Chaos, Solitons and Fractals 151 (2021) 111269

Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

Nonlinear Science, and Nonequilibrium and Complex Phenomena

journal homepage: www.elsevier.com/locate/chaos

Effects of dynamic-Win-Stay-Lose-Learn model with voluntary participation in social dilemma

Zhenyu Shi [a,b,c,d], Wei Wei [a,b,c,d,*], Xiangnan Feng [a,b,c,d], Ruizhi Zhang [a,b,c,d], Zhiming Zheng [a,b,c,d]

[a] School of Mathematical Sciences, Beihang University, Beijing, China
[b] Key Laboratory of Mathematics Informatics Behavioral Semantics, Ministry of Education, China
[c] Beijing Advanced Innovation Center for Big Data and Brain Computing, Beihang
[d] Peng Cheng Laboratory, Shenzhen, Guangdong, China

ARTICLE INFO

Article history:
Received 18 March 2021
Revised 1 July 2021
Accepted 9 July 2021

Keywords:
Prisoner's dilemma game
Dynamic aspiration
Cooperation
Spatial evolutionary game

ABSTRACT

In recent years, Win-Stay-Lose-Learn rule has attracted wide attention as an effective strategy updating rule, and voluntary participation is proposed by introducing a third strategy in Prisoner's dilemma game. Some researches show that combining Win-Stay-Lose-Learn rule with voluntary participation could promote cooperation more significantly under moderate temptation values, however, cooperators' survival under high aspiration levels and high temptation values is still a challenging problem. In this paper, inspired by Achievement Motivation Theory, a Dynamic-Win-Stay-Lose-Learn rule with voluntary participation is investigated, where a dynamic aspiration process is introduced to describe the co-evolution of individuals' strategies and aspirations. It is found that cooperation is extremely promoted and defection is almost extinct in our model, even when the initial aspiration levels and temptation values are high. The combination of dynamic aspiration and voluntary participation plays an active role since loners could survive under high initial aspiration levels and they will expand stably because of their fixed payoffs. The robustness of our model is also discussed and some adverse structures are found which should be alerted in the evolutionary process. Our work provides a more rational model and shows that cooperators may prevail defectors in an unfavorable initial environment.

1. Introduction

Though cooperative behavior is ubiquitous in biological, economic and social systems [1–3], how to explain its emergence and stability is still a valuable and challengeable problem in related fields [4]. Prisoner's dilemma (PD) game is a representative model to describe the social dilemmas among selfish individuals. In a typical PD game, two players choose cooperation or defection simultaneously without communication. One will always get a higher payoff if it chooses defection whichever its opponent chooses, but if both of them choose defection, their payoffs are lower than those of both choosing cooperation, which leads to a conflict between individual rationality and collective rationality. Evolutionary game theory has been thought of as a powerful mathematical framework to reveal the mechanisms for cooperation phenomenon in the competitive environment. In [5], Nowak found five typical mechanisms which are conducive to the existence of cooperation: kin selection, direct and indirect reciprocity, network reciprocity, and group selection, among which network reciprocity has received the most widespread attention. Series of researches focus on different kinds of network structures, such as lattice [6,7], small world networks [8,9] and scale-free networks [10,11]. Besides, many social mechanisms have been proved to enhance cooperation, such as punishment [12–15], game organizers [16,17], compassion [18], memory effects [19,20], and so on.

In recent years, aspiration-based strategy updating rules have got more and more attention from researchers [21–29]. A representative model is that compared to maximize their payoffs, individuals usually tend to keep their strategies when they feel satisfied, otherwise they try to learn what others do, which is called Win-Stay-Lose-Learn strategy updating rule and there have been some related studies in recent years [30–32]. In most of the above research, aspiration is a fixed value which is set for all individuals before the evolution process begins [21–24,30–32]. In fact, dynamic aspiration models meet the actual situation better and there are also some related researches [25–29]. But the mechanism how

* Corresponding author at: School of Mathematical Sciences, Beihang University, Beijing, China.
E-mail address: weiw@buaa.edu.cn (W. Wei).

https://doi.org/10.1016/j.chaos.2021.111269
0960-0779/

Table 1
Payoff matrix of OPD games.

	$\mathcal{C}$	$\mathcal{D}$	$\mathcal{L}$
$\mathcal{C}$	R	S	l
$\mathcal{D}$	T	P	l
$\mathcal{L}$	l	l	l

dynamic aspiration models impact the evolution process is still to be resolved. Our work focuses on what role dynamic aspiration models play to promote cooperation, which is compared to fixed aspiration model.

Besides cooperation and defection, there are still many other behaviors that could be observed in social dilemmas, such as tit-for-tat [33], punishment [34] and voluntary participation [35]. Among these, voluntary participation has been most widely introduced into the research of the PD games in recent years due to the universality of abstention in social dilemmas [35], which means an individual may choose to abstain the PD game and get a low but guaranteed payoff. It appropriately describes the phenomenon that some individuals may choose to reduce interactions in social dilemmas, such as a company will decide not to trade due to lack of trust in other companies [36]. Because of its spontaneity and spreadability, it is reasonable and necessary to introduce it as a third strategy besides of cooperation and defection.

Although Win-Stay-Lose-Learn rule and voluntary participation could promote cooperation under moderate temptation values, cooperators' survival under high aspiration levels and high temptation values is still a challenging problem. In this paper, combining both of their advantages, a Dynamic-Win-Stay-Lose-Learn rule is proposed in the Optional Prisoner's dilemma game. The model aims at promoting cooperation obviously with the appropriate initial structure, even when the temptation value T is large.

In the remainder of our paper, firstly we introduce our Dynamic-Win-Stay-Lose-Learn strategy updating rule with voluntary participation in the *Model* section. Then the main valuable phenomena is shown based on the results of Monte Carlo simulation, and more detailed analysis about why loners could promote cooperation under dynamic aspiration model and what are the necessary conditions for promoting cooperation is discussed, which is divide into two different parameter regions in the *Results* section. Finally the main conclusion and innovation of our work is summarized in the *Conclusion* section.

2. Model

Our work considers PD games with voluntary participation in which three strategies are included: cooperation($\mathcal{C}$), defection($\mathcal{D}$) and lone($\mathcal{L}$), which is also called Optional Prisoner's Dilemma(OPD) game. Individuals distribute in an $L \times L$ square lattice with periodic boundary conditions, in this paper, L is set to 100, and each of them will only interact with its four direct neighbors. The strategy and aspiration of an individual i are denoted as s_i and A_i respectively. In the evolutionary process, all individuals update their strategies and aspirations synchronously by discrete time steps. One complete step performs according to following rules:

(a) **Rule of game:** Each individual i plays OPD games with its four direct neighbors to get a payoff $P_i = \sum_{j \in \Omega_i} P_{ij}$, where Ω_i represents all direct neighbors of individual i. P_{ij} represents i's payoff for playing an OPD game with j, which could be got by Table 1. They will receive the reward R or punishment P if they both choose to cooperate or defect. If one of them chooses $\mathcal{C}$ but the other one chooses $\mathcal{D}$, the former will receive the sucker's payoff S and the the latter will receive the temptation value T. If at least one of them chooses $\mathcal{L}$, both of them will get the loner's payoff l. In OPD games, $S < P < R < T$ and $0 < l < 1$ should be meet.

In this paper, parameters are set as boundary game: $R = 1$, $P = S = 0$ and $T = b$ [6,30]. And l is set to 0.3 as what [31] does.

(b) **Rule of strategy's update:** Each individual i compares P_i with A_i. If $P_i \geq A_i$, i will keep its strategy. If $P_i < A_i$, i will select one of its direct neighbors j at random and imitate j's strategy with the Fermi updating rule:

$$W_{ij} = \frac{1}{1 + \exp[(P_i - P_j)/K]}, \tag{1}$$

where K represents the amplitude of noise [37] and is set to 0.1 in our model to characterize appropriate randomness [38,39].

