Dynamic Evolution of Complex Networks: A Reinforcement Learning Approach Applying Evolutionary Games to Community Structure

Bin Pi, *Student Member, IEEE,* Liang-Jian Deng, *Senior Member, IEEE,* Minyu Feng, *Senior Member, IEEE,* Matjaž Perc, *Member, IEEE,* and Jürgen Kurths

Abstract—Complex networks serve as abstract models for understanding real-world complex systems and provide frameworks for studying structured dynamical systems. This article addresses limitations in current studies on the exploration of individual birth-death and the development of community structures within dynamic systems. To bridge this gap, we propose a networked evolution model that includes the birth and death of individuals, incorporating reinforcement learning through games among individuals. Each individual has a lifespan following an arbitrary distribution, engages in games with network neighbors, selects actions using Q-learning in reinforcement learning, and moves within a two-dimensional space. The developed theories are validated through extensive experiments. Besides, we observe the evolution of cooperative behaviors and community structures in systems both with and without the birth-death process. The fitting of real-world populations and networks demonstrates the practicality of our model. Furthermore, comprehensive analyses of the model reveal that exploitation rates and payoff parameters determine the emergence of communities, learning rates affect the speed of community formation, discount factors influence stability, and two-dimensional space dimensions dictate community size. Our model offers a novel perspective on real-world community development and provides a valuable framework for studying population dynamics behaviors.

Index Terms—Complex networks, Evolutionary games, Reinforcement learning, Community structure, Stochastic process.

arXiv:2506.17925v1 [cs.SI] 22 Jun 2025

1 Introduction

From the internet to mobile communication networks [1], [2], social relationship networks [3], [4], biological networks [5], [6], and transportation networks [7], [8], etc. [9], [10], [11], complex networks have become ubiquitous and have an increasing impact on human production and life. How to establish a dynamic evolution mechanism for complex networks that accurately reflects real-world conditions is a key challenge in the field of complex networks. The origin of complex networks dates back to 1,960 when Erdős and Rényi [12] established the theory of random graphs, which pioneered the study of complex networks. The "six degrees of separation" small-world experiment conducted by Milgram [13] in 1,967 is a classic example of an empirical study of complex networks. At the end of the 20th century, Watts and Strogatz [14] proposed small-world networks, Barabási and Albert [15] developed scale-free networks. They respectively revealed the small-world and scale-free properties of complex networks, built corresponding models to elaborate the mechanisms of these properties, and successfully explained the widespread small-world properties [16], [17] and power-law distributions [18], [19] in reality. Since then, network science has started to emerge, and various novel complex network models have sprung up in people's vision.

An extensively developed complex network model not only provides insights into the evolutionary patterns of real network structures but also facilitates the investigation of various collective behaviors within structured populations, including synchronization [20], [21], evolutionary games [22], [23], propagation [24], [25], and so on [26], [27], [28]. This capability makes it a valuable tool for studying and understanding complex phenomena in diverse fields. In 1,992, Nowak and May [29] first introduced spatial chaos into the prisoner's dilemma, and their findings revealed that the cooperation frequency in the network remained stable at around 0.374, regardless of the specific values of the game parameters. Subsequent studies [30], [31] have further demonstrated that spatial structure has a positive impact on the evolution of cooperation in evolutionary games. However, the study published by Hauert and Doebeli [32]

B. Pi, L.-J. Deng, and M. Feng were supported in part by the National Nature Science Foundation of China (NSFC) under Grant Nos. 12271083, 62206230, and 62273077, in part by the Sichuan Province's Science and Technology Empowerment for Disaster Prevention, Mitigation, and Relief Project under Grant No. 2025YFNH0001, and in part by the Natural Science Foundation of Chongqing under Grant No. CSTB2023NSCQ-MSX0064. M.P. was supported by the Slovenian Research and Innovation Agency (Javna agencija za znanstvenoraziskovalno in inovacijsko dejavnost Republike Slovenije) under Grant Nos. P1-0403 and N1-0232.

Bin Pi and Liang-Jian Deng are with the School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: liangjian.deng@uestc.edu.cn).

Minyu Feng is with the College of Artificial Intelligence, Southwest University, Chongqing 400715, China (e-mail: myfeng@swu.edu.cn).

Matjaž Perc is with Faculty of Natural Sciences and Mathematics, University of Maribor, Koroška cesta 160, 2000 Maribor, Slovenia, with Community Healthcare Center Dr. Adolf Drolc Maribor, Ulica talcev 9, 2000 Maribor, Slovenia, with Department of Physics, Kyung Hee University, 26 Kyungheedae-ro, Dongdaemun-gu, Seoul 02447, Republic of Korea, with Complexity Science Hub, Metternichgasse 8, 1030 Vienna, Austria, and with University College, Korea University, 145 Anam-ro, Seongbuk-gu, Seoul 02841, Republic of Korea.

Jürgen Kurths is with the Department of Complexity Science, Potsdam Institute for Climate Impact Research, 14473 Potsdam, Germany, and also with the Institute of Physics, Humboldt University of Berlin, 12489 Berlin, Germany.

Corresponding authors: Liang-Jian Deng and Minyu Feng.

challenged this notion by suggesting that spatial structure actually inhibits the evolution of cooperation in the snowdrift game. Later, Santos and Pacheco [33] conducted studies on scale-free networks and demonstrated that cooperators play a dominant role in both the prisoner's dilemma game and snowdrift game, providing a plausible explanation for the existence of cooperation in real-world scenarios. With the development of complex network theory and evolutionary game theory, various mechanisms have been proposed to promote cooperative behavior. One of the most renowned contributions is the five rules introduced by Nowak [34]. Furthermore, more recent researches have uncovered that reputation [35], [36], reward and punishment mechanisms [37], [38] and game transitions [39], [40] also play significant roles in facilitating the persistence of cooperation. These findings have shed new light on the diverse strategies and mechanisms that contribute to the survival and evolution of cooperative behavior in various contexts.

Complex network modeling based on real-world phenomena is a classical idea and has been strongly studied. After the efforts in recent years, substantial progress has been made in the study of complex networks, numerous statistical features of complex networks have been discovered, and a growing number of complex network models that are more realistic have been proposed one after another [41], [42], [43]. For example, in recent years, some researchers found that the number of individuals in the network should not only increase but also decrease. Feng et al. [44] noticed this point and proposed a novel evolving network model that takes the growing and decreasing process into account based on the queueing system. Ref. [45] discussed the relationship between higher-order interactions and collective behavior based on the fact that interactions can often occur in groups of three or more individuals, which cannot be expressed simply in terms of pairwise interactions.

Despite the significant attention and fruitful results achieved in complex network modeling, existing research falls short in deriving the formation and evolutionary mechanisms of community structures from the game interactions among individuals. This phenomenon is pervasive in practice, ranging from interactions among small microorganisms [46] to large-scale strategic games among nations [47]. Moreover, community structures are a key feature of complex networks [48]. Numerous empirical studies have shown that many networks exhibit heterogeneity, i.e., they consist of various types of nodes rather than random connections among a vast number of nodes of the same nature, where there are more connections of the same type of nodes and relatively few connections of different types of nodes [49], [50]. We refer here to the subgraphs formed by nodes of the same type and the edges between these nodes as communities in the network. Communities observed in real-world networks represent distinct collections of different specific objects. For example, communities in citation networks consist of related papers on the same topic, communities in the World Wide Web are groups of websites discussing similar subjects, communities in social networks represent real social groups based on interests or backgrounds, and communities in biochemical networks or electronic circuit networks can denote specific functional units. The identification and analysis of communities in these networks provide valuable insights for understanding their structures and dynamics and optimizing their functionality.

To fill this gap and address the shortcomings of previous studies, this paper primarily applies networked evolutionary game theory and Q-learning in reinforcement learning to uncover the formation and development mechanisms of community structures in real-world systems. To the best of our knowledge, this is the first study to model complex networks based on individual game interactions using reinforcement learning. Concretely, we assume that the individuals within a system are intelligent and have the potential to alter their locations in the two-dimensional space at any given moment to maximize their payoffs. They learn from past experiences to determine which mobile actions will yield the highest rewards, but occasionally make random choices due to limited intelligence or knowledge. Simultaneously, we develop a mechanism for transforming the distribution of individuals in a two-dimensional space to the network structure in a high-dimensional space. We note that individual rewards are determined through the games they are engaged in with their network neighbors. To model this process, we apply Q-learning in reinforcement learning, a model-free reinforcement learning approach, which identifies the optimal action plan based on the individual's current state. The typical structure of the network generated according to our model is shown in Fig. 1, which exhibits clear community structures, where different colors indicate different communities. Moreover, we conduct a comparative analysis of the evolution of community and cooperative behaviors both in systems with and without the birth-death process, and we apply them to real populations and networks, yielding promising results. Additionally, we perform a comprehensive analysis of the model's influence on the community and assess our model in comparison to the heuristic model and classical networks.

