Received August 3, 2019, accepted August 23, 2019, date of publication August 29, 2019, date of current version September 13, 2019.

*Digital Object Identifier 10.1109/ACCESS.2019.2938254*

# An Evolutionary Game Coordinated Control Approach to Division of Labor in Multi-Agent Systems

**JINMING DU** [1,2,3], **(Member, IEEE)**
[1]Institute of Industrial and Systems Engineering, College of Information Science and Engineering, Northeastern University, Shenyang 110819, China
[2]Liaoning Engineering Laboratory of Operations Analytics and Optimization for Smart Industry, Northeastern University, Shenyang 110819, China
[3]Key Laboratory of Data Analytics and Optimization for Smart Industry (Northeastern University), Ministry of Education, Shenyang 110819, China

e-mail: dujinming@ise.neu.edu.cn

This work was supported in part by the National Natural Science Foundation of China under Grant 61703082, in part by the Fundamental Research Funds for the Central Universities under Grant N160403001, in part by the Major Program of the National Natural Science Foundation of China under Grant 71790614, in part by the National Key Research and Development Program of China under Grant 2016YFB0901900, in part by the Fund for Innovative Research Groups of the National Natural Science Foundation of China under Grant 71621061, in part by the Major International Joint Research Project of the National Natural Science Foundation of China under Grant 71520107004, and in part by the 111 Project under Grant B16009.

**ABSTRACT** In this paper, we propose an evolutionary game theoretic approach to coordinated control of multi-agent systems. In this mathematical framework, agents play games with their neighbors on the network, and update strategies through local interaction. In order to achieve a certain control objective of the system, we need to select the appropriate game type, design the calculation and evaluation methods of fitness, specify the interactive constraints and updating rules. During the evolutionary process of the system, agents have no predesigned dynamical equations. They adjust their behavior independently for the purpose of increasing their own benefits. The system achieves its final state in the process of individual interaction and autonomous decision-making. Taking division of labor problem as an example, we demonstrate the proposed control approach in detail. The performance of the theoretical method is verified by simulation on the regular graph, the general connected graphs, and heterogeneous scale-free networks, respectively.

**INDEX TERMS** Evolutionary game theory, multi-agent systems, coordinated control, complex system dynamics, division and cooperation.

## I. INTRODUCTION

The interaction among multiple agents are ubiquitous phenomena whatever in the animal world, microbial community, and also human society [1]–[4]. Individuals in a complex system are closely connected through interaction, communication, cooperation, adaptation, organization, learning and division of labor.

Multi-agent system is a typical model for studying the interactive behaviors among multiple individuals. It has been used to describe many theoretical and practical problems [5]. Among these issues, the problem of coordinated control of multiple agents has been extensively concerned [6]. It has a wide range of practical application background, such as moving, tracking, formation control of swarms, disaster rescue, multiple satellite cluster system and etc. Recently, some of the typical problems have been studied in the field of systems and control, such as consensus [7]–[13], coverage [14], formation [15]–[17], flocking [18], synchronization [19], controllability [20] and so on. Researchers have proposed various control methods, such as behavior based, virtual structure, leader-following, graph theory based and artificial potential field method [21]. In multi-agent systems, the connection between agents is determined by a fixed or time-variant topology. Agents communicate with their neighbors and exchange information. Thus information processing and decision-making are usually based on local information. The dynamics of agents are characterized by predesigned differential equations. According to such fixed control law, agents perform the predetermined action.

In many actual systems, however, agents are "intelligent" and they do not only passively follow predetermined

The associate editor coordinating the review of this article and approving it for publication was Zonghua Gu.

  

dynamical rules. Instead, they rationally consider their own energy, loss, behavior cost and other factors to pursue their own interests when they interact with other agents. Individual rationality has been widely confirmed in biological and social groups. The rational decision-making of one agent would improve its fitness. But at the same time, it may affect the interests of other agents. This may lead to the deviation of group system from the optimal performance. Therefore, to achieve the control objective of the whole multi-agent systems, it is worthwhile to study the competition and conflicts of interests among agents, the coordination between individual interests and group interests, and the organization of the whole system and all its parts.

When it comes to individuals with conflict of interests, game theory [22], [23] provides an effective mathematical tool and research framework for the study on the interaction among individuals. The coordinated control problem is different from the classical optimization problem. It is difficult to improve system performance by simply maximizing the given indicators under constraints. This is a more complex situation involving coordination among multiple agents whose interests may be conflicting. When the fitness of individual are closely related with the proportion of its strategy in the population, evolutionary game theory offers a general theoretical framework [24]–[26]. Such framework has been successfully used to study issues such as host parasite interaction, ecosystems, animal behavior, social evolution, and human language evolution [27]–[32].

In this paper, a mathematical framework is proposed to study the coordinated control of multi-agent systems based on evolutionary game theory. On the basis of the idea of natural selection, agents in the system act as the role of players in the game. Agents have different optional behaviors to choose as their strategies. They play games with their neighbors, and update strategies through local interaction. After each game, agents obtain payoffs based on the game matrix. The agents rationally adjust their strategies according to their own interests by evaluating their benefits. The system evolves while agents autonomously update their strategies. The overall performance of the system is determined by the behavior states of all the agents. In order to achieve the control objective of the system, such as consensus, synchronization and division of labor, the following elements which are the key to affecting the evolutionary direction of the system should be designed. We should select the appropriate game type, design the reasonable fitness calculation and evaluation methods, specify updating rule and relevant interactive restrictions. In addition, population size, strategy space, state set, communication topology, and other relevant parameters should be determined according to the practical problems. During the evolutionary process, there is no need to specify how a particular agent acts. Therefore we need not determine the dynamical equation of each individual. The agents have the capability of adaptive decision-making, rather than passively following a given differential equation. The control objective of the system could be achieved through adaptive evolution of the population. It is a controllable, intelligent and autonomous decision-making process.

## II. MULTI-AGENT SYSTEM MODEL

By utilizing evolutionary game theory, we propose a mathematical framework to realize the coordinated control of multi-agent systems.

For a generalized model, we consider a system consisting of $n$ agents. The population is determined by set $N = \{1, 2, \cdots, n\}$. For each agent $i \in N$, it is defined as follows:

$$Agent_i = (S_i, B_i, C_i, F_i), \tag{1}$$

where

- " $S_i$ (State) " reflects the state information of the focal agent at the current moment, such as the strategy, position, velocity, and etc.
- " $B_i$ (Behavior) " denotes the action or decision of the agent, which determines how agents adjust their state. For example, the rate of velocity change, the change of direction, and so on. For agents with a series of behaviors, it is defined as follows $B_i = (b_{i1}, b_{i2}, \cdots, b_{im})$. In some specific problems, the agent's behavior and state may be merged into account.
- " $C_i$ (Communication) " denotes the interactive relationship among agents. For example, it can be defined as a communication topology graph which reflects the information connection between agents and their neighbors. In particular, we can define a neighborhood set to show the neighbors of agents.
- " $F_i$ (Fitness)" depicts the ability of the agent to adapt to the competition under the sense of natural selection. The most straightforward form is the agent's payoff in the game. Through the assessment of fitness, the agent adjusts its strategy in order to obtain higher benefits.

The communication topology graph of multi-agent systems can be defined as $G = (N, \varepsilon)$, where the set of nodes is $N = \{1, 2, \cdots, n\}$, and the set of edges is $\varepsilon \subseteq N \times N$. Each agent, $i \in N$, independently chooses its behavior in the next step based on local information. In particular, when the communication topology is a directed graph, an ordered pair $(i, j)$ represents the directed edge in the graph. If the agent $i$ is able to receive the state information of the agent $j$, the agent $j$ is a neighbor of $i$. The neighborhood of agent $i$ is defined as $K_i = \{j \in N : (i, j) \in \varepsilon\}$, and we assume that $i$ is not its own neighbor. For undirected graphs, if $(i, j) \in \varepsilon$, then $(j, i) \in \varepsilon$.

We assume that agents play a symmetric $2 \times 2$ game with all their neighbors. Then they obtain payoffs according to the payoff matrix as follows:

$$\begin{array}{c|cc} & A & B \\ \hline A & a & b \\ B & c & d \end{array}. \tag{2}$$

For the game interaction, an $A$ strategy holder interacting with another $A$ player will receive the benefit of $a$. When $A$ holder plays against a $B$ player, the payoff will be $b$. Similarly, the

$B$ player receives $c$ from the $A$ player and $d$ from another $B$ player. Based on this, denoting that:

$$\Psi := \begin{bmatrix}\Psi_A \\ \Psi_B\end{bmatrix} := \begin{bmatrix}\mathrm{a} & \mathrm{b} \\ \mathrm{c} & \mathrm{d}\end{bmatrix}. \tag{3}$$

We define the state of the population as

$$\xi := [\xi_A, \xi_B]^T, \tag{4}$$

where each component $\xi_x$ $(x = A, B)$ represents the proportion of agents who take the strategy $x$ in the whole population, i.e.

$$\xi \in \Omega_0 := \{\xi | 0 \leq \xi_A \leq 1,\ 0 \leq \xi_B \leq 1,\ \xi_A + \xi_B = 1\}. \tag{5}$$

Thus, the change of each component in the population can be described by the replicator dynamics equation as follows [33]:

$$\dot{\xi_x} = r_x(\xi, \Psi)\,\xi_x, \quad x = A, B, \tag{6}$$

$$r_x(\xi, \Psi) : = \Psi_x\, \xi - \xi^T\, \Psi\, \xi, \tag{7}$$

where $\Psi_x\, \xi$ represents the expected payoff of strategy $x$ and $\xi^T\, \Psi\, \xi$ represents the average payoff of the whole population (also equals to the average payoff of a randomly selected agent). Therefore, the subpopulation of agents whose payoff higher than the average payoff will expand, while the lower ones will decrease. We can analyze the equilibrium points of the above replicator dynamics equations and determine the evolutionarily stable strategy of the system. Note that the replicator dynamics is evolving in the internal of the simplex. If a trajectory comes from the inside of the simplex, it will always keep in the internal of the simplex. Maybe it will converge to the boundary of the simplex, but it will never reach the boundary. Hence, new strategies could not happen.

The interaction among agents determines the fitness of individuals. At each time step $t \in \{0, 1, \cdots\}$, the agent $i$ obtains its fitness value according to the designed fitness function $F_i(\cdot)$ based on the information of the focal agent and its neighbors $j \in K_i$. The fitness function indicates that how agents handle the information. It reflects the adaptation ability of intelligent agents. Design of fitness function $\{F_i(\cdot)\}_{i\in N}$ is one of the main tasks in this control method.