(c) **Rule of aspiration's update:** Each individual i updates its aspiration according to the difference between P_i with A_i with the linear updating rule:

$$A_i(t+1) = A_i(t) + a * (P_i(t) - A_i(t)), \tag{2}$$

where $A_i(0) = A$ is the initial aspiration for all individuals. The dynamic-aspiration process is represented by a linear updating rule, where $a \in [0, 1]$ quantifies the evolution rate of aspiration [29]. When $a = 0$, the model is reduced to the fixed aspiration model. The upper bound of $a = 1$ ensures that one's aspiration will be close to but not over payoff. In general, an individual's aspiration might not be updated drastically, so a should set to a small value. In this paper, we set $a = 0.05$.

To ensure the network to be stable, above step will carry out 100,000 times repeatedly in a simulation. The final fractions of the three strategies denoting as $r_{\mathcal{C}}$, $r_{\mathcal{D}}$ and $r_{\mathcal{L}}$ are calculated by the average of the last 1,000 steps. For each pair of parameters, 20 independent simulations are performed to make the results more accurate.

3. Results

3.1. Overview

To begin our discussion, the result of randomly initialized network is shown, where all of the three strategies occupy one third of the network with random settings. The fractions of three strategies are denoted as $r_{\mathcal{C}0}$, $r_{\mathcal{D}0}$ and $r_{\mathcal{L}0}$. Besides, for all individuals, the initial aspiration A is of the same value. Fig. 1 shows $r_{\mathcal{C}}$, $r_{\mathcal{L}}$ and $r_{\mathcal{D}}$ as a function depending on b for different values of A. It could be found that there are significant differences between the results for different values of A. Two different phases could be easily observed, which are called *Stable Coexistence* and *Defection Suppression* respectively. When $A \leq 1.2$, cooperators, defectors and loners could coexist with moderate fractions respectively. For instance, $r_{\mathcal{C}} = r_{\mathcal{L}} = r_{\mathcal{D}} = 0.33$ when $A = 0$, $r_{\mathcal{C}} = 0.29$, $r_{\mathcal{L}} = 0.44$ and $r_{\mathcal{D}} = 0.27$ when $A = 0.8$, which are independent with the value of b. When $A > 1.2$, it can be observed that $r_{\mathcal{C}}$ decreases monotonously with the increase of the value of b and $r_{\mathcal{L}}$ is the opposite. More importantly, $r_{\mathcal{D}}$ always keeps a low level which is related to the value of A, even when $b > 1.8$ where the temptation value is so high that defectors' expansion is almost inevitable in many other models [29,31,40–42]. Compared to the results shown in [29] and [31], dynamic aspirations model with voluntary participation plays an important role to promote cooperation.

In [29], due to lack of participation of loners, defectors could occupy the whole network under most regions of parameters. When $A \leq 1.0$, cooperators and defectors could coexist. Since all individuals will get satisfied fast, the stability of the result is predictable and could be calculated theoretically, which is similar to our results. When $A > 1.0$, defectors will always occupy the whole network except when $1.0 < A \leq 2.0$ and $b \geq A$. The authors show that under the dynamic aspiration model, there are some typical

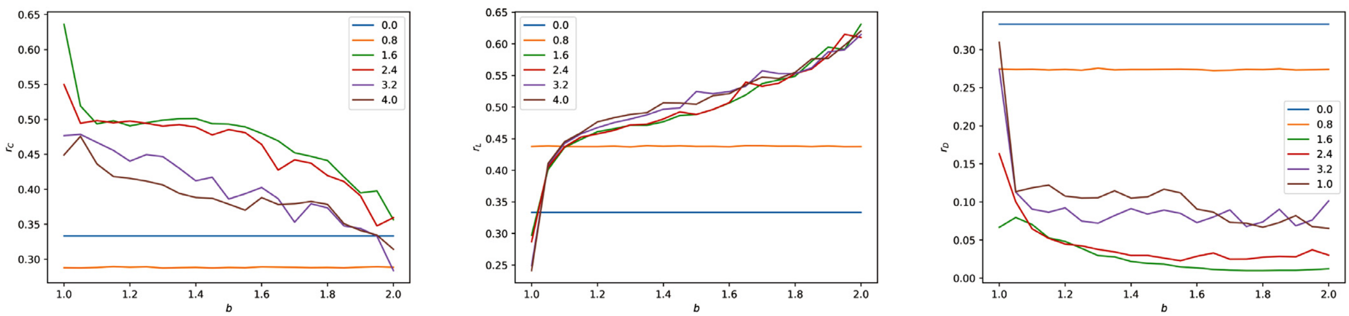

Fig. 1. Average fractions of cooperators, loners and defectors in the stable state in dependence on b at different values of the A, from left to right respectively.

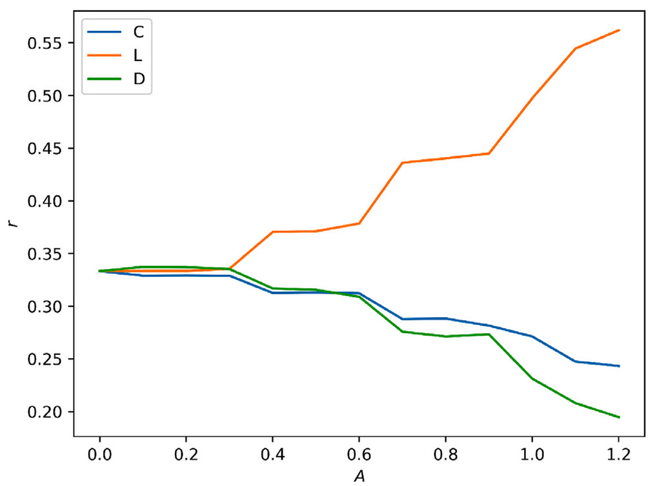

Fig. 2. Average fractions of cooperators, loners and defectors in the stable state in dependence on A when $b = 1.6$.

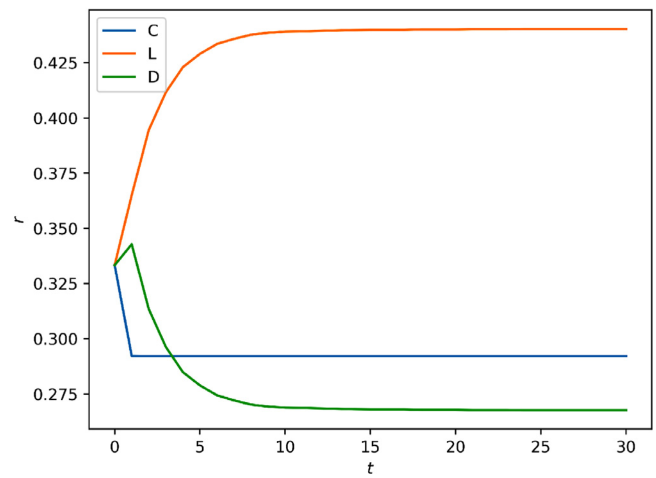

Fig. 3. Average fractions of cooperators, loners and defectors as a function of step t when $A = 0.8$ and $b = 1.9$.

initial local structures which could lead to expansion of defection under different regions of parameters. Individuals that evolve into cooperators and defectors repeatedly are called *Infectors* and their neighbors which have the same potential attributes are called *Infected nodes*. Their evolution process will eventually spread to the so called *High-risk cooperators* and cause chain phenomenon. However, in our model, the above process won't happen because of the participation of loners. Loners could suppress defectors' expansion when $A > 1.0$ for any value of b so that cooperation could be also promoted.

In [31], due to the fixed aspiration model, loners could play the role in suppressing defection only when $A \leq 1.2$ (assuming $l = 0.3$). When $A > 1.2$, loners are always dissatisfied so they are difficult to survive. As a result, the fraction of defectors couldn't be decreased to a quite low level. But in our dynamic aspiration model, even if $A > 1.2$, individuals' aspirations will decrease if they get dissatisfied, which provides a favorable environment for the survival of the loners. From above we know that the combination of dynamic aspiration and voluntary participation plays an active role since loners could survive under high initial aspiration levels and they will expand stably because of their fixed payoffs. In the following we will discuss more details according to different regions of parameters.