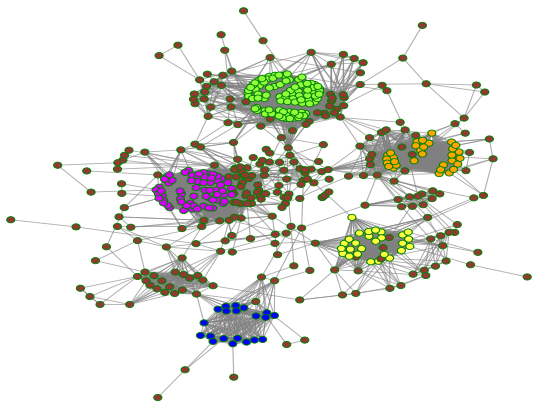


Fig. 1. **An example of the network structure generated according to our model.** The figure shows a network structure generated based on our model, where different colors denote different community structures, suggesting that the model proposed in this paper can provide an explanation for the formation and development of community structures in reality.

The model proposed in this paper offers a novel perspective for understanding the emergence and evolution of communities, bridges the gap left by previous studies in this aspect, advances the study of complex network modeling, and introduces a novel framework for examining the

dynamics on structured populations, such as evolutionary games, propagation, and synchronization.

In summary, the major contributions of this paper can be outlined as follows:

- We propose an innovative model for the dynamic evolution of complex networks based on a game between individuals with reinforcement learning. Besides, we establish a transition mechanism that maps the distribution of individuals in the two-dimensional space to the network structure in a high-dimensional space.
- We introduce a birth-death process of individuals to the system and determine the scale and distribution of the system at the steady state through both theoretical and experimental analyses. Furthermore, we compare it with systems without a birth-death process in numerical simulations.
- The practicality of the model is validated by fitting the population of different countries and the degree distribution of real networks. The comparison between the model proposed in this paper and classical networks in terms of network structure, degree distribution, clustering coefficient, etc. is also conducted.
- The model proposed in this paper provides a novel perspective on the emergence and development of community structure in reality, and the impacts of the parameters on community structure have been investigated from various aspects.

The remainder of this paper is structured as follows: In Section 2, we display the details of the complex network model based on the game of individuals with reinforcement learning. In Section 3, we conduct some simulations to validate our theory, observe the evolution of community structure and population cooperative behavior in systems both with and without the birth-death process, and fit the proposed model to population data and real networks. Subsequently, Section 4 offers a comprehensive analysis of the formation and evolution of community structures from various perspectives. Then, we provide a comparative analysis of the proposed models in Section 5. Finally, in Section 6, we summarize our study and provide some points for improvement.

2 Evolving Networks Based on a Game between Individuals with Reinforcement Learning

In this section, we introduce a novel evolving network that considers the birth-death of individuals using the queuing theory. We assume that each individual with self-learning will move in the two-dimensional space based on their past experiences and interactions with neighbors in a game and we transform these individual dynamics into the corresponding network structure in a high-dimensional space through the proposed mechanism.

2.1 Birth-Death Process of Individual

First, we consider a system with $r \times r$ locations, where each location can be represented by a lattice in a two-dimensional space (containing $r \times r$ lattices). Each individual in the system has a lifetime that follows a general distribution. Besides, new individuals are input to the system at a specific rate λ that follows an exponential distribution [44] with the probability distribution:

$$f_1(x;\lambda) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}. \tag{1}$$

The exponential distribution is employed to model individual births due to its memoryless property and mathematical simplicity, which make it an effective representation of independent random events in various natural processes. Additionally, the time intervals between events in many biological and physical phenomena align well with the assumptions of the exponential distribution [51], [52]. For instance, in large natural populations, the birth times of individuals typically occur as uncorrelated random events, with intervals that often exhibit the characteristics of an exponential distribution.

Then, the size of the system $N(t)$, i.e., the number of individuals in the system at time t, can be regarded as a continuous-time Markov chain with the state space $E = \{0, 1, 2, \cdots\}$. In addition, we assume that each individual in the system is a customer, and its behavior before leaving the system at the end of its lifespan is a service. Thus, this dynamic process of the input and output of individuals can be described by a $M/G/\infty$ queueing system, where M represents a Poisson process for the input of individuals, G indicates the lifespan of an individual obeying a general distribution, and ∞ denotes the number of individuals in the system belonging to $[0, \infty)$. We illustrate the impact of the birth and death of individuals on the system size in Fig. 2. The first and third rows depict the birth and death rates of the system, respectively, with birth intervals following an exponential distribution and lifespans adhering to a general distribution. The second row indicates that the size of the system falls within the range of $[0, \infty)$.

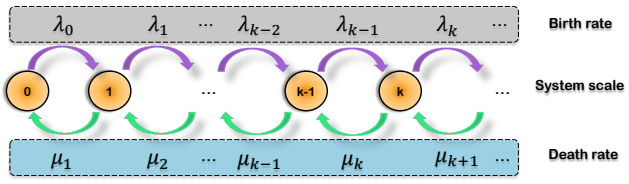


Fig. 2. **Phase transition of system scale.** The system scale undergoes a transition from $k-1$ to k at an exponential rate of λ_{k-1} or from k to $k-1$ at a general rate of μ_k, and the system size is constrained within the range of $[0, \infty)$.

Next, we derive the scale at which the system evolves to reach a plateau. Before presenting the results, we first give some illustrations. We suppose that S is the service time of an individual and that an individual staying in the system at time t indicates that the arrival time x of the individual obeys $t - S < x < t$. The individual arrives before time t is a uniform distribution with range $[0, t]$ since the arrival of individuals is a Poisson process. Therefore, we can deduce the probability that an individual arrives before time t and still remains in the system is:

$$X(t) = \int_0^t \frac{P[S > t - x]}{t} dx = \int_0^t \frac{1 - G(t-x)}{t} dx, \tag{2}$$

where G represents the distribution function of the service time.

We consider that the probability of the number of individuals at time t and those that arrive before t follows binomial distribution and the arrival time of each individual is i.i.d. Then, for those $m \leq n$, we get

$$\begin{aligned} P_{n,m}(N \mid A, t) &= C_n^m \left[X(t)\right]^m \left[1 - X(t)\right]^{n-m} \\ &= \frac{n! \left[X(t)\right]^m \left[1 - X(t)\right]^{n-m}}{m!(n-m)!}. \end{aligned} \tag{3}$$

Otherwise, i.e., for those $m > n$, it is obviously that

$$P_{n,m}(N \mid A, t) = 0. \tag{4}$$

Therefore, the number of individuals in the system at time t ($N(t)$) and those that arrive before t ($A(t)$) follows the conditional probability:

$$\begin{aligned} P_{n,m}(N \mid A, t) &= P\{N(t) = m \mid A(t) = n\} \\ &= \begin{cases} \frac{n!X(t)^m[1-X(t)]^{n-m}}{m!(n-m)!}, & m \leq n \\ 0, \ otherwise \end{cases}. \end{aligned} \tag{5}$$

Subsequently, we investigate the limiting probability distribution of the system size.

Theorem 1. *For the continuous-time Markov chain $N(t)$ with the state space E, assume that the expectation of death process $\{G(t), t \geq 0\}$ exists, its limiting probability $\{\pi_i, i = 1, 2, \cdots\}$ exists and follows*

$$\begin{aligned} \pi_i &= \lim_{t\to\infty} p_i(t) = \lim_{t\to\infty} P\{N(t) = i\} \\ &= \frac{\{\lambda E[G(t)]\}^i}{i!} e^{-\lambda E[G(t)]}. \end{aligned} \tag{6}$$

The proof of Thm. 1 is given in Section 1 of the supplementary material for those who are interested. According to Thm. 1, we can derive the probability distribution of the system size when it evolves to a steady state. Furthermore, we deduce certain statistical properties of the system size based on the limiting probability demonstrated in Thm. 1.

Theorem 2. *The average scale of the system is*

$$E[N(t)] = \lambda E[G(t)], \tag{7}$$

the variance of the scale is

$$D[N(t)] = \lambda E[G(t)], \tag{8}$$

and the average staying time of each individual is

$$E(T) = E[G(t)]. \tag{9}$$

The proof of Thm. 2 can be found in Section 2 of the supplementary material for those who are interested. Therefore, we can obtain the expected scale of the system, the dispersion of a set of scales, and the lifespan of an individual when the parameter λ and the death process $\{G(t), t \geq 0\}$ are known.

2.2 Individual Movement Based on Reinforcement Learning

We consider the case that each individual in the system has the potential to change its location at each moment to pursue a greater profit. Besides, every individual is intelligent, i.e., can get information from past experiences to determine which mobile actions will yield the highest benefits. In other words, the individual's choice of mobile behavior is based on learning experience, rather than obeying some simple rules. However, the intelligence of each individual is limited, i.e., at each time step, the individual can choose to move to a location that can bring the maximum payoff with the probability of δ (this behavior is called exploitation) and also has a random probability ϵ (this behavior is called exploration) to make a random choice of location to move to, where $\epsilon + \delta = 1$. Therefore, we employ Q-learning with an ϵ-greedy strategy to model the aforementioned process in this paper. This is a model-free (does not require a model of the environment) reinforcement learning that will find the best action according to the individual's current state. Concretely, an individual performing an action in a particular state provides it with a reward, and the goal of the individual is to maximize the cumulative reward at all steps. Each individual has a utility table called Q-table learned from its action experience through the Bellman equation, which leads the individual to take the best action in each state.

Hereby, we provide a detailed description of the process of individual movement. In the system, the state of each individual is the location it lives in, which can be represented by dividing the two-dimensional space into equal-sized grids, where a grid is regarded as a location. Each individual will choose to move during a one-time step based on its current state and past experience. Each individual has six optional movement actions: move left, move right, move up, move down, stay, and random move. The first four actions mean that the individual will move to a location adjacent to the location it is currently in, "stay" indicates that the individual will not change its location, and "random move" suggests that the individual will move to a location randomly selected from all locations. When an individual performs a specific action, it will receive a certain amount of reward, which is significant for the individual's future action choice.

