowak, M. A. & Sigmund, K. Tit for tat in heterogeneous populations. *Nature* **355**, 250–253 (1992).
- Milinski, M. Tit for tat in sticklebacks and the evolution of cooperation. *Nature* **325**, 433–435 (1987).
- Segal, U. & Sobel, J. Tit for tat: foundations of preferences for reciprocity in strategic settings. *J. Econ. Theor.* **136**, 197–216 (2007).
- Nowak, M. A. & Sigmund, K. Chaos and the evolution of cooperation. *Proc. Natl. Acad. Sci. USA* **90**, 5091–5094 (1993).
- Rapoport, A. & Chammah, A. M. (Chapter 8: Equilibrium Models with Adjustable Parameters) [Rapoport A. & Chammah A. M. (ed.)] [129–134]. *Prisoner's Dilemma*. (Michigan Univ. Press, 1965).
- Martinez-Vaquero, L. A., Cuesta, J. A. & Sanchez, A. Generosity pays in the presence of direct reciprocity: a comprehensive study of 2×2 repeated games. *PLoS ONE* **7**, e35135 (2012).
- Wang, Z., Szolnoki, A. & Perc, M. If players are sparse social dilemmas are too: Importance of percolation for evolution of cooperation. *Sci. Rep.* **2**, 369 (2012).
- Wang, Z., Szolnoki, A. & Perc, M. Percolation threshold determines the optimal population density for public cooperation. *Phys. Rev. E* **85**, 037101 (2012).
- Wang, Z., Szolnoki, A. & Perc, M. Evolution of public cooperation on interdependent networks: The impact of biased utility functions. *EPL* **97**, 48001 (2012).
- Wang, Z., Szolnoki, A. & Perc, M. Interdependent network reciprocity in evolutionary games. *Sci. Rep.* **3**, 1183 (2013).
- Jin, Q., Wang, L., Xia, C. Y. & Wang, Z. Spontaneous symmetry breaking in interdependent networked game. *Sci. Rep.* **4**, 4095 (2014).
- Wang, Z., Szolnoki, A. & Perc, M. Optimal interdependence between networks for the evolution of cooperation. *Sci. Rep.* **3**, 2470 (2013).
- Gómez-Gardeñes, J., Reinares, I., Arenas, A. & Floría, L. M. Evolution of cooperation in multiplex networks. *Sci. Rep.* **2**, 620 (2012).
- Hauert, C. & Doebeli, M. Spatial structure often inhibits the evolution of cooperation in the snowdrift game. *Nature* **428**, 643–646 (2004).
- Gracia-Lázaro, C. *et al.* Heterogeneous networks do not promote cooperation when humans play a prisoner's dilemma. *Proc. Natl. Acad. Sci. USA* **109**, 12922–12926 (2012).
- Nowak, M. A. & May, R. M. Evolutionary games and spatial chaos. *Nature* **359**, 826–829 (1992).
- Zhang, Y., Fu, F., Wu, T., Xie, G. & Wang, L. A tale of two contribution mechanisms for nonlinear public goods. *Sci. Rep.* **3**, 2021 (2013).
- Wu, T., Fu, F., Zhang, Y. & Wang, L. Adaptive role switching promotes fairness in networked ultimatum game. *Sci. Rep.* **3**, 1550 (2013).
- Perc, M. & Wang, Z. Heterogeneous Aspirations Promote Cooperation in the Prisoner's Dilemma Game. *PLoS ONE* **5**, e15117 (2010).
- Perc, M. Coherence resonance in a spatial prisoner's dilemma game. *New. J. Phys.* **8**, 22 (2006).
- Szolnoki, A., Perc, M., Szabó, G. & Stark, H. U. Impact of aging on the evolution of cooperation in the spatial prisoner's dilemma game. *Phys. Rev. E* **80**, 021901 (2009).
- Wang, Z., Murks, A., Du, W. B., Rong, Z. H. & Perc, M. Coveting thy neighbors fitness as a means to resolve social dilemmas. *J. Theor. Biol.* **277**, 19–26 (2011).
- Wang, Z., Szolnoki, A. & Perc, M. Rewarding evolutionary fitness with links between populations promotes cooperation. *J. Theor. Biol.* **349**, 50–56 (2014).
- Li, Q., Chen, M., Perc, M., Iqbal, A. & Abbott, D. Effects of adaptive degrees of trust on coevolution of quantum strategies on scale-free networks. *Sci. Rep.* **3**, 2949 (2013).
- Wu, T., Fu, F. & Wang, L. Moving away from nasty encounters enhances cooperation in ecological prisoner's dilemma game. *PLoS ONE* **6**, e27669 (2011).
- Hauert, C., Traulsen, A., Brandt, H., Nowak, M. A. & Sigmund, K. Via freedom to coercion: the emergence of costly punishment. *Science* **316**, 1905–1907 (2007).
- Mathew, S. & Boyd, R. When does optional participation allow the evolution of cooperation? *Proc. R. Soc. B* **276**, 1167–1174 (2009).
- Grujić, J., Eke, B., Cabrales, A., Cuesta, J. A. & Sánchez, A. Three is a crowd in iterated prisoner's dilemmas: experimental evidence on reciprocal behavior. *Sci. Rep.* **2**, 638 (2012).
- Wang, Z., Xia, C., Meloni, S., Zhou, C. & Moreno, Y. Impact of social punishment on cooperative behavior in complex networks. *Sci. Rep.* **3**, 3055 (2013).
- Szolnoki, A. & Perc, M. Reward and cooperation in the spatial public goods game. *EPL* **92**, 38003 (2010).
- Suzuki, S. & Kimura, H. Indirect reciprocity is sensitive to costs of information transfer. *Sci. Rep.* **3**, 1435 (2013).
- Gao, Y., Chen, Y. & Liu, K. Cooperation stimulation for multiuser cooperative communications using indirect reciprocity game. *IEEE Trans. Commun.* **60**, 3650–3661 (2012).
- Brafman, R. I. & Tennenholtz, M. On practically controlled multi-agent systems. *J. Artif. Intell. Res.* **4**, 477–507 (1996).
- Chen, T., Liu, X. & Lu, W. Pinning complex networks by a single controller. *IEEE Trans. Circuits-I* **54**, 1317–1326 (2007).
- Yu, W., Chen, G. & Lu, J. On pinning synchronization of complex dynamical networks. *Automatica* **45**, 429–435 (2009).
- Han, J., Li, M. & Guo, L. Soft control on collective behavior of a group of autonomous agents by a shill agent. *J. Syst. Sci. Complexity* **19**, 54–62 (2006).
- Wang, X., Han, J. & Han, H. W. Special agents can promote cooperation in the population. *PLoS ONE* **6**, e29182 (2011).
- Benoit, J. P. & Krishna, V. Finitely repeated games. *Econometrica* **17**, 317–320 (1985).
- Vicsek, T., Czirók, A., Ben-Jacob, E., Cohen, I. & Shochet, O. Novel type of phase transition in a system of self-driven particles. *Phys. Rev. Lett.* **75**, 1226 (1995).
- Duan, H. B., Shao, S., Su, B. W. & Zhang, L. New development thoughts on the bio-inspired intelligence based control for unmanned combat aerial vehicle. *SCI. CHINA Technol. Sc.* **53**, 2025–2031 (2010).
- Kennedy, J. & Eberhart, R. C. Particle swarm optimization. *Proc IEEE Int Conf Neural Networks, Perth, Australia* **4**, 1942–1948 (1995).
- Duan, H. B., Luo, Q. N., Ma, G. J. & Shi, Y. H. Hybrid particle swarm optimization and genetic algorithm for multi-UAVs formation reconfiguration. *IEEE Comput. Intell. Mag.* **8**, 16–27 (2013).
- Dorigo, M., Maniezzo, V. & Colnori, A. The ant system: optimization by a colony of cooperating agents. *IEEE Trans. Syst. Man. Cy.* **B** **26**, 29–41 (1996).
- Duan, H. B., Zhang, X. Y., Wu, J. & Ma, G. J. Max-min adaptive ant colony optimization approach to multi-UAVs coordinated trajectory re-planning in dynamic and uncertain environments. *J. Bionic. Eng.* **6**, 161–173 (2009).
- Shi, Y. H. An optimization algorithm based on brainstorming process. *Int. J. Swarm. Intell. Res.* **2**, 35–62 (2011).
- Sun, C. H., Duan, H. B. & Shi, Y. H. Optimal satellite formation reconfiguration based on closed-loop brain storm optimization. *IEEE Comput. Intell. Mag.* **8**, 39–51 (2013).