Besides, we discuss the result with initial settings that loners occupy a relatively small percentage of the network in Appendix A and result on the scale-free network is shown in Appendix B.

3.2. Stable coexistence ($A \leq 1.2$)

For small values of A, all of three strategies could survive with moderate fractions. Fig. 2 presents the fraction $r_{\mathcal{C}}$ of cooperators, the fraction $r_{\mathcal{L}}$ of loners and the fraction $r_{\mathcal{D}}$ of defectors when stable as a function of A when $b = 1.6$. It is shown that with A increasing, $r_{\mathcal{L}}$ increases but $r_{\mathcal{C}}$ and $r_{\mathcal{D}}$ decrease. Besides, there are some apparent discontinuous transitions which are A =0.3, 0.6, 0.9 and 1.0. These transitions can be explained as follows. In our model, the payoff of a loner is always 1.2 no matter what strategies its neighbors have. When $A \leq 1.2$, a loner's payoff is higher than its aspiration so it is always satisfied and never changes its strategy. As for a cooperator or a defector, their payoff can be written as $n_{\mathcal{C}} + 0.3n_{\mathcal{L}}$ and $bn_{\mathcal{C}} + 0.3n_{\mathcal{L}}$ respectively, where $n_{\mathcal{C}}$, $n_{\mathcal{L}}$ and $n_{\mathcal{D}}$ represent the number of a node's $\mathcal{C}$ neighbors, $\mathcal{L}$ neighbors and $\mathcal{D}$ neighbors and they should meet $n_{\mathcal{C}} + n_{\mathcal{L}} + n_{\mathcal{D}} = 4$. There might be three different states:

- When $n_{\mathcal{C}} > 1$, the node's payoff is higher than 2.0 and always satisfied.
- When $n_{\mathcal{C}} = 1$, the node might be dissatisfied only when $n_{\mathcal{L}} = 0$. Under this state, the payoffs of a cooperator and a defector are 1.0 and b respectively.
- When $n_{\mathcal{C}} = 0$, the node's payoff can only be 0, 0.3, 0.6, 0.9 or 1.2.

As mentioned above, the possible value of a node's payoff is in {0.3, 0.6, 0.9, 1.0}, which are consistent with the points that discontinuous transitions happen. When a cooperator or a defector is dissatisfied, it will change its strategy by imitating its neighbors. Furthermore, it might become satisfied only when it evolves into a loner. On the contrary, a loner will never change its strategy. So it is observed that $r_{\mathcal{L}}$ will be higher than 0.33 (the initial fraction of $\mathcal{L}$) when stable. The higher A is, the higher $r_{\mathcal{L}}$ will be

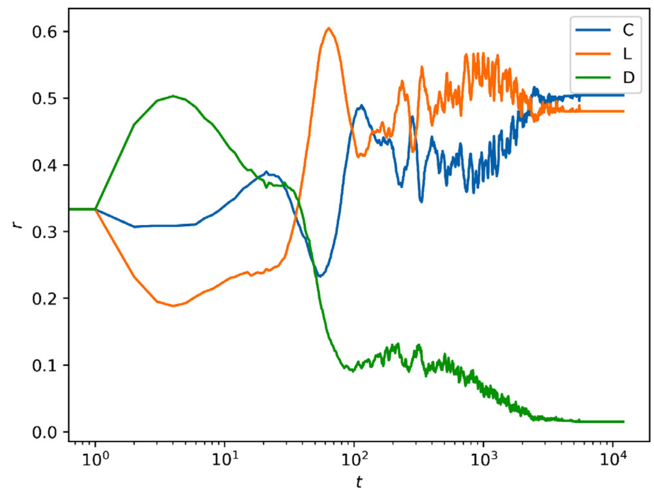

Fig. 4. Average fractions of cooperators, loners and defectors as a function of step t when $A = 1.6$ and $b = 1.6$.

when stable because more cooperators and defectors will be dissatisfied and evolve into loners finally, which is independent with the value of b. Fig. 3 presents $r_{\mathcal{C}}$, $r_{\mathcal{L}}$ and $r_{\mathcal{D}}$ with the time-evolution when $A = 0.8$ and $b = 1.9$. Some cooperators and defectors evolve into loners quickly then the network is stable even though under high value of b. Loners play an important role that they will never change their strategies if they are satisfied initially.

Besides, the theoretically values of $r_{\mathcal{C}}$ for different A could be calculated by the formula:

$$r_{\mathcal{C}} = r_{\mathcal{C}0} * (1 - \sum_{n=0}^{k} C_4^k r_{\mathcal{L}0}^k r_{\mathcal{D}0}^{4-k}), \tag{3}$$

where $k = \lceil A/0.3 \rceil - 1$.

For instance, when $A = 0.3$ ($k = 0$), a cooperator will be dissatisfied only when all its four neighbors are defectors, with probability of $r_{\mathcal{L}}^4$. So $r_{\mathcal{C}}$ could be calculated as:

$$r_{\mathcal{C}} = r_{\mathcal{C}0} * (1 - r_{\mathcal{D}0}^4) \approx 0.3291. \tag{4}$$

which is consistent with the Monte Carlo simulation result shown in Fig. 2.

3.3. Defection Suppression($A > 1.2$)

When A is large, cooperators and loners will expand and coexist while defectors' survival is greatly suppressed. Fig. 4 presents $r_{\mathcal{C}}$, $r_{\mathcal{L}}$ and $r_{\mathcal{D}}$ with the time-evolution when $A = 1.6$ and $b = 1.6$. It could be easily observed that there are four obvious phases:

- At first, because the initial aspiration of an individual is 1.6 and a loner's payoff is always 1.2, all the loners are dissatisfied and try to change their strategies. On the contrary, most of the cooperators and defectors are satisfied. Besides, defectors's payoffs are higher than cooperators' on average. So $r_{\mathcal{L}}$ decreases fast and defectors expand transitorily.
- With t growing, loners' aspirations become lower than 1.2, so they get satisfied and never change their strategies any more. Cooperators and defectors with low payoffs will try to evolve into loners and become satisfied, so loners could expand stably. At the same time, cooperators form some clusters gradually, which is like to the so-called END period [43,44].
- With t further growing, cooperators which still survive have formed some clusters. Dissatisfied defectors and loners neighboring with these clusters will evolve into cooperators and cause the chain phenomenon, which causes cooperators' expansion and it is called EXP period [43,44]. $r_{\mathcal{C}}$ increases while $r_{\mathcal{D}}$ and $r_{\mathcal{L}}$ decrease.
- Finally, all of the three strategies have formed some stable clusters which will never evolve any more, while there are some regions where three strategies mix well. In these regions, because of the cyclic dominance, Rock-Scissors-Paper-type cycles occur in three strategies. Since loners and cooperators' clusters are easier to expand, it could be found that $r_{\mathcal{C}}$ and $r_{\mathcal{L}}$ increase while $r_{\mathcal{D}}$ decreases concussively. The whole network will be stable at about $t = 2000$.