As we described before, each individual has a dynamic utility table to guide it in choosing the best action for each state, and the updating rule of the utility table at each time step can be expressed as

$$Q_{S_t}^{A_t}(t+1) = Q_{S_t}^{A_t}(t) + \eta[R_{S_t}^{A_t}(t+1) + \gamma \max_{a\in A} Q_{S_{t+1}}^{a}(t) - Q_{S_t}^{A_t}(t)], \tag{10}$$

where $Q_{S_t}^{A_t}(t)$ represents the utility gained by the individual in state S_t when executing action A_t at time t. $R_{S_t}^{A_t}(t+1)$ denotes the immediate payoff obtained by the individual in state S_t at time $t+1$ when performing action A_t at time t, and the calculation of the individual's immediate payoff will be explained in detail in the next subsection. In addition, the $\max_{a\in A} Q_{S_{t+1}}^{a}(t)$ means the maximum utility received by the individual in the future after moving to the state S_{t+1}. Moreover, η and γ are the learning rate and discount factor, respectively. A greater γ will cause the individual to focus more on past experiences, otherwise, the individual will focus more on immediate payoff. We note that the Q-table possessed by each individual is not static, and its internal values are evolving according to Eq. 10.

2.3 Payoff Calculation and Strategy Evolution of Individual

In this subsection, we introduce the calculation of the immediate payoff, which is based on the game between individuals and their neighbors. Before then, we illustrate the transition mechanism from the previously described movement of individuals between locations in the two-dimensional space to complex networks in high-dimensional space, which is primarily designed based on theories of memory decay [53] and strength of ties [54]. Specifically, there are two main situations:

i) One is individuals i and j are in distinct locations at time t. In this case, the rules for edge formation and edge weight updating are as follows:

(a) If two individuals have no edge at the last time $t-1$, then there will also be no edge between them at the current time. It reflects the principle in social networks that the formation of ties typically requires a history of interaction. In practice, connections between individuals are generally established based on previous interactions. Therefore, if no prior interactions or connections exist, it is unlikely that a new connection will form between them.

(b) If there is an edge with weight $w_{i,j}^{t-1}$ between individuals i and j at the previous time $t-1$, then the weight of the edge will be updated to $w_{i,j}^{t} = w_{i,j}^{t-1}/\beta$, where β is a decay factor. The decay mechanism aligns with the real-life principle of relationship maintenance, where relationships typically weaken over time, particularly when interactions between individuals diminish. If the updated weight exceeds the generation threshold of the edge σ, the edge will be retained; otherwise, it will be removed.

Therefore, the edge weight updating rule in this case can be expressed as

$$w_{i,j}^{t+1} = \begin{cases} 0, \text{ no edge between } i \text{ and } j \text{ at } t \\ w_{i,j}^{t}/\beta, \text{ otherwise} \end{cases}. \quad (11)$$

ii) The other scenario is individuals i and j are in the same state at time t. In this case, the rules for edge formation and the updating of edge weights are as follows:

(a) If there is no edge between the two individuals at the last time $t-1$, then a new edge will be established between them with an initial weight τ. This mechanism models the creation of new interactions and the formation of relationships. For example, in social networks, individuals might form new connections when they meet through common activities or shared interests.

(b) If individuals i and j have an edge with weight $w_{i,j}^{t-1}$ at the previous time $t-1$, then the edge will be preserved and its weight will be updated to $w_{i,j}^{t-1}/\beta + \tau$. Here, $w_{i,j}^{t-1}/\beta$ denotes the natural decay of the edge weight, simulating the natural weakening of the relationship over time. The addition of τ reflects the strengthening of the relationship due to new interactions or joint activities when i and j are in the same state. The updating mechanism aligns with the theory of the strength of ties, which posits that frequent interactions enhance the strength of connections between individuals.

Therefore, the edge weight updating rule in this case can be denoted as

$$w_{i,j}^{t+1} = \begin{cases} \tau, \text{ no edge between } i \text{ and } j \text{ at } t \\ w_{i,j}^{t}/\beta + \tau, \text{ otherwise} \end{cases}. \quad (12)$$

It is important to highlight that the decay factor β governs the gradual attenuation of edge weights over time. A larger value of β results in faster decay of the weights and a shorter memory effect of the network. This decay mechanism mirrors the natural weakening of relationships in real-world contexts or the diminishing impact of reduced interactions over time on the strength of relationships. Conversely, the weight τ represents the additional contribution from interactions when a relationship is either established or maintained between individuals. It reflects either the initial strength of a newly formed edge or the reinforcement of an existing relationship. In summary, the configuration of edge weights updating among individuals aligns with actual social behavior patterns, ensuring that the evolution of the network is consistent with real-world dynamics.

At each discrete time step, after all individuals have synchronously taken actions to reach new states, the networked structure is updated based on the newly adjusted weights among individuals. In the new network, each node represents an individual, and the edges between nodes indicate the interactions between pairs of individuals. Subsequently, each individual synchronously engages in the snowdrift game with its neighbors in the updated network to obtain payoffs. The interactions between different strategy pairs result in different payoffs, which can be represented by the following payoff matrix:

$$M = \begin{pmatrix} 1 & 1-r \\ 1+r & 0 \end{pmatrix}, \quad (13)$$

where r belongs to $[0,1]$, which is an adjustable payoff parameter for the snowdrift game. We utilize the strategy notation of two-dimensional unit vectors as $s_i = [1,0]^T$ for cooperate and $s_i = [0,1]^T$ for the defect. Accordingly, the payoff U_i obtained by individual i can be expressed as follows:

$$U_i = \sum_{j \in \Omega} s_i^T M s_j, \quad (14)$$

where Ω means the set of neighbors for individual i in the updated network structure. We emphasize that the cumulative payoff for all individuals in the same state corresponds to the immediate payoff obtained when an individual takes a specific action. Next, each individual performs the strategy evolution according to the Fermi function [55], [56]. Specifically, individuals synchronously and randomly select one of their neighbors for payoff comparison and adopt the neighbor's strategy in the next discrete time step with a probability given by:

$$P(s_i \leftarrow s_j) = \frac{1}{1+\exp\left[(U_i - U_j)/\kappa\right]}, \quad (15)$$

where κ denotes the noise intensity, which is utilized to characterize the uncertainty associated with the process of strategy evolution. A larger value of κ implies a higher tendency for individuals to adopt strategies in a more stochastic manner.

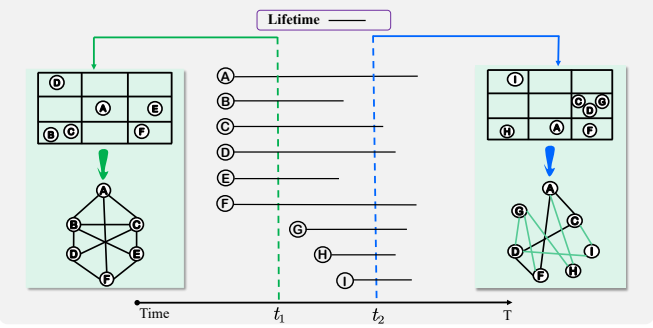


Fig. 3. **An illustration of the model.** The figure shows the evolution example of the complex network model. Black duration indicates the lifetime of the individual, which follows a general distribution. The individual will be removed from the system once its lifetime ends. We select t_1 and t_2 to observe the snapshots of the two-dimensional space and the network, which have a one-to-one correspondence between them. When $t = t_2$, individuals B and E in t_1 have died since their lifetimes are over, and new individuals G, H, and I are born in the system. The connections between newly born individuals and individuals already in the network are shown in green.

To enhance the clarity of our model, we give an example in Fig. 3. The black line indicates the lifetime of each individual. We choose evolution times t_1 and t_2 for observation. At $t = t_1$, there are six individuals in the system, and the corresponding networked structure can be obtained based on the locations and weights between individuals in the two-dimensional space. At $t = t_2$, individuals B and E in t_1 have passed away, because their lifespans ended. New individuals G, H, and I join the system and create a new network, since the addition of new individuals and individual movements, where interactions between new individuals and individuals already in the network are highlighted in green.

3 Simulations and Results

In this section, we conduct extensive simulation experiments to validate the previous theory. To begin with, we introduce the methods of simulation experiments. Subsequently, we study the evolution and statistical distribution of individual numbers under different death processes to validate the theoretical analysis. Next, we compare the evolution of cooperative behaviors, the emergence and development of community structures, and the effects of payoff parameters and exploitation rates on community structures in systems both with and without birth-death process (SWBD and SWOBD)[i]. It is worth noting that for the sake of clarity, for systems with the birth-death process, we only choose to show results where the death

i. The code of the SWBD and SWOBD is available at https://github.com/BinPi123/Dynamic-Evolution-of-Complex-Networks

process obeys a power-law distribution (SWBD with power-law distribution), and the interested reader can refer to Section 3 of the supplementary material to access all the results. Then, we apply our model to fit populations from four different countries and degree distributions of six distinct real networks.