During the evolutionary process, agents evaluate their fitness and then update their strategies. Different updating rules can be used to characterize the evolution of population, such as imitation process [34]–[36] and self-learning process [37]–[40]. In imitation process, the agent $i$ compares its fitness with its neighbors'. In self-learning process, the agent compares its fitness with an aspiration, which is a predetermined baseline. By comparison, the agent updates its strategy in order to improve the fitness. The update process could be either learning the behavior of a better neighbor, randomly switching strategies in the strategy space, or maintaining its current strategy. The design of updating rules is closely related to specific practical problems. It needs to be based not only on the objective of the whole system, but also on the actual constraints of agents, e.g. the communication ability. It is noteworthy that the self-learning updating rule requires less information than imitation. It does not need the fitness information of neighbor agents. In such case, the reasonable aspiration value should be designed to meet the requirements of the system.

Here, we present an adaptive updating rule based on the replicator dynamics equation (6), which is similar to the Win-Stay-Lose-Shift strategy [41], [42]. The state of strategy $x$ at the moment $t$ is defined as follows:

$$\eta(x, t) := \begin{cases} 1, & if\ r_x(\xi, \Psi) \geq 0 \\ 0, & if\ r_x(\xi, \Psi) < 0, \end{cases} \tag{8}$$

where $\eta(x, t) = 1$ and $\eta(x, t) = 0$ mean that the payoff of strategy $x$ is higher or lower than the average payoff, respectively. Agents make decisions based on the following conditional probabilities:

$$\begin{aligned} &Pr(S_i(t) = x \mid \eta(x, t) = 1) = 1 \\ &Pr(S_i(t) = switch(x) \mid \eta(x, t) = 1) = 0 \\ &Pr(S_i(t) = x \mid \eta(x, t) = 0) = 1 + \omega r_x(\xi, \Psi) \\ &Pr(S_i(t) = switch(x) \mid \eta(x, t) = 0) = -\omega r_x(\xi, \Psi), \end{aligned} \tag{9}$$

where $switch(x) = B$ if $x = A$; otherwise, $switch(x) = A$ if $x = B$. If the payoff is higher than the average payoff (i.e. $\eta(x, t) = 1$), the agent keeps its current strategy unchanged; otherwise (i.e. $\eta(x, t) = 0$), the agent switches its strategy with probability $-\omega\, r_x(\xi, \Psi)$. $\omega > 0$ represents the selection intensity. It reflects the influence of the difference between fitness on the individual decision-making [36], [43], [44], and satisfies $-\omega\, r_x(\xi, \Psi) < 1$. Thus, after one time step, the state of population changes:

$$\begin{bmatrix}\xi_A(t+1) \\ \xi_B(t+1)\end{bmatrix} = \begin{cases} \begin{bmatrix}\xi_A(t) - \omega\, r_B(\xi(t), \Psi)\,\xi_B(t) \\ \xi_B(t) + \omega\, r_B(\xi(t), \Psi)\,\xi_B(t)\end{bmatrix} \\ if\ \eta(A, t) = 1, \eta(B, t) = 0, \\ \begin{bmatrix}\xi_A(t) + \omega\, r_A(\xi(t), \Psi)\,\xi_A(t) \\ \xi_B(t) - \omega\, r_A(\xi(t), \Psi)\,\xi_A(t)\end{bmatrix} \\ if\ \eta(A, t) = 0, \eta(B, t) = 1. \end{cases} \tag{10}$$

Based on the above model, the system evolves while agents adjust their behaviors according to their own interests. Through reasonably designing all the elements in the game, the system can achieve the corresponding control objectives. In the following, we take the problem of division of labor for example to demonstrate how to realize the coordinated control of multi-agent systems by utilizing our framework.

## III. EVOLUTIONARY GAME CONTROL METHOD FOR DIVISION OF LABOR

### A. PROBLEM BACKGROUND

The applications of multi-agent systems often involve highly complex tasks where agents of different types or skills are required to work together. In social groups, this kind of cooperation for complex tasks is common, such as the division of labor in the human society [45]. In the animal world, such division of labor is also prevalent. For example, lionesses act as different roles spontaneously when they hunt together [46]. Two lionesses, the wings, attack a group of prey from either side panicking them to run forward. They run right into one or two other lionesses, positioned as centres, who are waiting for them. This kind of group hunt is highly successful. It is not possible with only one role of participants, but it is better with more roles. Similar cooperative hunting is also seen in other species, such as chimpanzee in the forest and wild dogs in Africa. In the insects population, the division of different types of work is common in the aggregation of bees and ants. In the artificial complex system, especially the swarm robotic systems, the division of labor is highly required [47]–[49]. For example, target tracking (similar to group hunting) [50], robot soccer race [51], maintenance of mechanical components with multiple steps, disaster detection and rescue, and etc.

It is worth noting that the characteristics of such examples are that the assigned tasks must be completed by different roles together, either role is unable to complete the task alone. This requires the agents in the system to spontaneously form two or more subgroups with different actions and cooperate with each other. Through specialized division of labor, the objective of the system could be completed. When dealing with such a coordinated control problem, an important task is how to divide the population on a complex connection network in order to achieve an effective strategy distribution and enable agents with different strategies to combine organically. For instance, a multi-agent system is built to complete a machine assembly task with two steps. It is necessary to put agents at the neighbor position of their complementary counterparts. A reasonable distribution helps to improve the overall efficiency of completing the assembly task.

In the following, we study how to realize the division of labor in multi-agent systems. Our aim is that different types of agents are distributed as evenly as possible around the complementary individuals.

### B. MATHEMATICAL DESCRIPTION

We consider a multi-agent system with a population size of $n$ ($N = \{1, 2, \cdots, n\}$), where agents are denoted as $\{Agent_i | i \in N\}$. Based on the design method described in Section II, we define the agent as $Agent_i = (S_i, B_i, C_i, F_i)$. The state, $S_i$, represents the action taken by the agent. Here the state space consists of two different behaviors, $S_i \in \{A, B\}$. The agents can choose strategy $A$ or $B$. The behavior, $B_i$, is determined by the evaluation of fitness. Agents make decisions in probability based on the comparison of payoffs among agents. In this model, we only consider the behavior choice of agents, but not pay attention to the change of the connection relationship between individuals. Thus the connection among agents is represented as the predefined static communication topology. The communication, $C_i$, denotes the neighbor set of $Agent_i$, that is, $C_i = \{j \in N | a_{ij} > 0\}$, where $a_{ij}$ is the element in the adjacency matrix $A = [a_{ij}] \in R^{n\times n}$. When $a_{ij} > 0$, $Agent_i$ can obtain the information of $Agent_j$, otherwise $a_{ij} = 0$.

The control objective of the division of labor problem is to make the agent and its neighbor show different states. Therefore, when the agents play games with their neighbors, they should obtain higher payoffs when their strategies are different than they hold the same strategy. Hence, we select the $2\times 2$ game whose Nash equilibria require each participant to take the opposite strategy of its counterpart. Thus, it should satisfy $a < c$ and $b > d$ in the payoff matrix (2). Snowdrift game (or hawk-dove game) [52], [53] may satisfy such condition, since the evolutionarily stable strategy of players in such game is to take the opposite strategy of its opponent.

We denote $\pi_{ij}$ as the payoff of $Agent_i$ when playing game with its neighbor $Agent_j$ ($j \in C_i$). Then the fitness of $Agent_i$ is ($k$ is the number of its neighbors):

$$F_i = \frac{1}{k} \sum_{j\in C_i} (a_{ij}\, \pi_{ij}). \tag{11}$$

We employ Fermi process as the updating rule. Individuals compare their fitness with their neighbors' average fitness ($\overline{F}$) and update their strategies. Here, we assume that the initial state of the population is randomly assigned and do not consider mutation or exploration of strategy. Thus, the probability for $Agent_i$ switching its state $S_i$ in the state space $\{A, B\}$ is: $[1 + \mathrm{e}^{-\omega(\overline{F}-F_i)}]^{-1}$, where $\omega$ is the selection intensity. Such stochastic learning rule is widely used in biological evolution and social learning [54], [55]. During the evolutionary process, we adopt asynchronous updating rules, that is, all the agents update their strategies successively at the same time step.

To show the effect of control, we define an indicator $J$, which represents the proportion of connection pairs with different strategies among all the connections in the topology network.

$$J = \frac{\Sigma_{i=1}^{n} k_Y}{\Sigma_{i=1}^{n} k}, \tag{12}$$

where

$$k_Y = \begin{cases} k_B & \text{if } S_i = A; \\ k_A & \text{if } S_i = B. \end{cases} \tag{13}$$

$k_A$ and $k_B$ are denoted as the number of $A$ and $B$ holders in one's neighbors, respectively. Obviously, $J \in [0, 1]$, and the closer $J$ approaches 1, the better the effect of division of labor is. In the case of limit, if $\lim_{t\to\infty} J = 1$ holds, it means that the system achieves a perfect division of labor.

## IV. THEORETICAL ANALYSIS

The objective of division of labor problem is to find out the conditions under which the system achieves an evenly distribution of agents with different strategies in the topology network. To further understand the evolution of the system, we discuss the evolutionary dynamics in the population endowed with two strategies.

We study a population of $n$ agents, all of which are distributed on the vertices of the communication topology, which is a static undirected graph. Each agent chooses one of the two strategies, $A$ and $B$, as its initial strategy. The game is played between the focal agent and all its neighbors. Its fitness is defined as the average of the payoffs obtained from all the games in which it is involved. The payoffs gained by each player are shown in the matrix (2).

Based on the system which is characterized by (3)-(7), evolution trends of the system can be analyzed. By studying the equilibrium points of the replicator dynamics (6), we can obtain the conditions under which the system is in a stable state. Several prominent examples of two-player games, which are motivated by many real-world systems and widely used in various disciplines, are considered. Among them are the prisoner's dilemma game, the snowdrift game, and the coordination game.

Further, we analyze the games in structured populations and obtain the conditions for the two strategies to be evenly distributed on the network.

### A. ONE STRATEGY DOMINATES THE OTHER

First, we discuss the situation in which one strategy is dominant in the game.