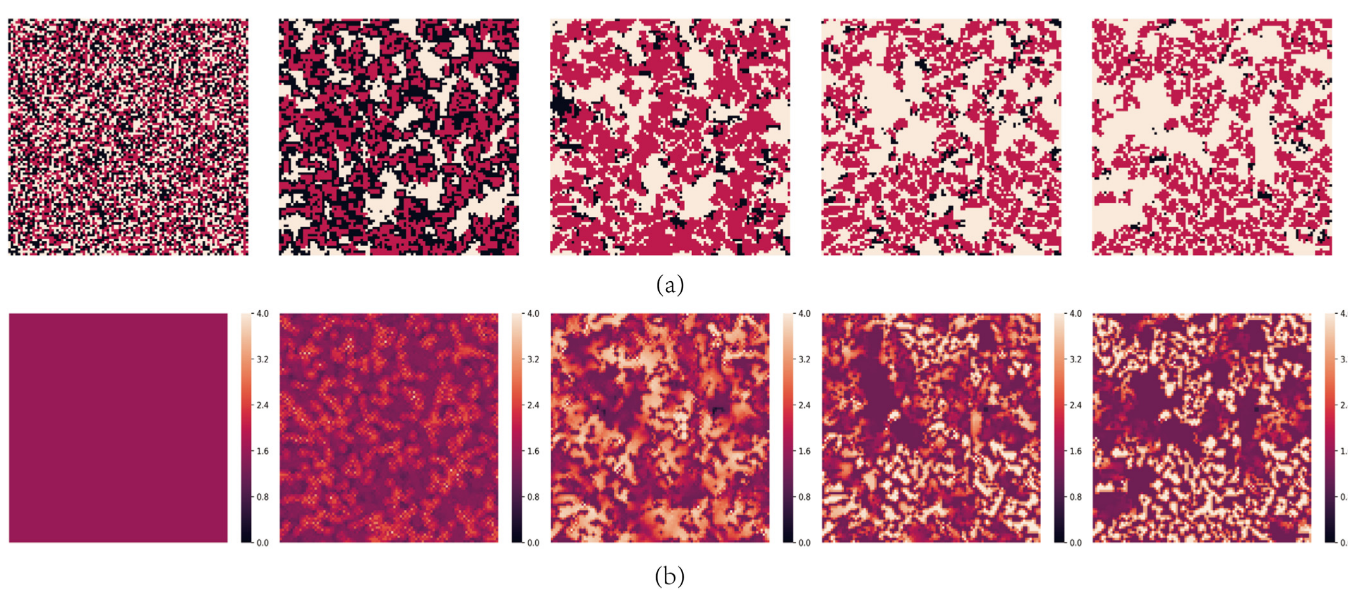

Fig. 5. (a) shows characteristic snapshots of cooperators (red), loners (white) and defectors (black) with time growing. (b) shows the heat map of aspiration distribution with time growing. The steps of them are t=0, 10, 100, 1000 and 2000 from left to right respectively. The results were got when $A = 1.6$ and $b = 1.6$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

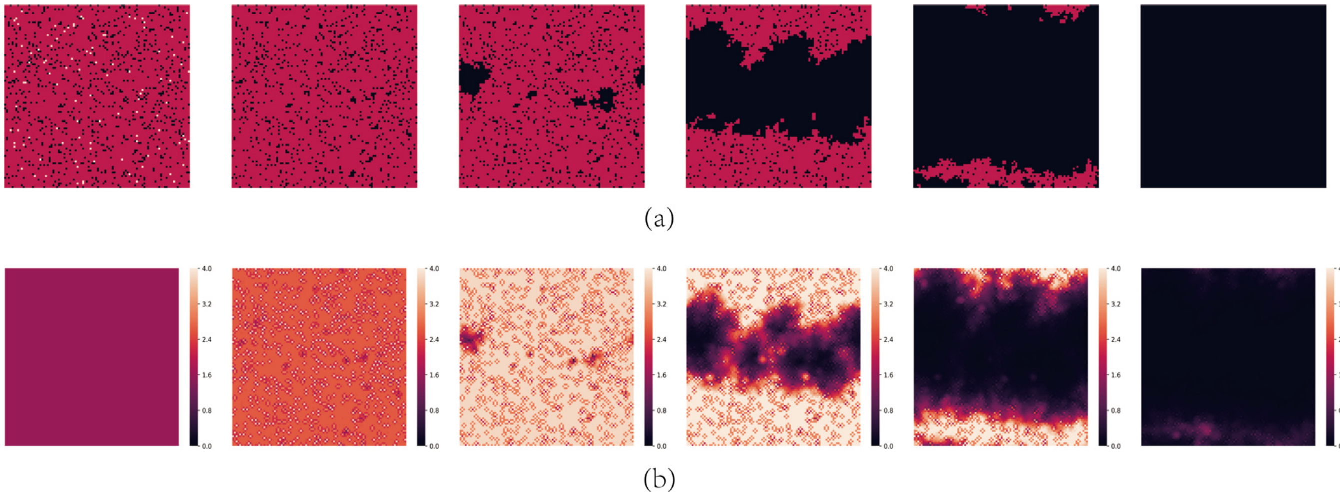

Fig. 6. (a) shows characteristic snapshots of cooperators (red), loners (white) and defectors (black) with time growing. (b) shows the heat map of aspiration distribution with time growing. The steps of them are t=0, 10, 50, 100, 150 and 200 from left to right respectively. The results were got when $A = 1.6$ and $b = 1.6$ with $r_{\mathcal{C}0} = 0.9$, $r_{\mathcal{D}0} = 0.09$ and $r_{\mathcal{L}0} = 0.01$ initially. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

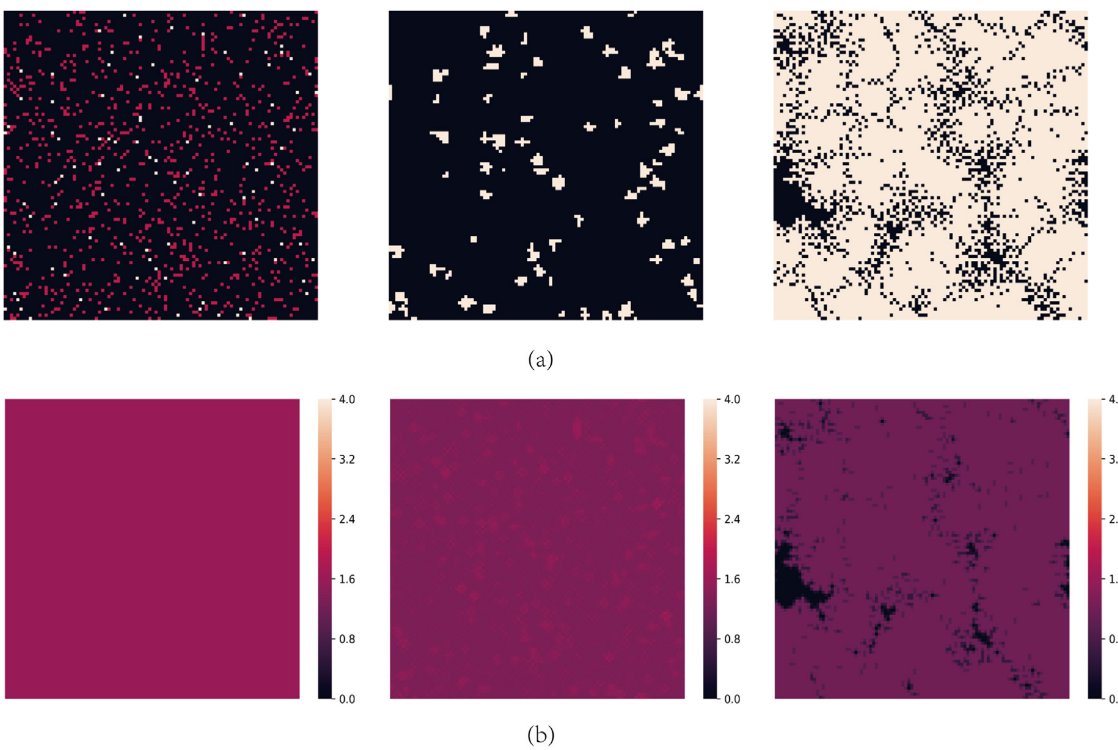

Fig. 7. (a) shows characteristic snapshots of cooperators (red), loners(white) and defectors (black) with time growing. (b) shows the heat map of aspiration distribution with time growing. The steps of them are t=0, 5 and 100 from left to right respectively. The results were got when $A = 1.6$ and $b = 1.6$ $r_{\mathcal{C}0} = 0.09$, $r_{\mathcal{D}0} = 0.9$ and $r_{\mathcal{L}0} = 0.01$ initially. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

In order to further discuss how cooperators, loners and defectors and their aspiration levels distribute in the network, Fig. 5 represents snapshots of strategies and aspirations for $A = 1.6$ and $b = 1.6$. As it is shown, cooperators and loners are separated by defectors at first and all of them are located in small-scale clusters. When the network is stable, loners have formed several large-scale clusters while cooperators are still located in many small-scale clusters. It is because that when neighboring directly, loners are dominant over defectors while defectors are dominant over cooperators.