3.1 Methods

Herein, we explain some methods for our subsequent simulations. At the initial moment, each individual is uniformly distributed among different locations, the utility tables are initialized to 0, and individuals select cooperation or defection as their initial strategy with the same probability. We divide the two-dimensional space into 10×10 grids of equal size, with each location corresponding to one grid. The initial network is a complete graph and then evolves dynamically as individuals move around. For the generation of systems with the birth-death process, we generate a Poisson process based on the fact that it has the property of an exponential time interval, which is generated by the function $expovariate$ of the package $random$. Once an individual is born, it will be connected to all individuals in the same state and is assigned a lifespan that obeys a specific probability distribution. In this study, we explore four distinct death processes that follow different distributions, all of which are widely used in practice. The first one is the power-law distribution, often observed in urban population dynamics, where a small number of large cities have significantly higher populations compared to the majority of smaller cities, which have relatively modest populations [57]. The second distribution is the uniform distribution, frequently employed in simple random sampling, where each individual has an equal chance of being selected for death without distinction [58]. The third is the exponential distribution, which is typically used to model the lifetimes of devices. The non-memory property of this distribution makes it appropriate for scenarios where the failure rate of a device remains constant over time [59]. Lastly, we consider the lognormal distribution, which is commonly observed in certain biological characteristics, such as the sizes of human organs, cell dimensions, and organism lifespans [60].

Next, we describe the generation for the four distributions. The power-law distribution is generated using the inverse transform technique. Concretely, the probability distribution of the power-law distribution [61] is defined as follows:

$$f_1(x;\alpha) = \frac{\alpha-1}{x_{min}^{1-\alpha}} x^{-\alpha},\ \alpha > 2, \tag{16}$$

where x_{min} and α indicate the lower bound and exponent of the distribution, respectively. Accordingly, its distribution function can be derived as follows:

$$F(x) = P\{X \le x\} = \int_{x_{min}}^{x} f_1(x;\alpha)dx = 1 - \frac{x^{1-\alpha}}{x_{\min}^{1-\alpha}}. \tag{17}$$

Denote $u \sim U(0,1)$ as a random variable following a uniform distribution in the range $(0,1)$ and $u = F(x)$, then we have

$$x = F^{-1}(u) = x_{\min}(1-u)^{\frac{1}{1-\alpha}}. \tag{18}$$

According to the inverse transform technique, the variable x in Eq. 18 follows a power-law distribution with a lower bound x_{min} and exponent α. It can be further simplified to $x = x_{\min}u^{\frac{1}{1-\alpha}}$, given that $1-u$ also follows a uniform distribution within the range $(0,1)$. Consequently, a power-law distribution can be generated using $x_{\min}u^{\frac{1}{1-\alpha}}$, where $u \sim U(0,1)$. For the generation of uniform and exponential distributions, we utilize the $uniform$ and $expovariate$ functions from the $random$ package, respectively. Besides, the lognormal distribution is generated using the $random.lognoraml$ function from the $numpy$ package.

Once the lifespan of an individual is over, it is removed from the network, along with its connections to its neighbors. In the subsequent simulations, we conduct experiments on both systems with and without the birth-death process to demonstrate the effect of this process on the results. All final results are carried out on $Python$ and are averaged over ten independent simulations to ensure high accuracy.

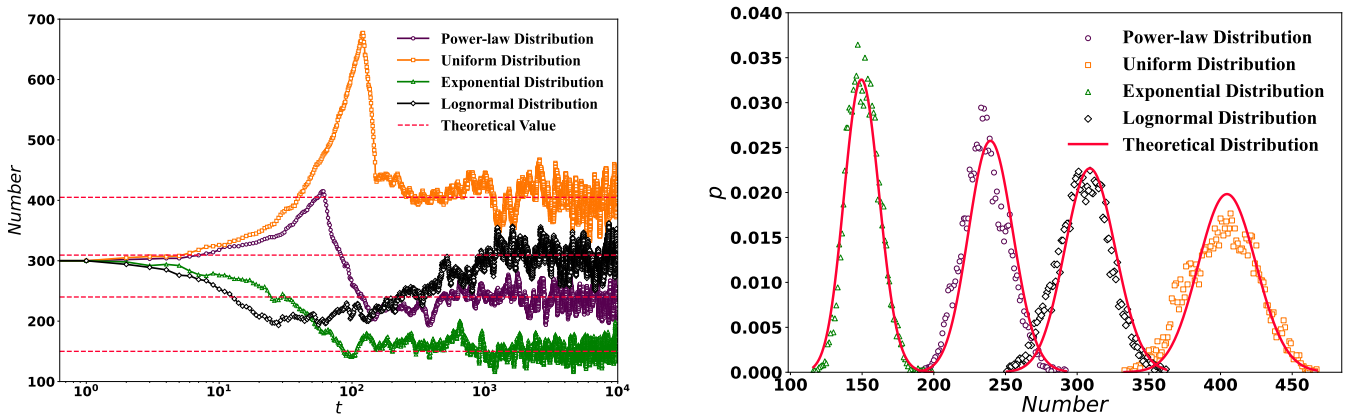
In brief, Algor. 1 illustrates the Monte Carlo process for the complex network modeling based on the game between individuals with Q-learning in reinforcement learning. It is important to note that, at each discrete time step, all individuals first move synchronously and then play the snowdrift game with their neighbors in the updated network structure. Following this, the Q-table and strategy of each individual are updated synchronously.

Algorithm 1 Monte Carlo Simulation

Initialize the current time $t_{now} = 0$, the update time $t_{update} = 0$, the weights among individuals $weight$, the network structure $\mathcal{G}$, each individual's state S, action a, utility table $Q(S,a)$, strategy s, and payoff U.
Initialize the death time of each individual and the birth time of the next individual.
while $t_{now} < MC_{step}$ **do**
  while $t_{update} < t_{now}$ **do**
    for each individual x **do**
      Choose action a from Q-table based on ϵ-greedy policy.
      Take action a and observe the new state S'.
    end for
    Update the weights among individuals $weight$ and obtain a new network structure $\mathcal{G}$.
    Calculate each individual's payoff U based on the game interactions.
    for each individual x **do**
      Update $Q(S,a)$: $Q(S,a) \leftarrow Q(S,a) + \eta[R + \gamma \max_{a'} Q(S',a') - Q(S,a)]$.
      Randomly select a neighbor y and generate a random number p between 0 and 1.
      if $p < P(s_x \leftarrow s_y)$ **then**
        Update strategy: $s_x \leftarrow s_y$.
      end if
    end for
    Increase update time: $t_{update} \leftarrow t_{update} + 1$.
  end while
  Determine the time t_{death} of the first individual to die and the time t_{birth} when the next individual j is born into the system.
  if $t_{birth} < t_{death}$ **then**
    Update current time: $t_{now} \leftarrow t_{birth}$.
    Determine the death time of individual j and the birth time of the subsequent new individual.
    Update the weights among individuals $weight$ and the network structure $\mathcal{G}$.
  else
    Update current time: $t_{now} \leftarrow t_{death}$.
    Update the weights among individuals $weight$ and the network structure $\mathcal{G}$.
  end if
end while

3.2 Evolution and Statistics of the Number of Individuals

To validate the theory presented in subsection 2.1, we begin



(a) Evolutionary curves (b) Statistical distributions

Fig. 4. **Evolutionary curves and statistical distributions of the number of individuals.** In this figure, we show both the (a) evolutionary curves and (b) statistical distribution of the individual number in the system with the birth-death process under various death processes, including power-law, uniform, exponential, and lognormal distributions. In addition, we utilize different markers and colors for comparison among different death processes. The red line indicates the theoretical value for each specific death process.

TABLE 1
The results of the stationary number of individuals under different death processes

Results	power-law distribution	uniform distribution	exponential distribution	lognormal distribution
Theoretical result	240	405	150	309.339
Experimental result	236.233	402.169	150.856	304.971
Relative error	1.569%	0.699%	0.571%	1.412%
Variance	242.141	545.242	127.547	336.829
Skewness	0.141	-0.062	0.208	-0.027
Kurtosis	-0.241	-0.481	-0.099	-0.152

by plotting the evolutionary curves and statistical distributions of the number of individuals in the system under different death processes over time, and the results are shown in Fig. 4. According to the theoretical results, we know that the final size of the system is determined solely by the expected value of the death process distribution and the input rate λ. Therefore, our analysis primarily focuses on the impact of various death process distributions on the number of individuals, including power-law, uniform, exponential, and lognormal distributions. For the power-law distribution, the parameters are set to $\lambda = 2, \alpha = 3$, and $x_{min} = 60$. For the uniform distribution, we set the birth rate to $\lambda = 3$, with the starting and ending values set to $a = 120$ and $b = 150$, respectively. In the case of the exponential distribution, the birth rate is $\lambda = 3$ and the death rate is $\kappa = 0.02$. Lastly, for the lognormal distribution, the birth rate, mean, and standard deviation are set as $\lambda = 5, \upsilon = 3$, and $\phi = 1.5$, respectively.