*Theorem 1:* For any given initial strategy distribution, $\xi_x \in (0, 1)$ $(x = A, B)$, if the condition $(a - c)(b - d) > 0$ is satisfied, then the system will evolve into the state of full of one strategy. Moreover,

(i) if $a < c$ and $b < d$, the system evolves into the state of full $B$, $\xi_A = 0$ and $\xi_B = 1$;

(ii) if $a > c$ and $b > d$, the system evolves into the state of full $A$, $\xi_A = 1$ and $\xi_B = 0$.

*Proof:* From (6), the fractions of strategies $A$ and $B$ follow the equations:

$$\begin{cases} \dot{\xi_A} = r_A(\xi, \Psi)\,\xi_A, \\ \dot{\xi_B} = r_B(\xi, \Psi)\,\xi_B. \end{cases} \tag{14}$$

By substituting $\xi_A + \xi_B = 1$ and inserting (3), (4) and (7) into (14),

$$\dot{\xi_A} = \xi_A(1 - \xi_A)[(a - b - c + d)\xi_A + (b - d)]. \tag{15}$$

When $a < c$ and $b < d$, $b - d < 0$. Then, if $\xi_A$ approaches zero from the right in the small neighborhood of zero, $\dot{\xi_A} < 0$. Therefore, state $\xi_A = 0$ is stable equilibrium. However, owing to $a - c < 0$, when $\xi_A$ approaches 1 from the left in the small neighborhood of 1, $\dot{\xi_A} < 0$. Hence, state $\xi_A = 1$ is unstable. For $\xi_A = \frac{d-b}{a-b-c+d}$, if $a - b - c + d > 0$, $\xi_A = \frac{d-b}{d-b+a-c} = 1 + \frac{c-a}{d-b+a-c} > 1$. Since $\xi_A \in [0, 1]$ is assumed, then $\xi_A = \frac{d-b}{d-b+a-c}$ is not the equilibrium. Thus, $(a-b-c+d)\xi_A + (b-d) < 0$ holds strictly. Hence, $\xi_A = 0$ is the unique stable equilibrium. Under such case, the system evolves into the state of full $B$.

Similarly, when $a > c$ and $b > d$, if $\xi_A$ approaches 1 from the left in the small neighborhood of 1, $\dot{\xi_A} > 0$, so $\xi_A = 1$ is the stable equilibrium. While $\xi_A$ approaches zero from the right in the small neighborhood of zero, $\dot{\xi_A} > 0$, namely $\xi_A = 0$ is unstable. Under this case, the system evolves into the state of full $A$. ∎

*Remark 1:* Theorem 1 depicts the population's decision-making behavior in the evolutionary dynamics when there exists a dominant strategy in the game depicted by (2). It indicates that if $(a - c)(b - d) > 0$, two different strategies can never coexist stably in such a system for any initial state and selection intensity. When $a < c$ and $b < d$, the payoff of $B$ strategy agent is higher than that of $A$ strategy player. Thus, in order to achieve higher fitness, all the agents would rationally adopt the $B$ strategy in the last. Under this case, the full $B$ state is the Nash equilibrium. The result is consistent with the classical game theory. The most typical example of such game is the prisoner's dilemma [56], in which defection ($B$ strategy here) is the only one pure Nash equilibrium.

### B. TWO STRATEGIES COEXIST

Next, we consider $a < c$ and $b > d$ in the payoff matrix (2). The game types represented by this matrix include the snowdrift game, hawk-dove game and chicken game [30], [52], [53]. The snowdrift game can be described as follows. We assume that two drivers are caught with their cars in a snowdrift on the way home. Each individual can choose whether or not to cooperate in shoveling a way out. Those who cooperate share the workload. Those who do not cooperate may take a rest while the others do the shoveling. When the snow is removed, everyone will manage to get to their destination. If one of them refuses to cooperate, the other driver is better off to cooperate unilaterally, rather than spend the night freezing. Hawk-dove game comes from a situation where animals fight for a territory: "Hawks" escalate the fight, risking serious injury, whereas "doves" flee when the opponent escalates. In the chicken game, two cars are heading for a collision. The loser chickens out, while the winner stays on track. Big loss occurs when both stay on track.

*Theorem 2:* For any given initial strategy distribution, $\xi_x \in (0, 1)$ $(x = A, B)$, if the condition $a < c$ and $b > d$ can be satisfied, the system will converge to the unique interior equilibrium. That is, $\lim_{t\to\infty} \xi_A = \frac{d-b}{a-b-c+d}$ holds.

*Proof:* Based on (15), the stability of equilibria, $\xi_A = 0$, $\xi_A = 1$ and $\xi_A = \frac{d-b}{a-b-c+d}$, can be analyzed. Owing to $a < c$ and $b > d$, we have $b-d > 0$ and $a-b-c+d < 0$. When $\xi_A$ approaches zero from the right side in the small neighborhood of zero, $\dot{\xi_A} > 0$. Therefore, state $\xi_A = 0$ is unstable equilibrium. Similarly, when $\xi_A$ approaches 1 from the left side in the small neighborhood of 1, $\dot{\xi_A} < 0$. Hence, state $\xi_A = 1$ is also unstable. For $\xi_A = \frac{d-b}{a-b-c+d} = \frac{b-d}{(b-d)+(c-a)}$, it belongs to $(0, 1)$. When $\xi_A$ approaches such equilibrium point from the right side in its small neighborhood, $\dot{\xi_A} < 0$. While $\xi_A$ approaches it from the left side in its small neighborhood, $\dot{\xi_A} > 0$. Therefore, $\xi_A = \frac{d-b}{a-b-c+d}$ is the stable equilibrium. Moreover it is the unique equilibrium of the system. ∎

*Remark 2:* According to the Theorem 2, when $a < c$ and $b > d$, the two different strategies have the possibility of coexistence in the system. Such payoff matrix of the game predicts the Nash equilibria as $(A, B)$ and $(B, A)$. That is, agents tend to choose strategies that are contrary to their co-players.

Different from the prisoner's dilemma game, where defection is a best reply no matter whether the co-player is cooperator or not. In the snowdrift game mentioned above, each strategy is a best reply to the other. The analysis in classical game theory shows that the Nash equilibria for the snowdrift game are: when your counterpart chooses to cooperate, you'd better choose defect; when the opponent chooses to defect, you'd better choose cooperate.

### C. TWO STRATEGIES ARE BISTABLE

Subsequently, we study the case of $a > c$ and $b < d$, which represents the coordination game [57]. A typical case for a coordination game is choosing the sides of the road upon which to drive. Two drivers meet on a narrow road. Both have to swerve in order to avoid a head-on collision. But if they choose differing maneuvers they will collide. Another type of coordination game commonly called battle of the sexes. In this game couples prefer engaging in the same activity over going alone, but their preferences differ over which activity they should engage in. The French philosopher, Jean Jacques Rousseau, presented a typical coordination game, called the stag hunt game [46]. Two hunters can either jointly hunt a stag or individually hunt a hare (less benefit). Hunting stags is quite challenging and requires mutual cooperation. If either hunts a stag alone, the chance of success is minimal. Hunting stags is most beneficial for society but requires a lot of trust among its members.

*Theorem 3:* If $a > c$ and $b < d$ are satisfied, the system will converge to one of the two equilibria: 0 and 1. Moreover, the initial strategy distribution influences the final state of the system.

(i) If initial fraction of strategy $A$ satisfies $0 < \xi_A < \frac{d-b}{a-b-c+d}$, $\lim_{t\to\infty} \xi_A = 0$ holds. The system evolves into the state of full $B$;

(ii) If initial fraction of strategy $A$ satisfies $1 > \xi_A > \frac{d-b}{a-b-c+d}$, $\lim_{t\to\infty} \xi_A = 1$ holds. The system evolves into the state of full $A$.

*Proof:* If $a > c$ and $b < d$, we have $a - b - c + d > 0$ and $b - d < 0$. Based on (15), we can study the stability of equilibria: $\xi_A = 0$, $\xi_A = 1$ and $\xi_A = \frac{d-b}{a-b-c+d}$. If $\xi_A$ approaches zero from the right side in the small neighborhood of zero, $\dot{\xi_A} < 0$. Therefore, state $\xi_A = 0$ is a stable equilibrium. Similarly, when $\xi_A$ approaches 1 from the left side in the small neighborhood of 1, $\dot{\xi_A} > 0$. Hence, state $\xi_A = 1$ is also stable. For $\xi_A = \frac{d-b}{(a-c)+(d-b)}$, when $a > c$ and $d > b$, it belongs to (0, 1). When $\xi_A$ approaches such equilibrium point from the right side in its small neighborhood, $\dot{\xi_A} > 0$. While $\xi_A$ approaches it from the left side in its small neighborhood, $\dot{\xi_A} < 0$. Hence, $\xi_A = \frac{d-b}{a-b-c+d}$ is an unstable equilibrium. When initial $\xi_A > \frac{d-b}{a-b-c+d}$, since $\dot{\xi_A} > 0$, the system evolves towards equilibrium $\xi_A = 1$. When initial $\xi_A < \frac{d-b}{a-b-c+d}$, owing to $\dot{\xi_A} < 0$, the system evolves towards equilibrium $\xi_A = 0$. ■

*Remark 3:* When $a > c$ and $b < d$, the best strategy choice for an agent is playing the same strategy with their co-player. Thus, the state of full $A$ and full $B$ are bistable. According to the Theorem 3, two strategies can not coexist in the system. The state with full of only one kind of strategy would be reached. Besides, if the updating rules are not limited to only imitating the strategies existing in the population, the initial strategy distribution cannot totally decide the final state of the system.

In the coordination game, or the stag-hunt game, there are two pure strategy equilibria. Both players prefer one equilibrium which Pareto dominates the other. However, the inefficient equilibrium is less risky as the payoff variance over the other player's strategies is lower. Specifically, one equilibrium is payoff-dominant while the other is risk-dominant. Suppose $B$ is risk dominant, which means that $a+b < c+d$. But $A$ is Pareto efficient, which means that $a > d$. Hence, strategy $B$ has the bigger basin of attraction, but a homogeneous population of $A$ has a higher payoff than that of $B$.

### D. THE CONDITION FOR STRATEGIES COEXISTING IN EQUAL PROPORTIONS ON THE NETWORK

In the following, we study the coexistence conditions of two strategies on the network structures. We utilize the method of pairwise approximation [58], [59] and perturbation theory [60] to analyze the expected proportion of strategy $A$ and $B$ respectively in the structured population. Using the pair approximation, a simplest decoupling approximation to take account of spatial correlation, we can obtain analytical results for stationary densities, and critical parameters for equilibrium.