58. Duan, H. B., Li, S. T. & Shi, Y. H. Predator-prey based brain storm optimization for DC brushless motor. *IEEE Trans. Magn.* **49**, 5336–5340 (2013).
59. Brede, M. Short versus long term benefits and the evolution of cooperation in the prisoner's dilemma game. *PLoS ONE* **8**, e56016 (2013).
60. Miyaji, K., Tanimoto, J., Wang, Z., Hagishima, A. & Ikegaya, N. Direct reciprocity in spatial populations enhances R-reciprocity as well as ST-reciprocity. *PLoS ONE* **8**, e71961 (2013).

Acknowledgments

This work was partially supported by Natural Science Foundation of China (NSFC) under grant #61333004, #61273054 and #60975072, National Key Basic Research Program of China (973 Project) under grant #2014CB046401 and #2013CB035503, National Magnetic Confinement Fusion Research Program of China under grant # 2012GB102006, Top-Notch Young Talents Program of China, Program for New Century Excellent Talents in University of China under grant #NCET-10-0021, and Aeronautical Foundation of China under grant #20135851042.

Author contributions

All authors have made contributions to this paper. H.D. and C.S. designed the whole work, produced all the data, and wrote the paper. H.D. supervised the whole work and contributed to the manuscript preparation. C.S. made simulations. All authors discussed the results and reviewed the manuscript.

Additional information

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Duan, H.B. & Sun, C.H. Swarm intelligence inspired skills and the evolution of cooperation. *Sci. Rep.* **4**, 5210; DOI:10.1038/srep05210 (2014).



This work is licensed under a Creative Commons Attribution 3.0 Unported License. The images in this article are included in the article's Creative Commons license, unless indicated otherwise in the image credit; if the image is not included under the Creative Commons license, users will need to obtain permission from the license holder in order to reproduce the image. To view a copy of this license, visit <http://creativecommons.org/licenses/by/3.0/>