We set the temporal coordinates of Fig. 4(a) to logarithmic coordinates to better observe the rising and smooth phases of individual population evolution. As time progresses, the evolutionary curves of the number of individuals become stable around $t = 1000$ in all four cases and subsequently fluctuate steadily around the theoretical value, as indicated by the red dashed line. However, the magnitude of these fluctuations varies among the different curves. To show the similarity between theoretical and experimental results more intuitively, we present the theoretical and experimental results for different death processes in Tab. 1, where the theoretical results are calculated according to Eq. 7, and we have $E[N(t)] = \frac{\lambda(\alpha-1)}{\alpha-2}x_{min}$ for power-law distribution, $E[N(t)] = \frac{\lambda}{2}(a+b)$ for uniform distribution, $E[N(t)] = \frac{\lambda}{\kappa}$ for the exponential distribution, and $E[N(t)] = \lambda e^{\upsilon+\phi^2/2}$ for lognormal distribution. The experimental results are averaged using the last 9,000 steps of a total of 10,000 evolutionary steps, and the relative error is obtained by $e = |x^* - x|/x$, where x^* is the experimental result and x is the theoretical one. Besides, we calculate the skewness to characterize the direction and degree of skewness of the data distribution, and the kurtosis to describe the steepness of the data distribution. It can be seen that the maximum relative error in Tab. 1 is just 1.569%, which illustrates that the experimental results are well matched with the theoretical ones. Moreover, the variance becomes larger as the size of the system increases, which is consistent with Eq. 8 and the phenomenon shown in Fig. 4(a), where the curve exhibits greater fluctuations as the scale of the system grows. By comparing the skewness and kurtosis of four data distributions, we get that the data distributions obtained with uniform and lognormal distributions are left-skewed, while the other two are right-skewed, and all the data distributions are thin-tailed.

Then, we capture the number of individuals in the last 9,000 steps of the total of 10,000 evolutionary steps, compute the frequency of each specific number separately, and plot the probability distribution of the four death processes in Fig. 4(b) by treating the frequencies as probabilities according to the law of large numbers. The solid red line is the theoretical distribution derived from Thm. 1. It is evident that the theoretical distributions fit well with the experimental ones. To illustrate this more intuitively, we respectively utilize the Kullback-Leibler (KL) divergence (Eq. 19), the Jensen-Shannon (JS) divergence (Eq. 20), the Pearson correlation coefficient (Eq. 21), and the cosine similarity (Eq. 22) to capture the distance between the theoretical and experimental distributions, and the

TABLE 2
The comparison between the theoretical and experimental distributions under different death processes

Results	power-law distribution	uniform distribution	exponential distribution	lognormal distribution
Kullback-Leibler divergence	0.039	0.045	0.022	0.038
Jensen-Shannon divergence	0.009	0.012	0.005	0.009
Pearson correlation coefficient	0.942	0.957	0.989	0.965
Cosine similarity	0.975	0.977	0.994	0.985

results are demonstrated in Tab. 2.

$$KL(A||B) = \sum A_i \log \frac{A_i}{B_i}, \tag{19}$$

where A and B respectively denote the theoretical and experimental distributions.

$$JS(A||B) = \frac{1}{2}KL(A||M) + \frac{1}{2}KL(A||M), \tag{20}$$

where $M = \frac{1}{2}(A + B)$.

$$r = \frac{E[(A - E(A))(B - E(B))]}{S(A)S(B)}, \tag{21}$$

where $E(X)$ and $S(X)$ represent the expectation and standard deviation of probability distribution X.

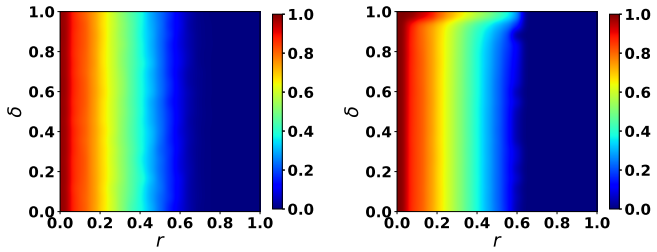
$$cos = \frac{\sum A_i B_i}{\sqrt{\sum A_i^2}\sqrt{\sum B_i^2}}. \tag{22}$$

It can be observed that the KL divergence (all less than 0.045) and JS divergence (all less than 0.012) are very small, while the Pearson correlation coefficient and cosine similarity are very close to 1 in all four death processes, which shows that the theoretical and experimental distributions are in very good agreement with each other, which also proves the validity of our theoretical analysis in subsection 2.1.

3.3 Evolution of Network Cooperation Behaviors

In order to explore the effect of exploitation rate δ and payoff parameter r on the cooperative behaviors of the system, we present the heat maps of the proportion of cooperators on SWBD with power-law distribution and SWOBD about δ and r in Figs. 5(a) and 5(b), respectively. Both simulations share identical parameter settings, differing only in the presence or absence of the birth-death process. Concretely, we respectively set the learning rate, discount factor, and weight fading factor of the edge of the two simulations to 0.7, 0.3, and 2 to ensure the comparability of the results. In addition, the total evolutionary time for both simulations is set to 5,000 long enough to allow the evolution of the cooperation ratio to reach stability.

From Fig. 5(a), we can see that the number of cooperators varies relatively uniformly about the payoff parameter, whereas the exploitation rate has almost no effect on the cooperative behavior for a particular payoff parameter r. However, in Fig. 5(b), a high exploitation rate can promote the emergence of cooperation at a specific r, which is a significant difference between the cooperative behavior of the system with the birth-death process and without. Furthermore, both Figs. 5(a) and 5(b) exhibit that a smaller payoff parameter r will be more favorable to the emergence and maintenance of cooperation. This is because according to the payoff matrix of the snowdrift game indicated in Eq. 13, a higher r results in a larger payoff to the defector and a smaller payoff to the cooperator, which in turn encourages more individuals to prefer the defective strategy. Additionally, we show the effect of δ and r on cooperative behaviors for the death process obeying uniform, exponential, and lognormal distributions, and the interested



(a) SWBD with power-law distribution (b) SWOBD

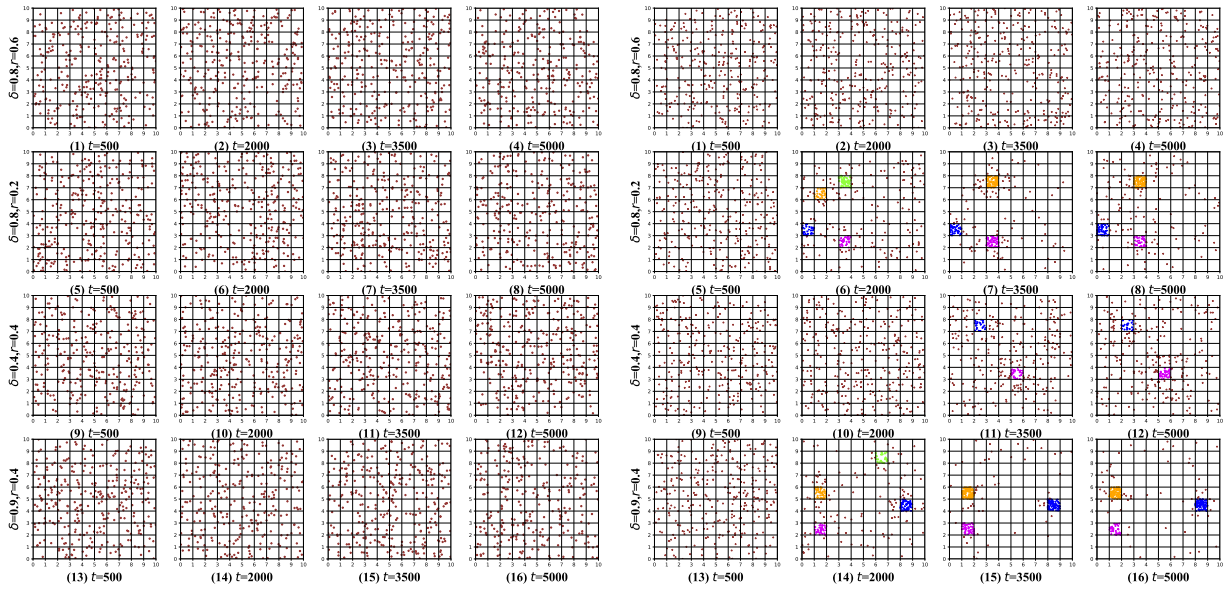
Fig. 5. **Heat maps of cooperation fraction about payoff parameter r and exploitation rate δ.** By setting the y-axis as the exploitation rate δ with a range [0, 1] and the x-axis as the payoff parameter r with a range [0, 1], we demonstrate the heat maps of cooperation fraction on the system with (in panel (a)) and without (in panel (b)) birth-death process with respect to payoff parameter r and exploitation rate δ. The learning rate, discount factor, and weight fading factor of the edge in both subfigures are all set to 0.7, 0.3, and 2.

reader can refer to subsection 3.1 of the supplementary material for more details.

3.4 Emergence and Evolution of Communities

In this subsection, we visually analyze the evolution of individual movement in a two-dimensional space to gain insights into the emergence of system communities under different parameter pairs (δ, r) of (0.8, 0.6), (0.8, 0.2), (0.4, 0.4), and (0.9, 0.4) in Figs. 6(a) and 6(b). Simultaneously, we transform the distribution of individuals in two dimensions to the corresponding network structure based on the weights between them in Figs. 2(a) and 2(b) of the supplementary material. We find that the formation and evolution of the communities in the network structure correspond well to the two-dimensional space. The scale of the system without birth-death process comprises 300 individuals, while the parameters associated with the scale of the system with birth-death process are set to $\lambda = 3$, $x_{min} = 80$, and $\alpha = 5$, i.e., the size of the system as it evolves to a steady state is 320, which can be obtained from Eq. 7 and is very close to the scale of the system without birth-death process.