*Theorem 4:* If the network structured system evolves to the equilibrium 1/2, namely $\lim_{t\to\infty} \xi_A = \lim_{t\to\infty} \xi_B = 1/2$, which means two strategies coexisting in the system and having the same fraction in the population, the following conditions should be satisfied:

(i) $a < c$ and $b > d$;

(ii) $a + b = c + d$.

*Proof:* The pair approximation technique, models space implicitly, by focusing on the interaction between nearest neighbors and tracking the dynamics of pairs of neighbors instead of single individuals.

Denoting $p_A$ and $p_B$ as the frequencies of strategy $A$ and $B$ in the population; $p_{AA}$, $p_{AB}$, $p_{BA}$ and $p_{BB}$ as the frequencies of pairs of two neighbor agents, $AA$, $AB$, $BA$ and $BB$, respectively. Let $q_{X|Y}$ denote the conditional probability to find an $X$-player given that the adjacent node is occupied by a $Y$-player. Here, both $X$ and $Y$ stand for $A$ or $B$.

The identities

$$\begin{aligned} &p_A + p_B = 1 \\ &q_{A|X} + q_{B|X} = 1 \\ &p_{XY} = q_{X|Y} \times p_Y \\ &p_{AB} = p_{BA}, \end{aligned} \tag{16}$$

imply that the whole system can be described by only two variables, $p_A$ and $p_{AA}$, in pair approximation.

The agent plays game with all its neighbors according to payoff matrix (2). The fitness of agents with $A$ and $B$ strategy are denoted as $F_A$ and $F_B$, respectively:

$$\begin{aligned} F_A &= \tfrac{1}{k}(k_A a + k_B b) \\ F_B &= \tfrac{1}{k}(k_A c + k_B d), \end{aligned} \tag{17}$$

where $k_A$ and $k_B$ ($k_A + k_B = k$) represent the number of individuals in the agent's neighborhood holding strategy $A$ and $B$, respectively.

Next, we calculate the probabilities that variables $p_A$ and $p_{AA}$ change during one time step. A $B$ strategy agent is randomly selected from the entire population with probability $p_B$. Therefore $p_A$ increase by $1/n$ with probability:

$$Prob(\triangle p_A = \frac{1}{n}) = p_B \sum_{k_A+k_B=k} \frac{k!}{k_A!\,k_B!} \frac{q_{A|B}^{k_A} q_{B|B}^{k_B}}{1+\mathrm{e}^{-\omega(\overline{F}-F_B)}}. \tag{18}$$

Regarding pairs, the number of $AA$-pairs increases by $k_A$ and therefore $p_{AA}$ increases by $k_A/(kn/2)$ with probability:

$$Prob(\triangle p_{AA} = \frac{2k_A}{kn}) = p_B \frac{k!}{k_A!\,k_B!} \frac{q_{A|B}^{k_A} q_{B|B}^{k_B}}{1+\mathrm{e}^{-\omega(\overline{F}-F_B)}}. \tag{19}$$

Similarly, an $A$ player is selected with probability $p_A$. Therefore $p_A$ decreases by $1/n$ with probability:

$$Prob(\triangle p_A = -\frac{1}{n}) = p_A \sum_{k_A+k_B=k} \frac{k!}{k_A!\,k_B!} \frac{q_{A|A}^{k_A} q_{B|A}^{k_B}}{1+\mathrm{e}^{-\omega(\overline{F}-F_A)}}. \tag{20}$$

Regarding pairs, the number of $AA$-pairs decreases by $k_A$ and therefore $p_{AA}$ decreases by $k_A/(kn/2)$ with probability:

$$Prob(\triangle p_{AA} = -\frac{2k_A}{kn}) = p_A \frac{k!}{k_A!\,k_B!} \frac{q_{A|A}^{k_A} q_{B|A}^{k_B}}{1+\mathrm{e}^{-\omega(\overline{F}-F_A)}}. \tag{21}$$

Let us suppose that one replacement event takes place in one unit of time. When the population size is much larger than the average degree of the network and under weak selection, the time derivatives of $p_A$ and $p_{AA}$ are given by:

$$\begin{aligned} \dot{p_A} &= \frac{1}{n} Prob(\triangle p_A = \frac{1}{n}) + (-\frac{1}{n}) Prob(\triangle p_A = -\frac{1}{n}) \\ &= \frac{p_B - p_A}{2n} + \omega \times \frac{1}{4n}[\overline{F}(p_B - p_A) - c\,p_{AB} \\ &\quad - d\,p_{BB} + a\,p_{AA} + b\,p_{BA}] + O(\omega^2), \end{aligned} \tag{22}$$

and

$$\begin{aligned} \dot{p_{AA}} &= \sum_{k_A=0}^{k} \frac{2k_A}{kn} Prob(\triangle p_{AA} = \frac{2k_A}{kn}) \\ &\quad + \sum_{k_A=0}^{k} (-\frac{2k_A}{kn}) Prob(\triangle p_{AA} = -\frac{2k_A}{kn}) \\ &= \frac{p_{AB} - p_{AA}}{n} + \frac{\partial p_{AA}}{\partial \omega}\Big|_{\omega=0} \omega + O(\omega^2), \end{aligned} \tag{23}$$

where

$$\begin{aligned} \frac{\partial p_{AA}}{\partial \omega}\Big|_{\omega=0} &= \sum_{k_A=0}^{k} \frac{2k_A}{kn} p_B \frac{k!}{k_A!\,k_B!} q_{A|B}^{k_A} q_{B|B}^{k_B} \frac{\overline{F} - F_B}{4} \\ &\quad - \sum_{k_A=0}^{k} \frac{2k_A}{kn} p_A \frac{k!}{k_A!\,k_B!} q_{A|A}^{k_A} q_{B|A}^{k_B} \frac{\overline{F} - F_A}{4} \\ &= \frac{1}{2n}[\overline{F}(p_{AB} - p_{AA}) - c p_{AB} q_{A|B} \frac{k-1}{k} \\ &\quad - c p_{AB} \frac{1}{k} - d p_{AB} q_{B|B} \frac{k-1}{k} + a p_{AA} q_{A|A} \frac{k-1}{k} \\ &\quad + a p_{AA} \frac{1}{k} + b p_{AA} q_{B|A} \frac{k-1}{k}]. \end{aligned} \tag{24}$$

For $\omega = 0$, the equilibrium of (22)-(23), $(p_A^*(\omega), p_{AA}^*(\omega))$, is given by $(1/2, 1/4)$. By (16), we further have $p_A^*(\omega) = p_B^*(\omega) = 1/2$ and $p_{AA}^*(\omega) = p_{AB}^*(\omega) = p_{BA}^*(\omega) = p_{BB}^*(\omega) = 1/4$. According to the perturbation theory, the equilibrium $(p_A^*(\omega), p_{AA}^*(\omega))$ always exists under weak selection. Thus, there exists function $f_{XY}$, $X, Y \in \{A, B\}$, which is only dependent on payoff entries, degree of the communication topology graph and population size, such that:

$$\begin{aligned} p_{AA}^*(\omega) &= \frac{1}{4} + \omega \times f_{AA} \\ p_{AB}^*(\omega) &= \frac{1}{4} + \omega \times f_{AB} \\ p_{BA}^*(\omega) &= \frac{1}{4} + \omega \times f_{BA} \\ p_{BB}^*(\omega) &= \frac{1}{4} + \omega \times f_{BB}. \end{aligned} \tag{25}$$

Inserting (25) into (22), it leads to

$$\begin{aligned} \dot{p_A} = \frac{1-2p_A}{2n} + \omega \times \frac{1}{4n}[\overline{F}(1-2p_A) \\ - \frac{1}{4}(c - d + a + b)] + O(\omega^2). \end{aligned} \tag{26}$$

The equilibrium can be obtained from $\dot{p_A}$. Since the selection intensity $\omega$ is weak, we neglect the higher order term $O(\omega^2)$. Thus the equilibrium $p_A^*(\omega)$ fulfills $\frac{1-2p_A^*(\omega)}{2n} + \omega \times \frac{1}{4n}[\overline{F}(1 - 2p_A^*(\omega)) - \frac{1}{4}(c - d + a + b)] = 0$. Therefore

$$p_A^*(\omega) = \frac{1}{2} + \frac{\omega(a+b-c-d)}{8(2+\omega\overline{F})}. \tag{27}$$

If we choose the entries in the payoff matrix as satisfying $a + b = c + d$, then $p_A^*(\omega) = 1/2$.

To achieve the objective that agents with two different strategies coexist with the same proportion in the system, according to Theorem 2, the entries in the payoff matrix should satisfy $a < c$ and $b > d$. ∎

*Remark 4:* Theorem 4 gives the characteristics of a game matrix that enables a system to evolve into two strategies