According to the above analysis, voluntary participation plays an important role to promote cooperation because of loner's fixed payoff. When loners get satisfied, they are certain to survive and expand. But it should be noticed that loners might be extinct before part of them get satisfied. As Fig. 6 showing, loners will be extinct soon when the value of $r_{\mathcal{L}0}$ is too small. Then the network degenerates to the two-strategies condition where defectors could expand to the whole network under such parameters. In this case, cooperators couldn't survive with the help of loners.

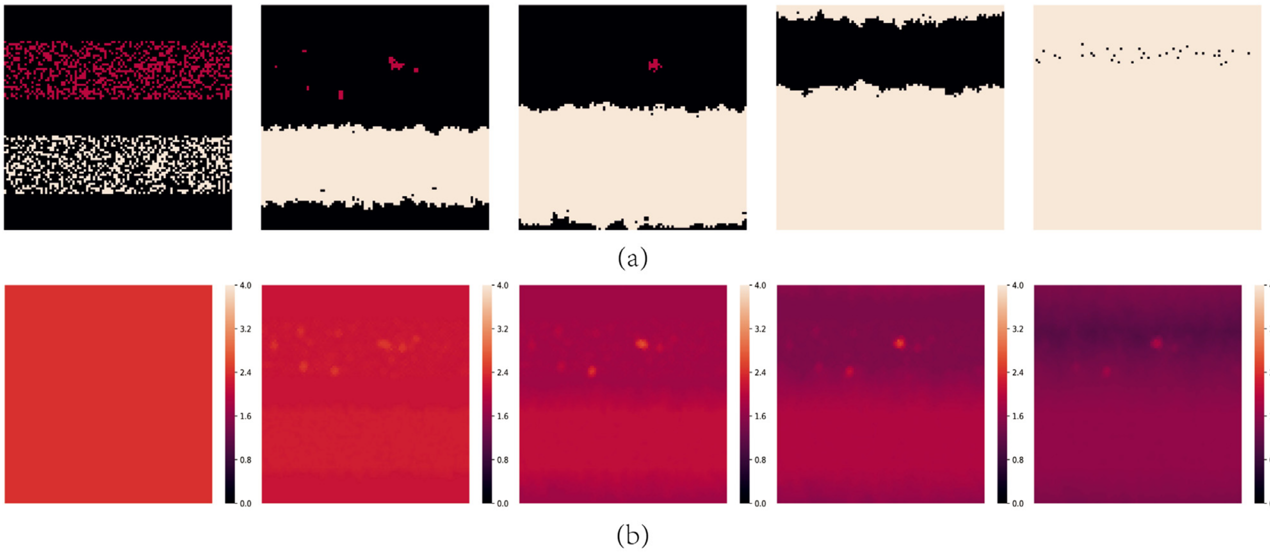

Fig. 8. (a) shows characteristic snapshots of cooperators (red), loners (white) and defectors (black) with time growing. (b) shows the heat map of aspiration distribution with time growing. The steps of them are t=0, 10, 30, 50 and 100 from left to right respectively. The results were got when $A = 2.4$ and $b = 1.6$ with $r_{\mathcal{C}0} = 0.1$, $r_{\mathcal{D}0} = 0.8$ and $r_{\mathcal{L}0} = 0.1$ initially, where cooperators and loners are separated by defectors. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

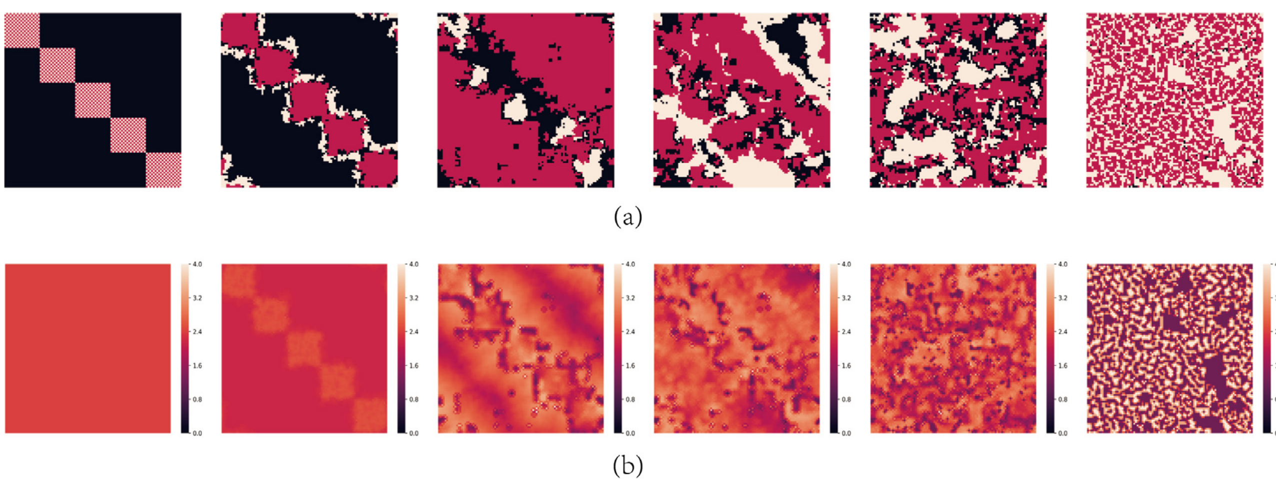

Fig. 9. (a) shows characteristic snapshots of cooperators (red), loners (white) and defectors (black) with time growing. (b) shows the heat map of aspiration distribution with time growing. The steps of them are t=0, 10, 100, 500, 1000 and 10000 from left to right respectively. The results were got when $A = 2.4$ and $b = 1.6$ with $r_{\mathcal{C}0} = 0.1$, $r_{\mathcal{D}0} = 0.8$ and $r_{\mathcal{L}0} = 0.1$ initially, where cooperators and loners are mixed well. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Besides, whether cooperators could survive is also worth considering. Fig. 7 shows another adverse condition that cooperators will be extinct. At first, nearly all cooperators are surrounded by defectors and loners. They are dissatisfied and get lower payoffs than their $\mathcal{D}$ and $\mathcal{L}$ neighbors, so cooperators are extinct more quickly than loners. Then loners could expand easily among defectors because a loner's payoff will always be larger than a defector's payoff if no cooperator exists. As shown in Fig. 7 (a), loners will occupy most of the network when stable, but no cooperator exists.

From the above, voluntary participation could promote cooperation only when both loners and cooperators are not extinct during the evolution process. To further understand this condition, some special setup should be considered. Fig. 8 shows a typical special setup in which cooperators and loners are separated by defectors. It is shown from Fig. 8 that in the upper half of the network, all cooperators are surrounded by defectors and they couldn't survive when $b = 1.6$. Meanwhile in the lower half of the network, loners expand because of the higher payoffs. But when loners expand into the upper half of the network, cooperators have become extinct, so the loners will finally occupy the whole network and cooperation is not promoted. Fig. 9 shows another initial setting for the same parameters and fractions of three strategies with Fig. 8. Cooperators and loners mix well in five clusters initially, and it could be observed that cooperators expand on the border of the clusters fast by the influence of loners. When loners form clusters and get satisfied, they will be stable, then cooperators and loners will coexist with moderate fractions. Well-mixed cooperators and loners could expand together, which is the main reason why voluntary participation could promote cooperation.

To conclude, for small values of A ($A \leq 1.2$), three kinds of strategies could coexist. Loners will never change their strategies because of their fixed payoffs, while part of cooperators and defectors are dissatisfied and evolve into loners, which depends on the value of A but is independent with the value of b. Then the cyclic Rock-Scissors-Paper-type of dominance happens and three strategies could coexist. For large values of A ($A > 1.2$), cooperators and loners could expand and coexist while defectors' survival is greatly suppressed. Loners could form some large stable clusters where all individuals are always satisfied so they could expand unconditionally. Most of the defectors are hard to survive when there are a large number of loners in the network because they have no chance to get the temptation value T, even when it is large. On the contrary, cooperators could coexist with loners easily by forming many small clusters. Besides, it should be noticed that initial distribution has a significant impact on promoting cooperation. Cooperators or loners may be extinct under the adverse initial distribution. Loners should be fully adjacent to both cooperators and defectors, which is a necessary condition for promoting cooperation. In traditional fixed aspiration model, voluntary participation plays a rather limited role in promoting cooperation under high aspiration levels, where loners are always dissatisfied so they are hard to survive. Defection cannot be suppressed when b is large. However, dynamic aspiration model provides a favorable environment for loners to survive and expand, an individual' aspiration could decrease to a low level if they get dissatisfied for a long term. When an individual' aspiration becomes lower than 1.2, it could keep $\mathcal{L}$ as its strategy because of the fixed payoff.