As displayed in Fig. 6(a), the individuals are consistently evenly distributed on the plane regardless of the values of the exploitation rates and payoff parameters. In addition, the individuals of the corresponding network in Fig. 2(a) of the supplementary material are also distributed randomly with no distinguishing features. However, in Fig. 6(b), when the exploitation rate is high and the payoff parameter is low, we observe that individuals do not maintain an even distribution across locations. Instead, they begin to cluster, forming multiple communities as time progresses and we adopt different colors to mark these different communities that are formed. In Fig. 2(b) of the supplementary material, the communities corresponding to (δ, r) = (0.8, 0.2) and (0.9, 0.4) are clearly identifiable, but exhibit an instability at $t = 2,000$, which are continuing to rapidly evolve and become stable again at $t = 3,500$. Furthermore, by comparing the first and second



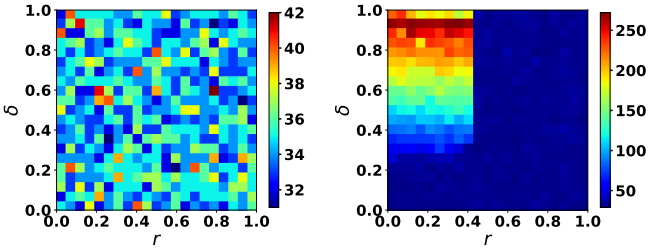
(a) Individual distribution of SWBD with power-law distribution (b) Individual distribution of SWOBD

Fig. 6. **Evolutionary snapshots in the two-dimensional space under different parameter pairs (δ, r).** By setting the parameter pairs (δ, r) as (0.8, 0.6), (0.8, 0.2), (0.4, 0.4), and (0.9, 0.4) from top to bottom and fixing the time step at 500, 2000, 3500, and 5000 from left to right, we present the evolutionary snapshots of individual movement in the two-dimensional space on the system with and without birth-death process in subplots (a) and (b), respectively. All other parameters in the two subplots remain consistent for comparison.

rows (where δ is the same but r differs) of Fig. 6(b), we conclude that a small payoff parameter can promote the emergence and maintenance of communities. Although communities are formed in both the third and fourth rows (where δ differs but r is the same) of Fig. 6(b), the number of individuals in each community structure in (δ, r) = (0.9, 0.4) is clearly larger and appears earlier than that in (δ, r) = (0.4, 0.4). It suggests that a larger exploitation rate is more conducive to the evolution and maintenance of communities. These conclusions are further corroborated by the corresponding network structures shown in Fig. 2(b) of the supplementary material, but the communities are not observed in the system with the birth-death process. In this system, individuals die and new individuals join the system at each moment but individuals cannot form a community in a short period of time, therefore the joining of new individuals is equivalent to entering a location at random. Besides, we investigate the emergence and evolution of communities under the death process following uniform, exponential, and lognormal distributions, and the interested reader can refer to subsection 3.2 of the supplementary material for further details. Furthermore, we examine the emergence and evolution of communities when the stag hunt game is adopted as the underlying game model. Our findings reveal that the SWOBD within the stag hunt game remains capable of fostering the formation of communities. This suggests that the game model employed in this study is not limited to the snowdrift game but can be readily extended to other game scenarios by modifying the payoff matrix. Readers interested in further details are referred to subsection 3.2 of the supplementary material.

3.5 Effect of Payoff Parameter and Exploitation Ratio on Community Structures

Subsequently, to investigate the impact of the payoff parameter and the exploitation rate on community structures, we use the sum of the number of individuals in the four locations with the highest populations throughout the evolutionary process as a measure of community structure (denoted as N_c). In this measure, a larger value indicates a more significant community structure and vice versa. The results of SWBD with power-law distribution and SWOBD are displayed in Figs. 7(a) and 7(b).



(a) SWBD with power-law distribution (b) SWOBD

Fig. 7. **Heat maps of the sum of the first four highest numbers of individuals with respect to payoff parameter r and exploitation rate δ.** By setting the x-axis as the payoff parameter r with a range [0, 1] and the y-axis as the exploitation rate δ with a range [0, 1], we show the impact of r and δ on the sum of the first four highest numbers of individuals on SWBD with power-law distribution (in panel (a)) and SWOBD (in panel (b)).

From Fig. 7(a), we can see that the value of N_c for the system with birth-death process falls within the range of 30 to 42, which is considerably small compared to the scale of the system at the steady state. It suggests that, in this scenario, individuals cannot self-organize themselves to form a community structure. Additionally, the exploitation rates and payoff parameters appear to have minimal impact on the community structure. For the system without the birth-death process, the result shown in Fig. 7(b) illustrates that when the payoff parameter $r < 0.45$ and the exploitation rate $\delta < 0.95$, the value of N_c increases as the exploitation rate grows. In some cases, the maximum value can exceed 250, indicating that more than 83% of the individuals in the system are self-organizing into communities. However, it is worth noting that the value of N_c with $\delta = 1$ is smaller than that with $\delta = 0.95$, demonstrating that the community structure will be better served by an appropriate random selection of individuals as they move around the location. Meanwhile, when $r > 0.45$, the value of N_c remains around 50 regardless of the value of the exploitation rate adopted, i.e., the system does not form communities under these conditions. Moreover, we explore the effect of r and δ on community structures under

the death process obeying uniform, exponential, and lognormal distributions, and the interested reader can refer to subsection 3.3 of the supplementary material for more details.

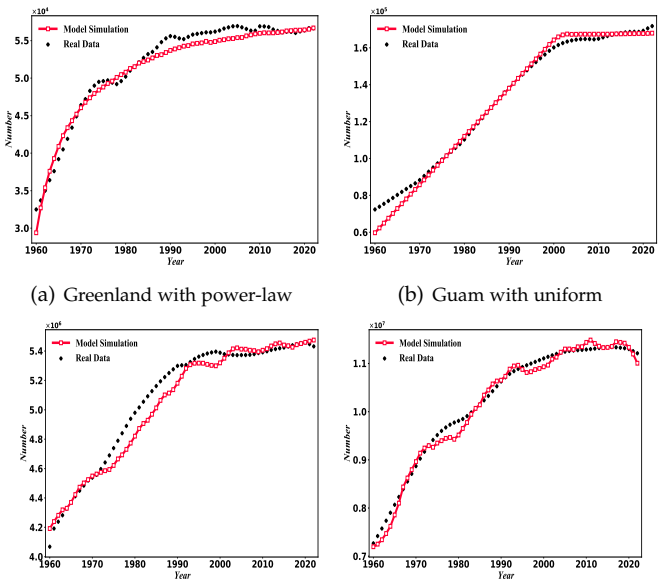
3.6 Fit of the Proposed Model on Real Data

To demonstrate the practical importance of the model proposed in this paper, in this subsection, we apply the scale of the system with the birth-death process under various death distributions to model the populations of four different countries. Furthermore, the degree distributions of networks generated by SWBD and SWOBD with reinforcement learning are employed to fit the degree distributions of six different real networks.

3.6.1 Fit of the Birth-Death Process on Real Populations

We select the populations of Greenland[ii], Guam[iii], the Slovak Republic[iv], and Cuba[v] from 1960 to 2022, spanning a total of 62 years, and model them using the scale of the system with birth-death process under various death distributions. Specifically, we utilize a power-law distribution to model the population of Greenland, setting the input rate at $\lambda = 600$, with the power-law distribution parameters $x_{min} = 50$ and $\alpha = 3$. For Guam's population, we employ a uniform distribution, with the input rate set to 525, and the starting and ending values set to 310 and 330, respectively. The population of the Slovak Republic is modeled using an exponential distribution, with an input rate of 89,600 and a death rate of 1/73. For the population of Cuba, we use a lognormal distribution with an input rate of 291,200, a mean of 2.5, and a standard deviation of 1.8. The fitting results for these four regions are presented and compared with the real data in Fig. 8. All figures show that the populations of the four countries gradually increase and then tend to stabilize, and the simulation results closely match the real data.

To provide a more intuitive demonstration of the fit, we calculate the expected populations for the four countries alongside the simulation results. We then determine the relative error using the equation: $e = \mid x - x^* \mid /x$, where x represents the actual population and x^* denotes the simulated population. Furthermore, we compute the correlation coefficient between the simulation results and the actual populations. The results are summarized in Tab. 3, with the relative errors for Greenland, Guam, the Slovak Republic, and Cuba being 0.96%, 0.48%, 0.71%, and 0.38%, respectively, which are all below 1%. Besides, the correlation coefficients for all four fits exceed 98.39%, further confirming the accuracy of the fit, as illustrated in Fig. 8.



(a) Greenland with power-law (b) Guam with uniform

(c) Slovak Republic with exponential (d) Cuba with lognormal

Fig. 8. **Fit of the birth-death process under different death distributions on the populations of four different countries.** In this figure, we simulate the populations of four countries using different death process distributions. Specifically, we use the power-law distribution for Greenland (subplot (a)), and the uniform distribution for Guam (subplot (b)). Subplots (c) and (d) illustrate the simulations for the Slovak Republic and Cuba, where the death processes follow exponential and lognormal distributions. The actual population data is represented by black diamonds, while the simulated curves are depicted with red squares.