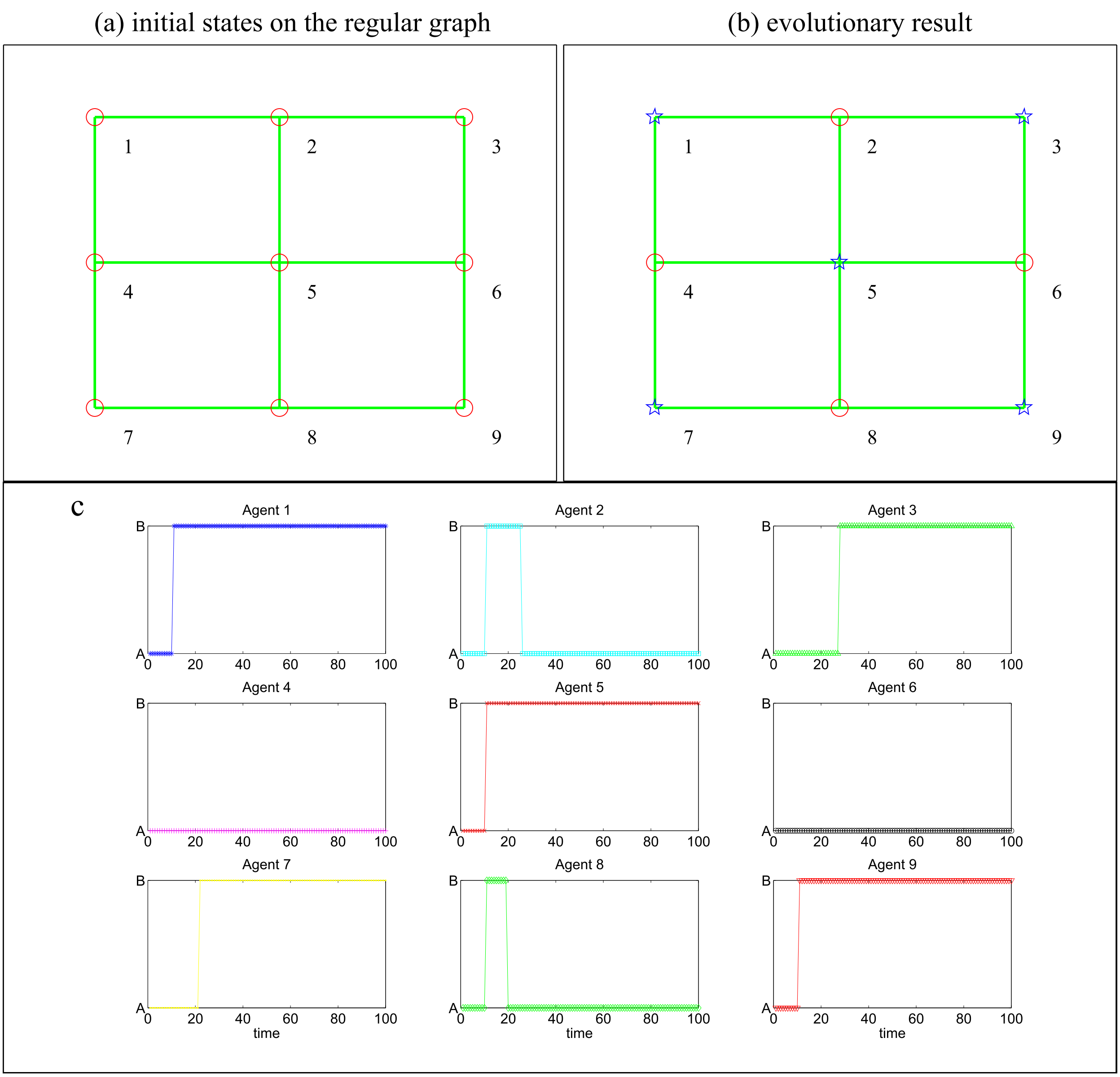


**FIGURE 1.** Simulation results on the regular graph.

coexisting and having the same fraction. The two conditions correspond to the following functions. (i) The two strategies can coexist, corresponding to $a < c$ and $b > d$; (ii) At the equilibrium point, the proportions of the two strategies are the same, corresponding to $a + b = c + d$. Therefore we can choose reasonable entry values of the payoff matrix to build a game both satisfying these conditions to achieve the basic conditions required by the problem of division of labor. However, even if the above conditions are met, it is not always sufficient to ensure that the networked system can evolve to the state that all the agents have the opposite state with their neighbors. More factors, such as randomness and network structures, need to be taken into account to regulate a perfect strategy allocation.

## V. SIMULATION

We verify the effectiveness of the proposed evolutionary game method for multi-agent systems through simulation. We first construct the system by using the conditions derived from the theoretical analysis section, and verify the feasibility of the related theories on the regular graph and the general connected graphs respectively. In particular, for a special class of connected graphs, it is impossible to completely implement all the neighbor nodes holding different strategies, and the proposed method can also give satisfactory results. For heterogeneous networks (scale-free networks), which are very common and widely used in practice, we compare the evolutionary results of several classical game types, including prisoner's dilemma, coordination game and also snowdrift

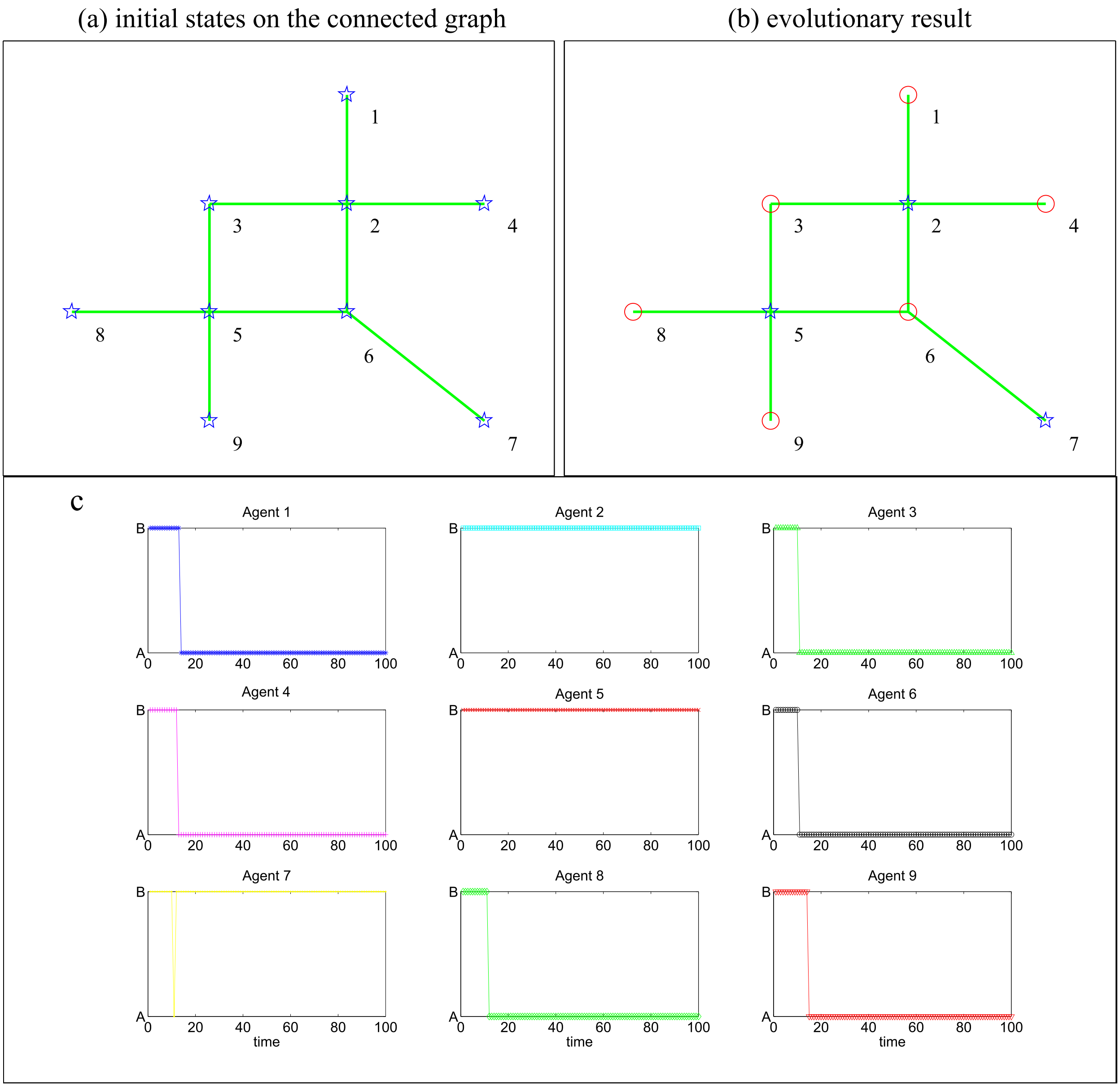


**FIGURE 2. Simulation results on the general connected graph.**

game. It is demonstrated that the proposed theoretic method can effectively guide the selection of game types and payoff matrices, so as to achieve an effective division of labor.

### A. REGULAR GRAPH

Firstly, we consider the case of the regular graph. For a multi-agent system of population size $n = 9$, agents are denoted as $Agent_i = (S_i, B_i, C_i, F_i)$ $(i \in N = \{1, 2, \cdots, 9\})$. The state $S_i \in \{A, B\}$, where the agent has two strategies to choose, $A$ or $B$. The behavior $B_i$, is based on the difference between one's payoff and the average payoff of its neighbors. The communication $C_i$ is denoted by the neighbor set of $Agent_i$, that is, $C_i = \{j \in N | a_{ij} = 1\}$, where $a_{ij}$ is the element in the adjacency matrix $A = [a_{ij}] \in R^{n \times n}$. In particular, for the $a_{ij} > 0$ in the adjacency matrix, we simply consider $a_{ij} = 1$. When $a_{ij} = 1$, $Agent_i$ can obtain the information of $Agent_j$; otherwise $a_{ij} = 0$. All the agents are put on the vertex of a regular graph with a von Neumann neighborhood (4 neighbors lattice), and the adjacency matrix $A$ is given as follows:

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}. \quad (28)$$

The control objective is to put the agents with different states as close as possible. According to the theoretical