4. Conclusion

In summary, this paper discusses how voluntary participation impacts PD games with Dynamic-Win-Stay-Lose-Learn strategy updating rule. This dynamic model is adjusted by a single parameter a. It is found that the proposal of strategy $\mathcal{L}$ could promote cooperation and suppress defection within a wide range of parameters, especially when b is large. We also studied how initial distribution influences the evolutionary process to reveal some adverse initial distribution. In OPD games, the best way for loners and cooperators to survive is mixing well and forming clusters gradually, which could be easily achieved with Dynamic-Win-Stay-Lose-Learn strategy updating rule.

Our work combines voluntary participation with dynamic aspiration model to provide a new perspective on how voluntary participation promotes cooperation in PD games. In terms of the broader relevance of our research, comparing with fixed aspiration model, dynamic aspiration model is more in line with the law of evolutionary games. Dynamic aspiration is the natural tendency among humans, so it is expected that our work provides some reference values for solving the social dilemma in the real world [45,46].

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Zhenyu Shi: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition. **Wei Wei:** Conceptualization, Formal analysis, Writing – review & editing, Project administration, Funding acquisition. **Xiangnan Feng:** Validation, Formal analysis, Investigation, Writing – review & editing, Visualization, Supervision. **Ruizhi Zhang:** Methodology, Data curation, Investigation, Writing – original draft, Writing – review & editing, Visualization, Supervision. **Zhiming Zheng:** Conceptualization, Supervision, Project administration, Funding acquisition.

Acknowledgments

This work is supported by the Fundamental Research Funds for the Central Universities, the Research and Development Program of China (No.2018AAA0101100), the Beijing Natural Science Foundation (1192012, Z180005) and National Natural Science Foundation of China (No.62050132).

Appendix A. Results with small value of $r_{\mathcal{L}}$

Though voluntary participation is spontaneity and spreadability, it will not exist in large quantities in initial in many practical scenarios besides of the initial settings that all three strategies occupy one third of the network as shown in Fig. 1, it is also necessary to discuss the result with initial settings that loners occupy a relatively small percentage of the network. Fig. A.1 shows $r_{\mathcal{C}}$, $r_{\mathcal{L}}$ and $r_{\mathcal{D}}$ as a function depending on $r_{\mathcal{L}0}$ for $A = 1.6$ and $b = 1.6$, where the value of $r_{\mathcal{L}0}$ is from 0 to 0.25 and $r_{\mathcal{C}0} = r_{\mathcal{D}0}$. For each pair of parameters, we also did 20 independent simulations to make the results more accurate. As it shown, compared to $r_{\mathcal{L}0} = 0$, where $r_{\mathcal{C}}$ is close to 0 when the network is stable, only a few loners being set initially could protect cooperation. Both $r_{\mathcal{C}}$ and $r_{\mathcal{L}}$ are greater than $r_{\mathcal{D}}$ only if $r_{\mathcal{L}0} > 0.04$. And when $r_{\mathcal{L}0} > 0.10$, the results are similar to the condition that all three strategies occupy one third of the network in initial.

Figs. A.2 and A.3 present $r_{\mathcal{C}}$, $r_{\mathcal{L}}$ and $r_{\mathcal{D}}$ with the time-evolution when $A = 1.6$ and $b = 1.6$, with $r_{\mathcal{L}0} = 0.04$ and $r_{\mathcal{L}0} = 0.1$ respectively. It could be observed the same four obvious phases as Fig. 4 shown, and the only difference is the quantity. An interesting phenomenon is when $r_{\mathcal{L}0} = 0.04$, the final fraction $r_{\mathcal{L}}$ is greater than when $r_{\mathcal{L}0} = 0.1$ and 0.33. From Fig. A.2 we could observe that the main reason is after $r_{\mathcal{L}}$ increases in the second stage, it will only decrease slightly in the third stage. Due to lack of loners initially, defectors expand fast so that the average payoffs and aspirations of the whole network are much lower. Once an individual evolves into a loner, it will get satisfied with high probability because of its low aspiration level, so it will keep its strategy $\mathcal{L}$. The above analysis further clarifies how defectors are suppressed by loners, though there is only a very small proportion of loners in initial.

To conclude, the emergence of cooperation and the suppression of defection don't require loners to exist in sufficiently

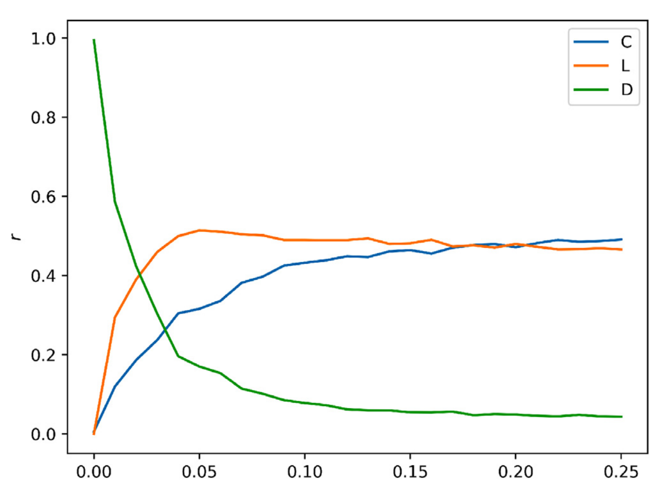

Fig. A1. Average fractions of cooperators, loners and defectors in the stable state as a function depending on $r_{\mathcal{L}0}$ for $A = 1.6$ and $b = 1.6$, where the value of $r_{\mathcal{L}0}$ is from 0 to 0.25 and $r_{\mathcal{C}0} = r_{\mathcal{D}0}$.

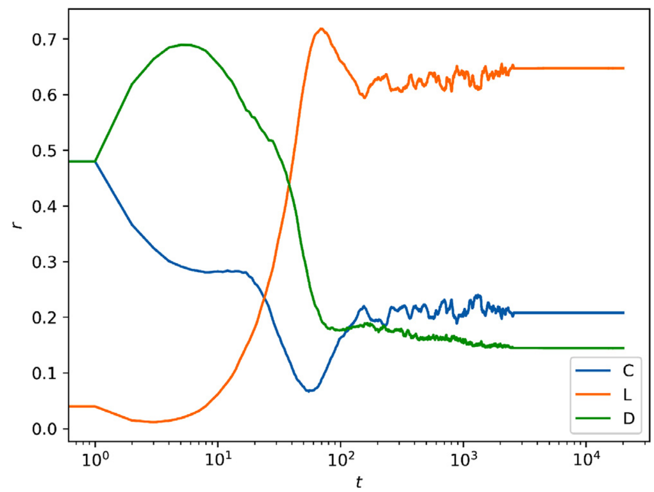

Fig. A2. Average fractions of cooperators, loners and defectors as a function of step t when $A = 1.6$ and $b = 1.6$, with $r_{\mathcal{L}0} = 0.04$ and $r_{\mathcal{C}0} = r_{\mathcal{D}0}$.