3.6.2 Fit of the Degree Distribution using Reinforcement Learning on Real Networks

Next, we analyze the degree distributions of networks generated by both SWOBD and SWBD with reinforcement learning and utilize them to fit with the degree distributions of six distinct real-world networks [62]: HS-HT, TWITTER-COPEN, AVES-WEAVER-SOCIAL, DWT-607, MAMMALIA-DOLPHIN, and NETSCIENCE. The HS-HT network is a biological network, where nodes represent genes and edges signify connections between genes. The TWITTER-COPEN network is a retweet network consisting of 761 users as nodes and 1,029 retweets/mentions as edges. AVES-WEAVER-SOCIAL and MAMMALIA-DOLPHIN are animal social networks, where nodes denote animals and edges represent interactions between them. DWT-607 is a miscellaneous network based on a symmetric connection table from DTNSRDC, Washington. Finally, NETSCIENCE is a collaboration network containing 379 researchers as nodes and 914 co-authorships as edges. Some statistical information about these six real networks, such as the number of nodes, number of edges, assortativity, clustering coefficients, etc., are given in Tab. 4. The statistics reveal that all networks have relatively low densities and the NETSCIENCE network has the largest clustering coefficient among the six networks, indicating that individuals in the NETSCIENCE network tend to have the closest interactions.

We utilize the network generated by the SWOBD model to fit HS-HT and TWITTER-COPEN networks, with the fitting results shown in Figs. 7(a) and 7(b) of the supplementary material, respectively. Additionally, we apply the SWBD model to fit the real networks based on different death processes, with the results displayed in Figs. 7(c)-(f) of the supplementary material, where these death processes separately follow power-law, uniform, exponential, and lognormal distributions. As observed in Fig. 7 of the supplementary material, the degree distributions of HS-HT and TWITTER-COPEN networks, along with the network generated without the birth-death process, exhibit characteristics consistent with the power-law distribution. In contrast, the degree distributions of the AVES-WEAVER-SOCIAL network and the network generated with a birth-death process that follows power-law distribution approximate the normal distribution. Notably, the network generated by SWBD can also obey a power-law distribution, as illustrated in Fig. 7(f) of the supplementary material. For the fitting of DWT-607 and MAMMALIA-DOLPHIN networks, we model the death process in the SWBD model with uniform and exponential distributions, respectively, and the degree distributions of the resulting networks display irregular patterns, as depicted in Figs. 7(d) and 7(e) of the supplementary material.

To more intuitively assess the fit of the SWOBD and SWBD models, we calculate cosine similarity (Eq. 22) and Jensen-Shannon divergence (Eq. 20) to quantify the similarity between the real and simulated data, as shown in Tab. 4. The results indicate that the Jensen-Shannon divergence between the simulated data and real data is relatively small, not exceeding 0.0997. Conversely, the cosine similarity is notably high, with values

ii. https://data.worldbank.org/indicator/SP.POP.TOTL?locations=GL
iii. https://data.worldbank.org/indicator/SP.POP.TOTL?locations=GU
iv. https://data.worldbank.org/indicator/SP.POP.TOTL?locations=SK
v. https://data.worldbank.org/indicator/SP.POP.TOTL?locations=CU

TABLE 3
The comparison of models under different death processes in fitting real populations

Results	Simulation population	Real population	Relative error	Correlation coefficient
Greenland	51,216.46	51,713.24	0.96%	98.96%
Guam	131,359.27	131,996.69	0.48%	99.72%
Slovak Republic	5,039,330.92	5,075,572.91	0.71%	98.39%
Cuba	10,170,945.35	10,209,542.11	0.38%	99.49%

TABLE 4
Statistical results of real networks and similarity between simulated degree distributions and real degree distributions

Results	Nodes	Edges	Assortativity	Clustering coefficient	Density	JS divergence	Cosine similarity
HS-HT	2,570	13,691	0.2943	0.1695	0.0042	0.0268	0.9877
TWITTER-COPEN	761	1,029	-0.099	0.0759	0.0036	0.0884	0.9732
AVES-WEAVER-SOCIAL	445	1,423	0.2295	0.6924	0.0144	0.0997	0.7563
DWT-607	607	2,262	0.1861	0.4792	0.0123	0.0398	0.9823
MAMMALIA-DOLPHIN	62	159	-0.0436	0.2589	0.0841	0.0095	0.9852
NETSCIENCE	379	914	-0.0817	0.7412	0.0128	0.0502	0.9645

no less than 0.7563. These findings suggest that the difference between the degree distribution generated by our models and the real data is minimal.

Therefore, based on the fitting of our model to real-world data, we can conclude that our model demonstrates strong consistency with actual population dynamics and real network structures. This consistency indicates that the model proposed in this paper has significant practical value.

4 Analysis of the Community Structures

As we have shown before, community structures predominantly emerge in systems, when there is no process of birth and death. In this section, our focus shifts to analyzing the effect of several parameters, including the learning rate, the discount factor in Q-learning, and the dimensions in the two-dimensional space (number of rows and columns) on community formation and development in SWOBD. In alignment with the approach detailed in subsection 3.5, we employ the sum of the individuals in the four locations with the highest populations as a metric for community structure (denoted as N_c). A higher value in this measure signifies a more substantial community structure, and vice versa.

4.1 Impact of Learning Rate on Community Structures

In this subsection, we study the influence of the learning rate η in Eq. 10 on the community structure. The dimension of the two-dimensional space is set to 10×10, with an initial population of 5 individuals at each location, i.e., the total scale of the system is 500. The exploitation rate, discount factor, and payoff parameter are respectively set to 0.7, 0.2, and 0.1. In addition, we set the evolutionary time to 10,000, ensuring sufficient time for the evolution to stabilize. The evolutionary curve of N_c over time under different learning rates is depicted in Fig. 9(a). Figs. 9(b)-(d), (e)-(g), and (h)-(i) illustrate the corresponding network structure at selected instants of time with $\eta = 0.1$, 0.2, and 0.7.

From Fig. 9(a), we observe that the sum of the number of individuals in the four locations with the highest populations grows as time progresses and eventually stabilizes around a certain value. The difference between the various curves lies in the speed at which they evolve to a steady state. Specifically, the curve with $\eta = 0.1$ reaches a steady state around $t = 6,000$, $\eta = 0.2$ achieves stability around $t = 5,000$, $\eta = 0.3$ reaches a steady state around $t = 4,000$, while the curves with $\eta = 0.7$ and 0.9 achieve stability around $t = 3,000$. This trend is visually demonstrated by examining the corresponding network structures. For instance, the network structure associated with $\eta = 0.1$ in Figs. 9(b)-(d) reveals 1, 5, and 4 communities at $t = 3,000$, $t = 5,000$, and $t = 7,000$ respectively, and the corresponding values of N_c at these times are notably distinct, which suggests that its evolution has not yet reached a steady state at $t = 5,000$. On the contrary, the network structure with $\eta = 0.7$ in Figs. 9(h)-(i) exhibits 4 communities at both $t = 3,000$ and 7,000, with similar values for N_c, which indicates that it has reached a stable state already at around $t = 3,000$. Hence, $\eta = 0.2$ in Figs. 9(e)-(g) can be obtained by a similar analysis. It is noteworthy that although the rate of evolution varies among the curves, the value of N_c upon reaching a stable state is similar. Therefore, we derive that the learning rate influences the speed of community structure formation and evolution, with a large learning rate facilitating a quicker emergence and stabilization of community structures. This is attributed to a larger learning rate enabling individuals to learn more rapidly from their actions, resulting in a faster convergence of their Q-tables.

4.2 Effect of Discount Factor on Community Structures

Subsequently, we examine the impact of the discount factor γ in Eq. 10 on the community structure. We present the evolutionary curves of the value of N_c about time t in Fig. 10(a), along with some individual distributions in the two-dimensional space in Figs. 10(b)-(h). The exploitation rate, learning rate, and payoff parameter are separately set to 0.7, 0.9, and 0.1. The scale of the two-dimensional space is 10×10 and there are 5