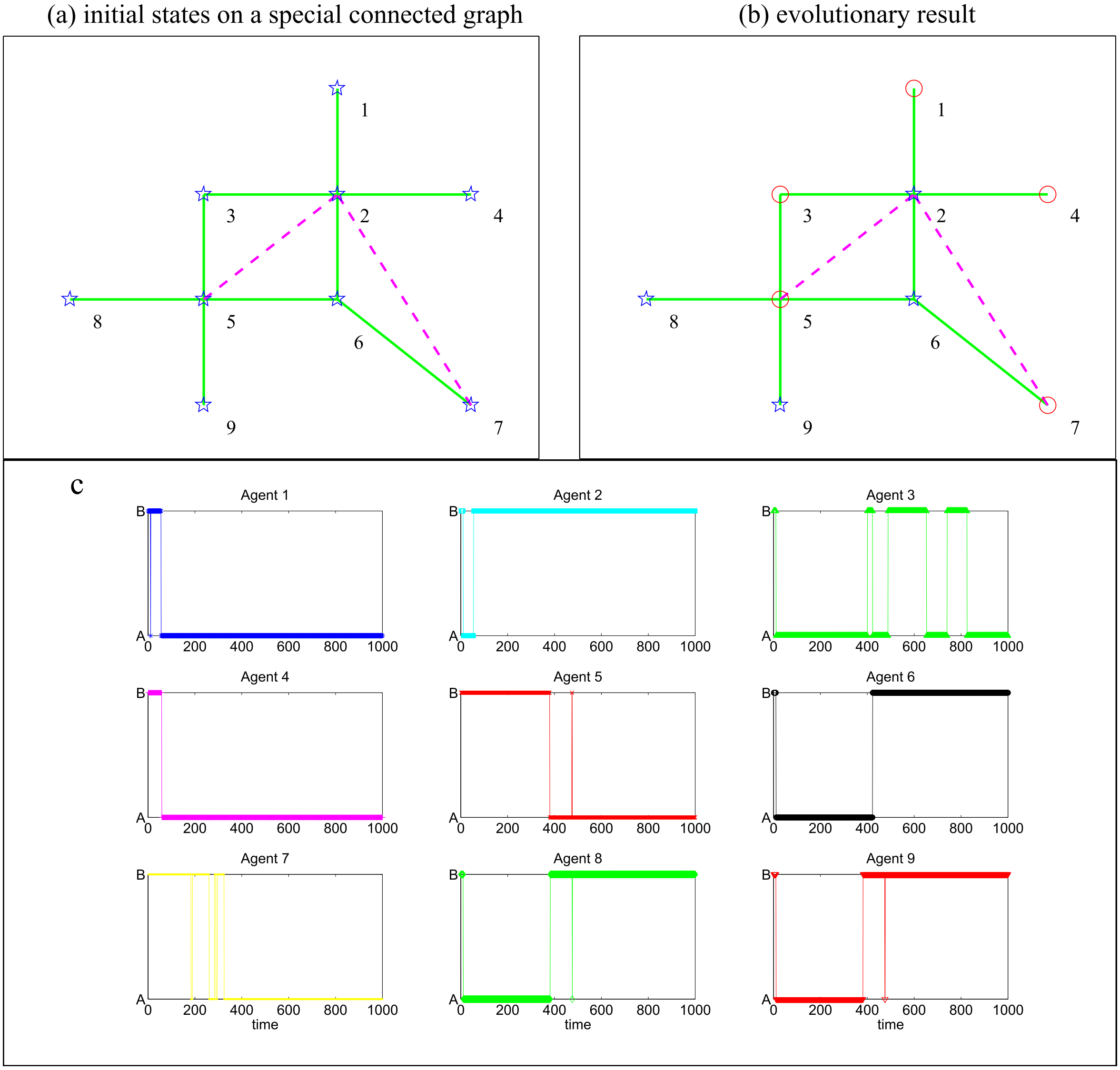


**FIGURE 3. Simulation results on a special connected graph.**

analysis, we select the payoff matrix as follows:

$$\begin{array}{c|cc} & A & B \\ \hline A & 0 & 1 \\ B & 1 & 0 \end{array}. \tag{29}$$

This simple payoff matrix fulfills the condition that $a + b = c + d$ and $a < c$, $b > d$. It reflects that agents obtain higher payoff when choosing the opposite strategy from their opponents.

The evolutionary results are shown in Fig. 1, in which Fig. 1($a$) represents the communication topology and an initial population distribution. Fig. 1($b$) shows the evolutionary result (the distribution of agents on the graph) of the two strategies (the red circle represents the strategy $A$, the blue star represents the strategy $B$). Fig. 1($c$) shows the changes of state of each agent with time. Finally, the state of the agents in the system evolves into an optimal distribution, and each agent holds a different strategy from its neighbors. The control objective is achieved.

### B. GENERAL CONNECTED GRAPH

Secondly, we consider the case of a general connected graph. The system still contains 9 agents. The communication topology of the system is represented by the following adjacency matrix:

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{30}$$

In this case, the connection between agents is no longer completely regular. The evolutionary results are shown in Fig. 2, where Fig. 2($a$) represents the communication topology and initial population distribution. Fig. 2($b$) shows

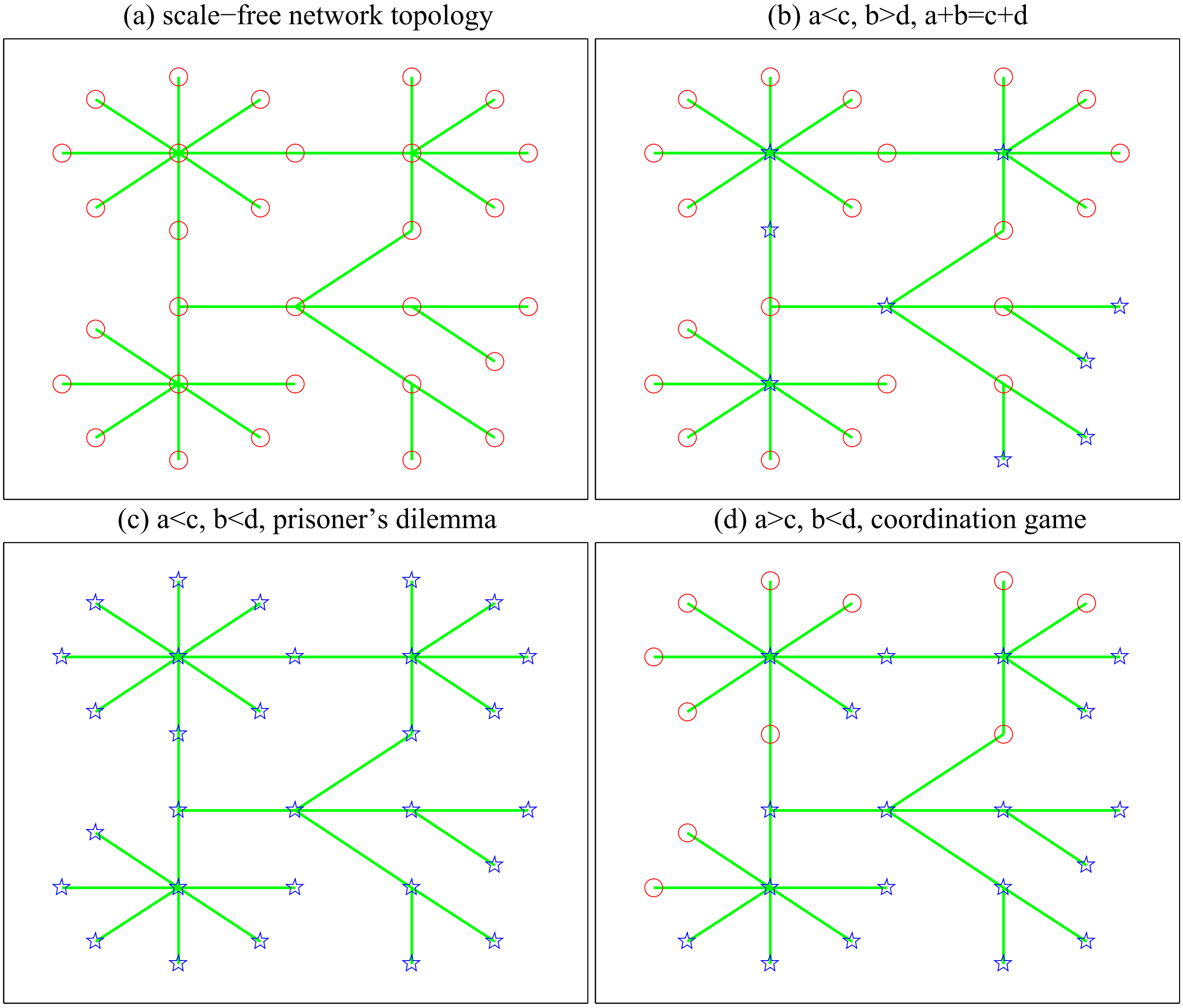


**FIGURE 4.** **Simulation results on the scale-free network.**

the evolutionary result of the two strategies (the red circle represents the strategy $A$, the blue star represents the strategy $B$). Fig. 2($c$) shows the changes of state of each agent with time. The results show that the control objectives can also be achieved.

Thirdly, we consider a more complex case of topology on the basis of the previous example. In this case, two connection edges (pink dotted line) are added on the basis of the topology in Fig. 2($a$), which leads to cycles and some nodes with odd connections. The adjacency matrix is shown as follows:

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}. \quad (31)$$

It is found that the system in this situation could not evolve into a perfect state as in regular graphs where agents with different strategies connect each other. But in this case, our evolutionary game approach can still give an approximately even distribution result (see Fig. 3). Fig. 3($a$) represents communication topology and initial population distribution. Fig. 3($b$) shows the evolutionary result of the two strategies (the red circle represents the strategy $A$, the blue star represents the strategy $B$). Fig. 3($c$) shows the trends of state change of each agent with time. Obviously, the result is not stable. If different connections can be assigned with different weights (i.e. non-zero entries in adjacency matrix $A$ not only equal to 1), it may reinforce the state of relatively important pair of nodes.

### C. SCALE-FREE NETWORK

Next, we study a kind of heterogeneous network topology, scale-free network, which is more common in practice [61], [62]. In order to reflect the validity of the selection conditions obtained from our theoretical analysis, we compare the evolutionary results of several different game types. Fig. 4($a$) depicts the initial scale-free network topology. Such kind of network is highly uneven and irregular. Most of its nodes