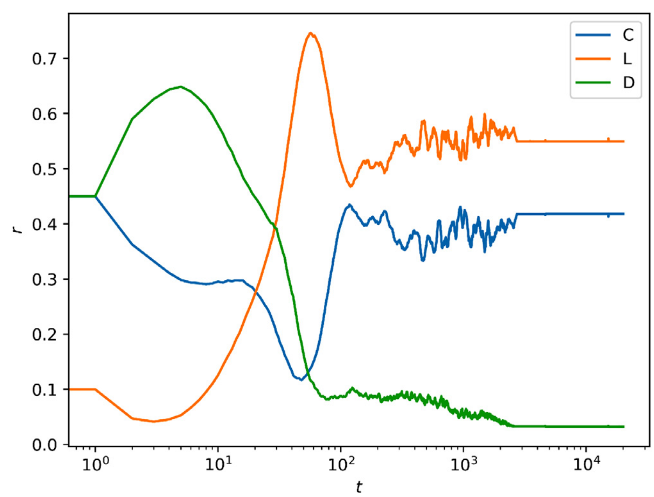

Fig. A3. Average fractions of cooperators, loners and defectors as a function of step t when $A = 1.6$ and $b = 1.6$, with $r_{\mathcal{L}0} = 0.1$ and $r_{\mathcal{C}0} = r_{\mathcal{D}0}$.

large densities in initial. The key principle is dynamic aspiration model provides a favorable environment for loners to survive and expand.

Appendix B. Results on the scale-free networks

The main results of our model on the lattice have been shown. But in the real world, the spatial structure of population is often irregular, so the result on other irregular networks is also worth studying. Scale-free network is one of classic network models on which most nodes have few neighbors while a small minority of nodes have a large number of neighbors [47]. It assumes that individuals tend to connect with others with more neighbors, which is in line with many practical problems. In our paper, a classic algorithm named Barabasi-Albert (BA) algorithm is used to generate random scale-free networks [48]. In BA algorithm, parameter m represents the number of neighbors a new node will connect with, and N represents the total number of the nodes. At first, a network with m nodes and no edge is generate. Then the remaining $N - m$ nodes are added into the network one by one, and each of them will connect with other m already existing nodes randomly, probability of which is proportional to degrees of already existing nodes. We set $N = 3000$ to ensure the network is large enough, and m is set to 1,2,4 and 8 respectively to investigate the impact of different network connectivity on the results. Different from the process in the *Model* section, each individual updates its strategy and aspiration asynchronously, which means in each step there is only one node updating its strategy and aspiration. The step will carry out 150,000 times repeatedly in a simulation, and 20 independent simulations are performed to make the results more accurate. Fig. B1 shows $r_{\mathcal{C}}$, $r_{\mathcal{L}}$ and $r_{\mathcal{D}}$ as a function depending on b for different values of A, where the network is randomly generated and each of the strategies occupy one third of the network. It is shown that the value of m has a significant impact on the results, and all the results could be compared with $A = 0$ in each sub-figure where all individuals are always satisfied.

- When $m = 1$ and $m = 2$, cooperation could be promoted unless $A = 0.2$, and as A increases, $r_{\mathcal{C}}$ also increases. It is worth noting that when $A = 0.2$, the results are similar as what we have got in lattice, and not related to the value of m. The reason for it is also the same: all loners and the vast majority of defectors are satisfied initially, only a part of cooperators are dissatisfied and evolve into other strategies then the network is stable. The above phenomenon is universal when $A \leq 4l$, so we will not repeat the discussion in the subsequent analysis.
- When $m = 4$, cooperation could not be promoted for every value of A, when $A = 0.4$, loners will dominate the network. When $A = 0.6$ and $A = 0.8$, $r_{\mathcal{L}}$ and $r_{\mathcal{D}}$ are close. Lower value of A could make loners get satisfied more quickly so $r_{\mathcal{L}}$ is larger than when the value of A is higher.
- When $m = 8$, survival of cooperation will become more difficult. Defectors will occupy the whole network when $A = 0.4$. While when $A = 0.6$ and $A = 0.8$, defectors will dominate the network when $b < 1.4$, and with b increasing, $r_{\mathcal{C}}$ and $r_{\mathcal{D}}$ decrease while $r_{\mathcal{L}}$ increases gradually. When $b > 1.55$, loners will occupy the whole network.

To conclude, the combination of low value of m and high initial aspirations A could promote cooperation in scale-free network, while moderate and high value of m will always unable to protect cooperators' survival. The more detailed mechanism how complex networks affect cooperation phenomenon in our dynamic aspiration model will be the focus of our future research.

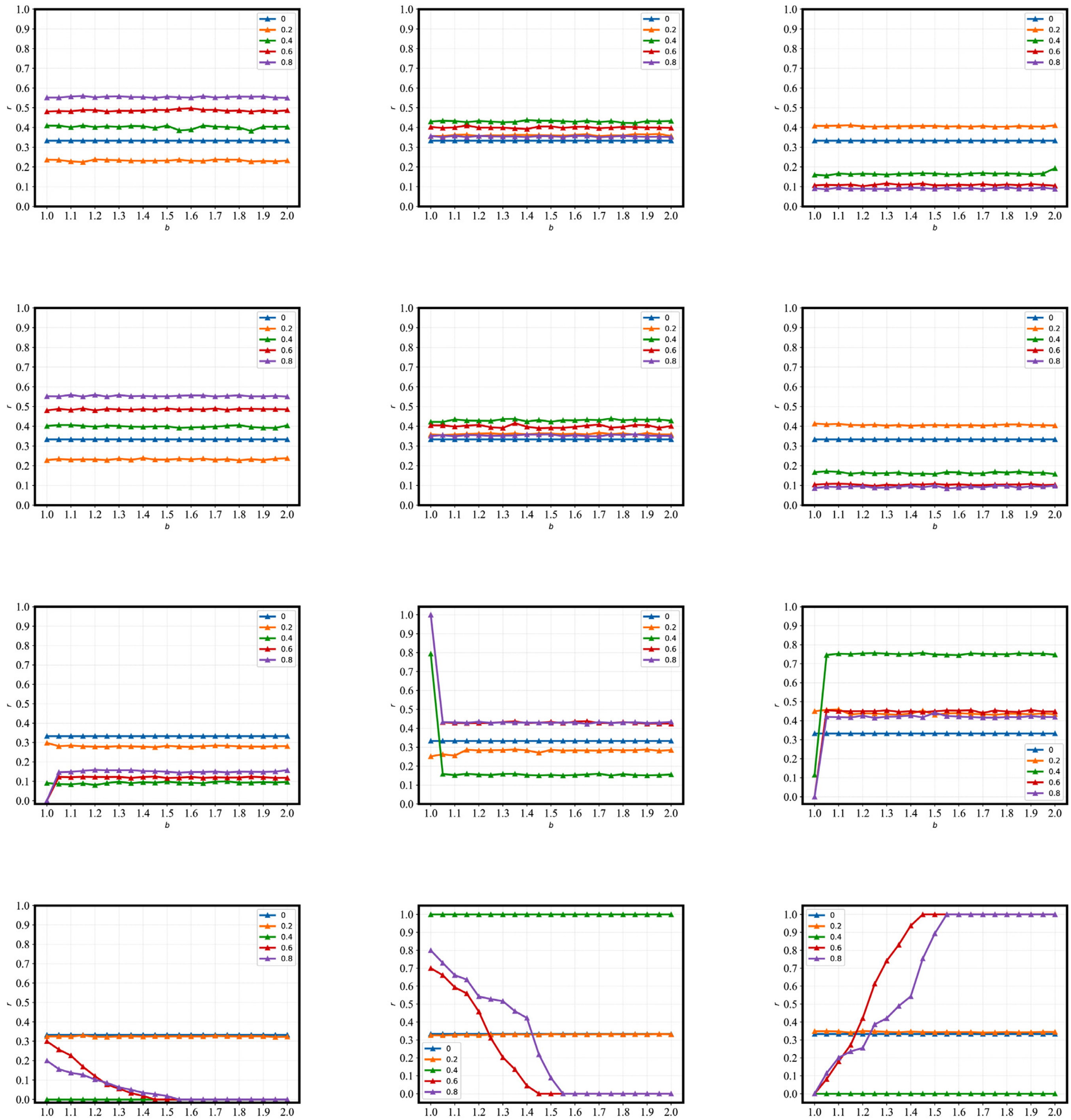

Fig. B1. Average fractions of three strategies in the stable state in dependence on b at different values of the A in BA networks, from left to right are fractions of cooperators, loners and defectors respectively, and from top to bottom are m=1, 2, 4 and 8 respectively.