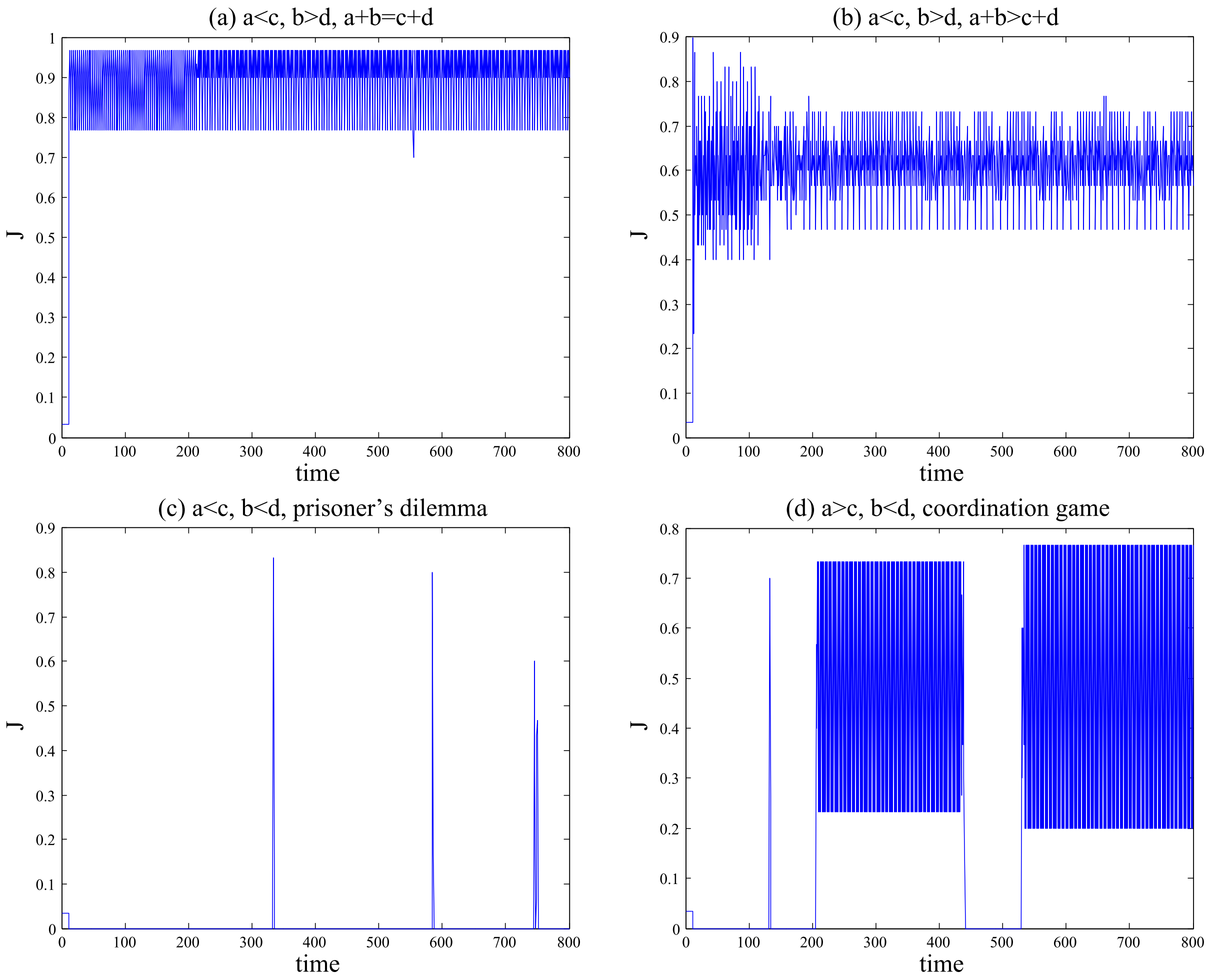


**FIGURE 5.** **Control index *J* for different game matrices.**

have a small connection degree while a few hub nodes have a large degree. Hub nodes form several clusters in the network. The characteristic path length between nodes and clustering coefficient of the scale-free network are quite different from those of the regular graph.

Fig. 4($b$)-($d$) show the results of evolution by choosing different types of game matrices. Obviously, Fig. 4($b$), based on the conditions we obtained in the theoretical analysis section ($a < c$, $b > d$ and $a + b = c + d$), presents the best division of labor. The result in Fig. 4($c$) is obtained when $a < c$ and $b < d$ are adopted (matrix (32)), which represents the prisoner's dilemma game.

| | $A$ | $B$ |
|---|---|---|
| $A$ | 3 | 0 |
| $B$ | 5 | 1 |

(32)

It is shown that all the agents in the population have evolved to strategy $B$, which is the dominant strategy in prisoner's dilemma. Owing to the existence of a strategy which is completely dominant, it is difficult for such game type to achieve division of labor. The result of the case $a > c$ and $b < d$ is shown in Fig. 4($d$). Matrix (33) is a coordination game and both strategies are Nash equilibria.

| | $A$ | $B$ |
|---|---|---|
| $A$ | 5 | 0 |
| $B$ | 3 | 1 |

(33)

As a result, the system can not be stable in the coexistence of strategies, and always evolves towards the direction of full $A$ or full $B$. In this case, it is also difficult to reach an effective division of labor.

Because of the complexity of the network, it is usually difficult to make all the neighbors present different strategic states in practical problems. This requires us to utilize a control index to evaluate the effect of division of labor of the system. The index $J$ presented in (12) represents the proportion of strategy pairs constituted by different strategies to the total number of connections in the network. Fig. 5 shows the change of $J$ in different game types. Fig. 5($a$) corresponds to the case that $a < c$, $b > d$ and $a + b = c + d$ (matrix (29)). Fig. 5($b$) shows another case of $a < c$ and

$b > d$, snowdrift game, in which $a + b > c + d$, see (34):

$$\begin{array}{c|cc} & A & B \\ \hline A & 4 & 3 \\ B & 5 & 0 \end{array} \tag{34}$$

Although the two strategies can coexist in this case, the effect of division of labor is much worse than that of $a + b = c + d$. Fig. 5($c$) and Fig. 5($d$) show the change of $J$ in cases of prisoner's dilemma (matrix (32)) and coordination game (matrix (33)), respectively. Owing to the randomness in the system evolution, even if the system reaches the evolutionarily stable state, the strategy of one agent may switch to another. In particular, for the coordination game, the system may switch repeatedly between two states: full $A$ and full $B$.

## VI. CONCLUSION

In this paper, a coordinated control approach for multi-agent systems based on evolutionary game theory is proposed. In this mathematical framework, agents rationally evaluate their payoffs and autonomously update their strategies through local interaction for the purpose of increasing their fitness in the game. The system evolves while the agents compete with each other in the game. During the evolutionary process, we do not need to specify certain agents' dynamics or assign them what to do. The overall target of the system could be reached only through the adaptive evolution of the population, but not by supervision or enforcement. Taking the problem of division of labor as an example, we explain how to implement the proposed control method in detail. Theoretical analysis and simulation verify the effect of proposed evolutionary game method on the coordinated control of multi-agent systems.

## REFERENCES


[1] J. M. Smith and G. R. Price, "The logic of animal conflict," *Nature*, vol. 246, no. 5427, pp. 15–18, Nov. 1973. doi: 10.1038/246015a0.

[2] I. L. Bajec and F. H. Heppner, "Organized flight in birds," *Animal Behav.*, vol. 78, no. 4, pp. 777–789, Oct. 2009. doi: 10.1016/j.anbehav.2009.07.007.

[3] G. J. Velicer and M. Vos, "Sociobiology of the myxobacteria," *Annu. Rev. Microbiol.*, vol. 63, no. 1, pp. 599–623, Oct. 2009. doi: 10.1146/annurev.micro.091208.073158.

[4] D. G. Rand and M. A. Nowak, "Human cooperation," *Trends Cogn. Sci.*, vol. 17, no. 8, pp. 413–425, Aug. 2013. doi: 10.1016/j.tics.2013.06.003.

[5] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proc. IEEE*, vol. 95, no. 1, pp. 215–233, Jan. 2007. doi: 10.1109/JPROC.2006.887293.

[6] Y. Cao, W. Yu, W. Ren, and G. Chen, "An overview of recent progress in the study of distributed multi-agent coordination," *IEEE Trans. Ind. Informat.*, vol. 9, no. 1, pp. 427–438, Feb. 2013. doi: 10.1109/TII.2012.2219061.

[7] W. Ren, R. W. Beard, and E. M. Atkins, "Information consensus in multivehicle cooperative control," *IEEE Control Syst. Mag.*, vol. 27, no. 2, pp. 71–82, Apr. 2007. doi: 10.1109/MCS.2007.338264.

[8] L. Wang and F. Xiao, "Finite-time consensus problems for networks of dynamic agents," *IEEE Trans. Autom. Control*, vol. 55, no. 4, pp. 950–955, Apr. 2010. doi: 10.1109/TAC.2010.2041610.

[9] Y. Zheng, J. Ma, and L. Wang, "Consensus of hybrid multi-agent systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 4, pp. 1359–1365, Apr. 2018. doi: 10.1109/TNNLS.2017.2651402.

[10] Y. Zhu, S. Li, J. Ma, and Y. Zheng, "Bipartite consensus in networks of agents with antagonistic interactions and quantization," *IEEE Trans. Circuits Syst., II, Exp. Briefs*, vol. 65, no. 12, pp. 2012–2016, Dec. 2018. doi: 10.1109/TCSII.2018.2811803.

[11] J. Ma, M. Ye, Y. Zheng, and Y. Zhu, "Consensus analysis of hybrid multiagent systems: A game-theoretic approach," *Int. J. Robust Nonlinear Control*, vol. 29, no. 6, pp. 1840–1853, Apr. 2019. doi: 10.1002/rnc.4462.

[12] J. Liu, Y. Zhang, C. Sun, and Y. Yu, "Fixed-time consensus of multi-agent systems with input delay and uncertain disturbances via event-triggered control," *Inf. Sci.*, vol. 480, pp. 261–272, Apr. 2019. doi: 10.1016/j.ins.2018.12.037.

[13] J. Liu, Y. Zhang, Y. Yu, and C. Sun, "Fixed-time event-triggered consensus for nonlinear multiagent systems without continuous communications," *IEEE Trans. Syst., Man, Cybern. Syst.*, to be published. doi: 10.1109/TSMC.2018.2876334.

[14] H. Choset, "Coverage for robotics—A survey of recent results," *Ann. Math. Artif. Intel.*, vol. 31, no. 1, pp. 113–126, Oct. 2001. doi: 10.1023/A:1016639210559.

[15] F. Xiao, J. Chen, and L. Wang, "Finite-time formation control for multi-agent systems," *Automatica*, vol. 45, no. 11, pp. 2605–2611, 2009. doi: 10.1016/j.automatica.2009.07.012.

[16] G. Jing, G. Zhang, H. W. J. Lee, and L. Wang, "Weak rigidity theory and its application to formation stabilization," *SIAM J. Control Optim.*, vol. 56, no. 3, pp. 2248–2273, Jan. 2018. doi: 10.1137/17M1122049.

[17] J. Yu, X. Dong, Q. Li, and Z. Ren, "Practical time-varying formation tracking for second-order nonlinear multiagent systems with multiple leaders using adaptive neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 12, pp. 6015–6025, Dec. 2018. doi: 10.1109/TNNLS.2018.2817880.

[18] G. Jing, Y. Zheng, and L. Wang, "Flocking of multi-agent systems with multiple groups," *Int. J. Control*, vol. 87, no. 12, pp. 2573–2582, Dec. 2014. doi: 10.1080/00207179.2014.935485.

[19] J. Lunze, "Synchronization of heterogeneous agents," *IEEE Trans. Autom. Control*, vol. 57, no. 11, pp. 2885–2890, Nov. 2012. doi: 10.1109/TAC.2012.2191332.