References

[1] Myerson RB. Game theory: analysis of conflict. Harvard University Press; 1991.
[2] Schuster R, Perelberg A. Why cooperate?: an economic perspective is not enough. Behav Processes 2004;66(3):261–77.
[3] Gibbons RS. Game theory for applied economists. Princeton University Press; 1992.
[4] Axelrod R, Hamilton WD. The evolution of cooperation. Science 1981;211(4489):1390–6.
[5] Nowak MA. Five rules for the evolution of cooperation. Science 2006;314(5805):1560–3.
[6] Nowak MA, May RM. Evolutionary games and spatial chaos. Nature 1992;359(6398):826–9.
[7] Szabo G, T?Ke C. Evolutionary prisoner's dilemma game on a square lattice. Phys Rev E 1998;58(1):69–73.
[8] Wu Z, Xu X, Chen Y, Wang Y. Spatial prisoner's dilemma game with volunteering in Newman-Watts small-world networks. Phys Rev E 2005;71(3):037103.
[9] Ren J, Wang WX, Qi F. Randomness enhances cooperation: a resonance-type phenomenon in evolutionary games. Phys Rev E 2007;75(4 Pt 2):045101.
[10] Rong Z, Li X, Wang X. Roles of mixing patterns in cooperation on a scale-free networked game.. Phys Rev E 2007;76(2):027101.
[11] Du WB, Cao XB, Lin Z, Hu MB. Evolutionary games on scale-free networks with a preferential selection mechanism. Physica A 2009;388(20):4509–14.
[12] Herrmann B, Thöni C, Gächter S. Antisocial punishment across societies. Science 2008;319(5868):1362–7.

[13] Helbing D, Szolnoki A, Perc M, Szabó G. Punish, but not too hard: how costly punishment spreads in the spatial public goods game. N J Phys 2010;12(8):083005.
[14] Chen X, Szolnoki A, Perc M. Probabilistic sharing solves the problem of costly punishment. N J Phys 2014;16(8):083016.
[15] Chen X, Sasaki T, Brännström Å, Dieckmann U. First carrot, then stick: how the adaptive hybridization of incentives promotes cooperation. J R Soc Interface 2015;12(102):20140935.
[16] Szolnoki A, Perc M. Conformity enhances network reciprocity in evolutionary social dilemmas. J R Soc Interface 2015;12(103):20141299.
[17] Szolnoki A, Perc M. Leaders should not be conformists in evolutionary social dilemmas. Sci Rep 2016;6:23633.
[18] Li Y, Zhang J, Perc M. Effects of compassion on the evolution of cooperation in spatial social dilemmas. Appl Math Comput 2018;320:437–43.
[19] Wang WX, Ren J, Chen G, Wang BH. Memory-based snowdrift game on networks. Phys Rev E 2006;74(5 Pt 2):056113.
[20] Qin S, Chen Y, Zhao X, Shi J. Effect of memory on the Prisoner's dilemma game in a square lattice. Phys Rev E 2008;78(4).
[21] Yang HX, Rong Z, Lu PM, Zeng YZ. Effects of aspiration on public cooperation in structured populations. Physica A 2012;391(15):4043–9.
[22] Wu T, Fu F, Wang L. Coevolutionary dynamics of aspiration and strategy in spatial repeated public goods games. N J Phys 2018;20(6).
[23] Chu C, Mu C, Liu J, Liu C, Boccaletti S, Shi L, et al. Aspiration-based coevolution of node weights promotes cooperation in the spatial prisoner's dilemma game. N J Phys 2019.
[24] Zhang L, Huang C, Li H, Dai Q. Aspiration-dependent strategy persistence promotes cooperation in spatial prisoner's dilemma game. EPL 2019;126(1):18001.
[25] Posch M, Pichler A, Sigmund K. The efficiency of adapting aspiration levels. Proc R Soc London SerB 1999;266(1427):1427–35.
[26] Amaral MA, Wardil L, Perc M, da Silva JK. Stochastic win-stay-lose-shift strategy with dynamic aspirations in evolutionary social dilemmas. Phys Rev E 2016;94(3):032317.
[27] Arefin MR, Tanimoto J. Evolution of cooperation in social dilemmas under the coexistence of aspiration and imitation mechanisms. Phys Rev E 2020;102(3):032120.
[28] Li C, Fan S. A dynamic aspiration-based interaction strategy blocks the spread of defections in social dilemma. EPL 2020;129(4):48002(7pp).
[29] Shi Z, Wei W, Feng X, Li X, Zheng Z. Dynamic aspiration based on win-stay-lose-learn rule in spatial prisoner's dilemma game. Plos One 2021;16(1):e0244814.
[30] Liu Y, Chen X, Zhang L, Wang L, Perc M. Win-stay-lose-learn promotes cooperation in the spatial prisoner's dilemma game. PloS One 2012;7(2):e30689.
[31] Chu C, Liu J, Shen C, Jin J, Shi L. Win-stay-lose-learn promotes cooperation in the prisoner's dilemma game with voluntary participation. Plos One 2017;12(2):e0171680.
[32] Fu M-J, Yang H-X. Stochastic win-stay-lose-learn promotes cooperation in the spatial public goods game. Int J Modern Phys C 2018;29(04):1850034.
[33] Szolnoki A, Perc M, Szabo G. Phase diagrams for three-strategy evolutionary prisoner's dilemma games on regular graphs. Phys Rev E 2009;80(5):056104.
[34] Szolnoki A, Chen X. Alliance formation with exclusion in the spatial public goods game. Physreve 2017;95(5-1):052316.
[35] Cardinot M, O'Riordan C, Griffith J. The optional prisoner's dilemma in a spatial environment: coevolving game strategy and link weights. In: 8th International conference on evolutionary computation theory and applications; 2016.
[36] Myerson RB. Game theory. Harvard university press; 2013.
[37] Szabó G, Tőke C. Evolutionary prisoner's dilemma game on a square lattice. Phys Rev E 1998;58(1):69.
[38] Szabó G, Vukov J, Szolnoki A. Phase diagrams for an evolutionary prisoner's dilemma game on two-dimensional lattices. Phys Rev E 2005;72(4):047107.
[39] Perc M. Coherence resonance in a spatial prisoner's dilemma game. N J Phys 2006;8(2):22.
[40] Geng Y, Hu K, Shen C, Shi L, Wang Z. Aspiration induced interdependence leads to optimal cooperation level. Chaos 2019;29(8):083114.
[41] Zhang L, Huang C, Li H, Dai Q, Yang J. Cooperation guided by imitation, aspiration and conformity-driven dynamics in evolutionary games. Physica A 2021;561:125260.
[42] Wu T, Fu F, Wang L. Coevolutionary dynamics of aspiration and strategy in spatial repeated public goods games. N J Phys 2018;20(6):063007.
[43] Ogasawara T, Tanimoto J, Fukuda E, Hagishima A, Ikegaya N. Effect of a large gaming neighborhood and a strategy adaptation neighborhood for bolstering network reciprocity in a prisoner's dilemma game. J Stat Mech 2014;2014(12):P12024.
[44] Kabir KA, Tanimoto J, Wang Z. Influence of bolstering network reciprocity in the evolutionary spatial prisoner's dilemma game: a perspective. Eur Phys J B 2018;91(12):312.
[45] Wang Z, Jusup M, Wang R-W, Shi L, Iwasa Y, Moreno Y, et al. Onymity promotes cooperation in social dilemma experiments. Sci Adv 2017;3(3):e1601444.
[46] Wang Z, Bauch CT, Bhattacharyya S, d'Onofrio A, Manfredi P, Perc M, et al. Statistical physics of vaccination. Phys Rep 2016;664:1–113.
[47] Barabási A-L, Bonabeau E. Scale-free networks. Sci Am 2003;288(5):60–9.
[48] Barabási A-L, Albert R. Emergence of scaling in random networks. Science 1999;286(5439):509–12.