[20] L. Wang, F. Jiang, G. Xie, and Z. Ji, "Controllability of multi-agent systems based on agreement protocols," *Sci. China F, Inf. Sci.*, vol. 52, no. 11, pp. 2074–2088, Nov. 2009. doi: 10.1007/s11432-009-0185-7.

[21] W. Dong and J. A. Farrell, "Cooperative control of multiple nonholonomic mobile agents," *IEEE Trans. Autom. Control*, vol. 53, no. 6, pp. 1434–1448, Jul. 2008. doi: 10.1109/TAC.2008.925852.

[22] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*. Princeton, NJ, USA: Princeton Univ. Press, 1944.

[23] J. F. Nash, Jr., "Equilibrium points in n-person games," *Proc. Nat. Acad. Sci. USA*, vol. 36, no. 1, pp. 48–49, Jan. 1950. doi: 10.1073/pnas.36.1.48.

[24] J. M. Smith, *Evolution and the Theory of Games*. Cambridge, MA, USA: Cambridge Univ. Press, 1982.

[25] J. W. Weibull, *Evolutionary Game Theory*. Cambridge, MA, USA: MIT Press, 1995.

[26] H. Gintis, *Game Theory Evolving: A Problem-Centered Introduction to Modeling Strategic Interaction*. Princeton, NJ, USA: Princeton Univ. Press, 2000.

[27] J. Hofbauer and K. Sigmund, *Evolution Games Population Dynamics*. Cambridge, U.K.: Cambridge Univ. Press, 1998.

[28] M. A. Nowak, *Evolutionary Dynamics: Exploring the Equations of Life*. Cambridge, MA, USA: Harvard Univ. Press, 2006.

[29] J. Du and B. Wang, "Evolution of global cooperation in multi-level threshold public goods games with income redistribution," *Frontiers Phys.*, vol. 6, Jul. 2018, Art. no. 67. doi: 10.3389/fphy.2018.00067.

[30] M. A. Nowak and K. Sigmund, "Evolutionary dynamics of biological games," *Science*, vol. 303, no. 5659, pp. 793–799, Feb. 2004. doi: 10.1126/science.1093411.

[31] T. Reichenbach, M. Mobilia, and E. Frey, "Mobility promotes and jeopardizes biodiversity in rock–paper–scissors games," *Nature*, vol. 448, no. 7157, pp. 1046–1049, Aug. 2007. doi: 10.1038/nature06095.

[32] M. A. Nowak, N. L. Komarova, and P. Niyogi, "Computational and evolutionary aspects of language," *Nature*, vol. 417, no. 6889, pp. 611–617, Jun. 2002. doi: 10.1038/nature00771.

[33] P. D. Taylor and L. B. Jonker, "Evolutionary stable strategies and game dynamics," *Math. Biosci.*, vol. 40, nos. 1–2, pp. 145–156, Jul. 1978. doi: 10.1016/0025-5564(78)90077-9.

[34] J. Du and L. Tang, "Evolution of global contribution in multi-level threshold public goods games with insurance compensation," *J. Stat. Mech.*, vol. 2018, no. 1, Jan. 2018, Art. no. 013403. doi: 10.1088/1742-5468/aa9bb6.

[35] D. Fudenberg and L. A. Imhof, "Imitation processes with small mutations," *J. Econ. Theory*, vol. 131, no. 1, pp. 251–262, Nov. 2006. doi: 10.1016/j.jet.2005.04.006.

[36] A. Traulsen, J. M. Pacheco, and M. A. Nowak, "Pairwise comparison and selection temperature in evolutionary game dynamics," *J. Theor. Biol.*, vol. 246, no. 3, pp. 522–529, Jun. 2007. doi: 10.1016/j.jtbi.2007.01.002.

[37] J. Du, B. Wu, P. M. Altrock, and L. Wang, "Aspiration dynamics of multi-player games in finite populations," *J. Roy. Soc. Interface*, vol. 11, no. 94, May 2014, Art. no. 20140077. doi: 10.1098/rsif.2014.0077.

[38] J. Du, B. Wu, and L. Wang, "Aspiration dynamics in structured population acts as if in a well-mixed one," *Sci. Rep.*, vol. 5, Jan. 2015, Art. no. 8014. doi: 10.1038/srep08014.

[39] X. Chen and L. Wang, "Promotion of cooperation induced by appropriate payoff aspirations in a small-world networked game," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 77, no. 1, Jan. 2008, Art. no. 017103. doi: 10.1103/PhysRevE.77.017103.

[40] J. Du, "Redistribution promotes cooperation in spatial public goods games under aspiration dynamics," *Appl. Math. Comput.*, vol. 363, Dec. 2019, Art. no. 124629. doi: 10.1016/j.amc.2019.124629.

[41] M. A. Nowak and K. Sigmund, "A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game," *Nature*, vol. 364, no. 6432, pp. 56–58, Jul. 1993. doi: 10.1038/364056a0.

[42] M. Posch, "Win–stay, lose–shift strategies for repeated games—Memory length, aspiration levels and noise," *J. Theor. Biol.*, vol. 198, no. 2, pp. 183–195, May 1999. doi: 10.1006/jtbi.1999.0909.

[43] F. Fu, L. Wang, M. A. Nowak, and C. Hauert, "Evolutionary dynamics on graphs: Efficient method for weak selection," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 79, no. 4, Apr. 2009, Art. no. 046707. doi: 10.1103/PhysRevE.79.046707.

[44] B. Wu, P. M. Altrock, L. Wang, and A. Traulsen, "Universality of weak selection," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 82, no. 4, Oct. 2010, Art. no. 046106. doi: 10.1103/PhysRevE.82.046106.

[45] B. Skyrms, *The Stag Hunt and the Evolution of Social Structure*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[46] J. M. Pacheco, F. C. Santos, M. O. Souza, and B. Skyrms, "Evolutionary dynamics of collective action in N-person stag hunt dilemmas," *Proc. Roy. Soc. B, Biol. Sci.*, vol. 276, no. 1655, pp. 315–321, Jan. 2009. doi: 10.1098/rspb.2008.1126.

[47] R. M. Murray, "Recent research in cooperative control of multivehicle systems," *J. Dyn. Syst., Meas., Control*, vol. 129, no. 5, pp. 571–583, 2007. doi: 10.1115/1.2766721.

[48] F. Bullo, J. Cortés, and S. Martínez, *Distributed Control of Robotic Networks: A Mathematical Approach to Motion Coordination Algorithms*. Princeton, NJ, USA: Princeton Univ. Press, 2009.

[49] Z. Qu, *Cooperative Control of Dynamical Systems: Applications to Autonomous Vehicles*. New York, NY, USA: Springer, 2009.

[50] Y. Hu, W. Zhao, and L. Wang, "Vision-based target tracking and collision avoidance for two autonomous robotic fish," *IEEE Trans. Ind. Electron.*, vol. 56, no. 5, pp. 1401–1410, May 2009. doi: 10.1109/TIE.2009.2014675.

[51] K.-S. Hwang, S.-W. Tan, and C.-C. Chen, "Cooperative strategy based on adaptive Q-learning for robot soccer systems," *IEEE Trans. Fuzzy Syst.*, vol. 12, no. 4, pp. 569–576, Aug. 2004. doi: 10.1109/TFUZZ.2004.832523.

[52] D.-F. Zheng, H.-P. Yin, C.-H. Chan, and P. M. Hui, "Cooperative behavior in a model of evolutionary snowdrift games with N-person interactions," *Europhys. Lett.*, vol. 80, no. 1, Sep. 2007, Art. no. 18002. doi: 10.1209/0295-5075/80/18002.

[53] M. D. Santos, F. L. Pinheiro, F. C. Santos, and J. M. Pacheco, "Dynamics of n-person snowdrift games in structured populations," *J. Theor. Biol.*, vol. 315, pp. 81–86, Dec. 2012. doi: 10.1016/j.jtbi.2012.09.001.

[54] A. Traulsen, D. Semmann, R. D. Sommerfeld, H.-J. Krambeck, and M. Milinski, "Human strategy updating in evolutionary games," *Proc. Nat. Acad. Sci. USA*, vol. 107, no. 7, pp. 2962–2966, Feb. 2010. doi: 10.1073/pnas.0912515107.

[55] B. Wu, J. García, C. Hauert, and A. Traulsen, "Extrapolating weak selection in evolutionary games," *PLoS Comput. Biol.*, vol. 9, no. 12, Dec. 2013, Art. no. e1003381. doi: 10.1371/journal.pcbi.1003381.

[56] R. Axelrod and W. D. Hamilton, "The evolution of cooperation," *Science*, vol. 211, no. 4489, pp. 1390–1396, 1981. doi: 10.1126/science.7466396.

[57] G. Ellison, "Learning, local interaction, and coordination," *Econometrica*, vol. 61, no. 5, pp. 1047–1071, Sep. 1993. doi: 10.2307/2951493.

[58] H. Matsuda, N. Ogita, A. Sasaki, and K. Satō, "Statistical mechanics of population: The lattice Lotka-Volterra model," *Prog. Theor. Phys.*, vol. 88, no. 6, pp. 1035–1049, Dec. 1992. doi: 10.1143/ptp/88.6.1035.

[59] H. Ohtsuki, C. Hauert, E. Lieberman, and M. A. Nowak, "A simple rule for the evolution of cooperation on graphs and social networks," *Nature*, vol. 441, no. 7092, pp. 502–505, May 2006. doi: 10.1038/nature04605.

[60] T. Antal, H. Ohtsuki, J. Wakeley, P. D. Taylor, and M. A. Nowak, "Evolution of cooperation by phenotypic similarity," *Proc. Nat. Acad. Sci. USA*, vol. 106, no. 21, pp. 8597–8600, May 2009. doi: 10.1073/pnas.0902528106.

[61] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, Oct. 1999. doi: 10.1126/science.286.5439.509.

[62] A.-L. Barabási, "Scale-free networks: A decade and beyond," *Science*, vol. 325, no. 5939, pp. 412–413, Jul. 2009. doi: 10.1126/science.1173299.



**JINMING DU** (M'18) was born in Shenyang, Liaoning, China, in 1987. He received the B.S. degree in automation from Northeastern University, Shenyang, in 2010, and the Ph.D. degree in general mechanics and foundation of mechanics from Peking University, Beijing, China, in 2016.

He is currently an Assistant Professor with the Institute of Industrial and Systems Engineering, College of Information Science and Engineering, Northeastern University. His research interests include evolutionary game dynamics, networked control systems, swarm intelligence, and complex systems modeling and control.