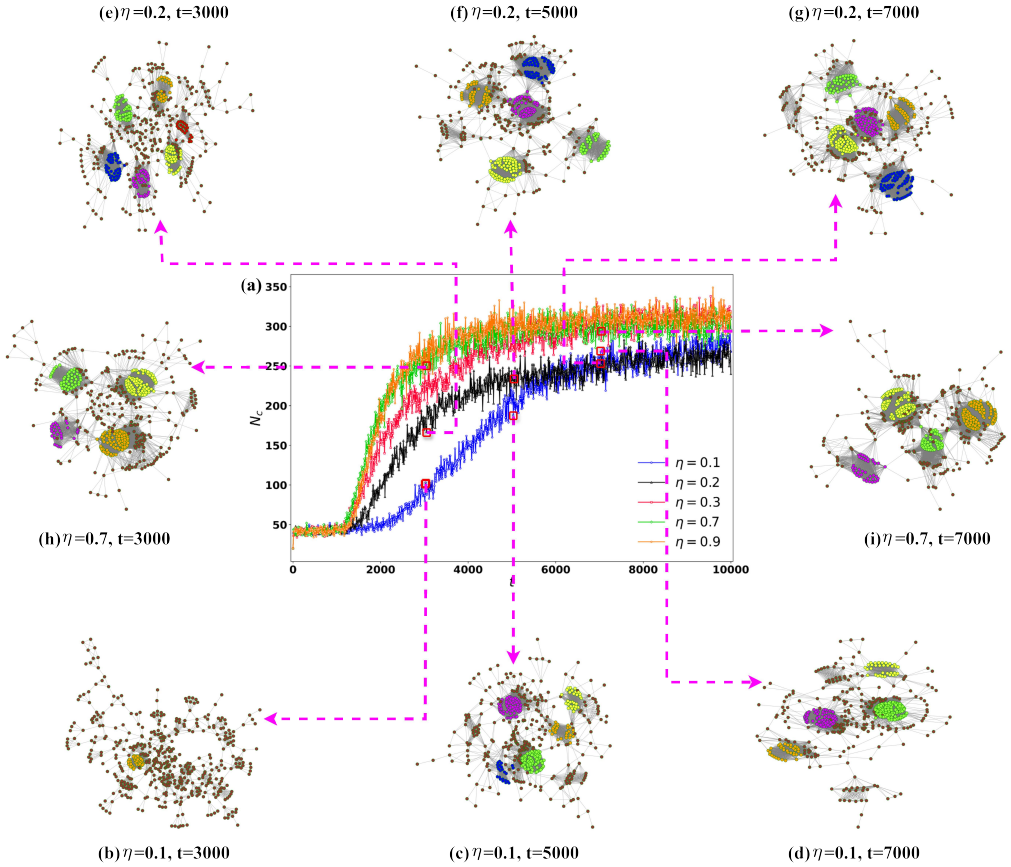


Fig. 9. **Impact of learning rate on the formation and evolution of community structure.** (a) The sum of the number of individuals in the four locations with the highest populations (denoted as N_c) as a function of time t under different learning rates with $\eta = 0.1, 0.2, 0.3, 0.7$, and 0.9. (b)-(i) The corresponding networks at different given learning rates and times, utilizing distinct colors to demonstrate the community structures.

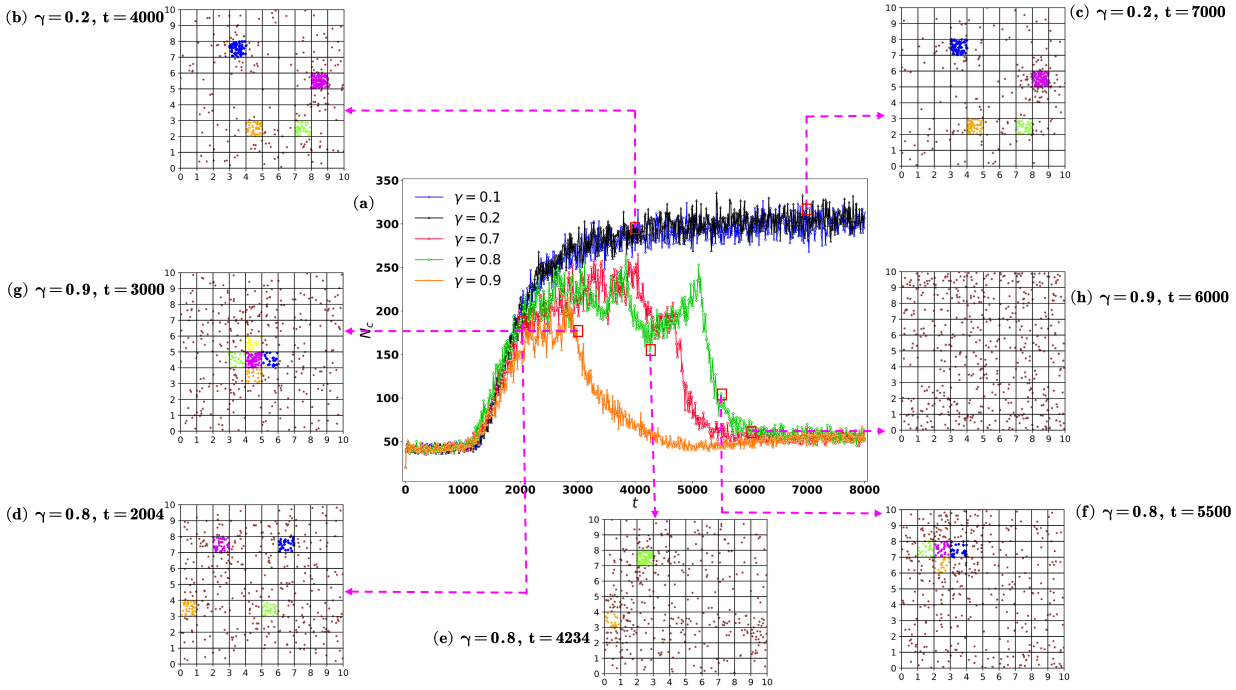


Fig. 10. **Effect of discount factor on the formation and development of community structure.** Subplot (a) demonstrates the evolution of N_c as time evolves under different discount factors with $\gamma = 0.1, 0.2, 0.7, 0.8$, and 0.9. The snapshots of individual distributions in the two-dimensional are depicted in (b)-(c) for $\gamma = 0.2$ at $t = 4,000$ and $t = 7,000$, in (d)-(f) for $\gamma = 0.8$ at $t = 2,004$, $t = 4,234$, and $t = 5,500$, and in (g)-(h) for $\gamma = 0.9$ at $t = 3,000$ and $t = 6,000$.

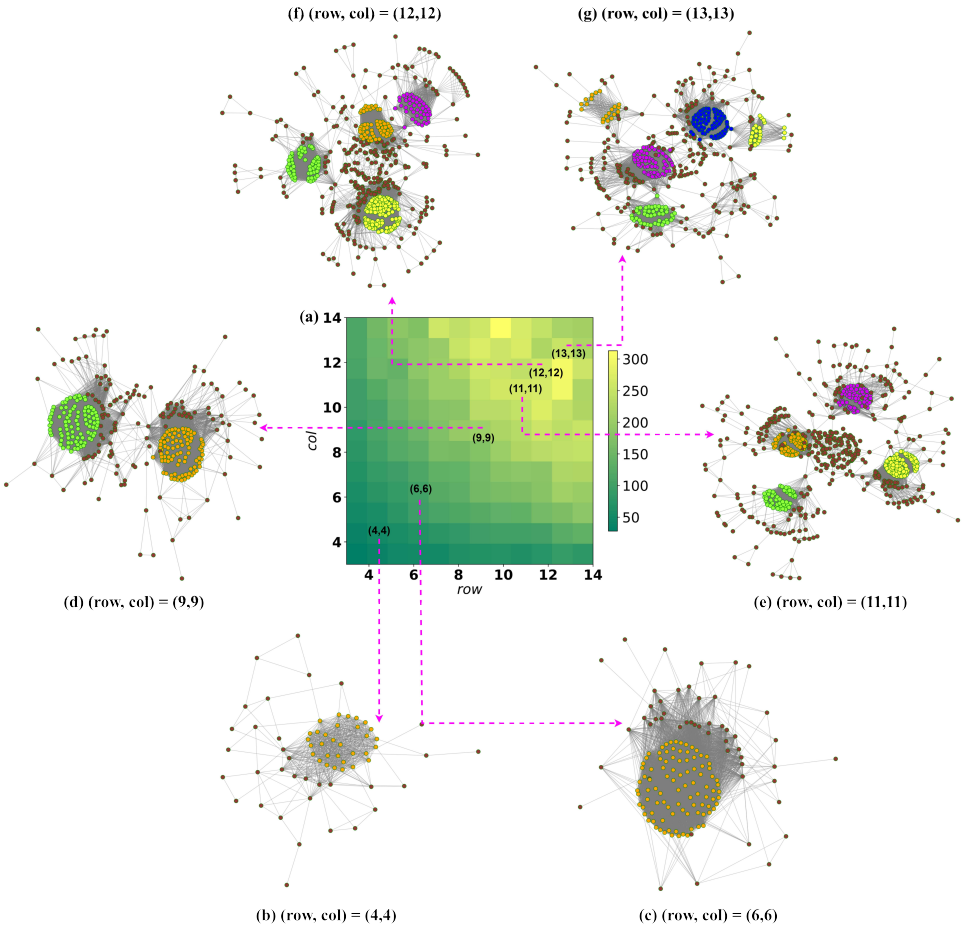


Fig. 11. **Influence of the number of rows and columns in the two-dimensional space on the emergence and development of community structure.** Panel (a) presents the heat map of the sum of the first four highest numbers of individuals in relation to the number of rows and columns. Panels (b)-(g) show the network structure under different values [(4, 4), (6, 6), (9, 9), (11, 11), (12, 12), (13, 13)] of the number of rows and columns (row, col), respectively.

individuals at each location at the beginning, leading to a total system scale of 500.

As depicted in Fig. 10(a), when the discount factor γ is relatively small (γ = 0.1, 0.2), N_c gradually increases and then plateaus with the passage of time. Conversely, when γ takes larger values (γ = 0.7, 0.8, 0.9), N_c exhibits an initial increase, followed by significant fluctuations over time, ultimately reducing to approximately 50. This indicates a lack of sustained community formation, a point corroborated by Fig. 10(h). Furthermore, the snapshots demonstrated in Figs. 10(b)-(c) illustrate that there is already a substantial community formation at $t = 4,000$, which remains stable as time progresses. However, for larger values of γ, although communities emerge during the evolution process, the structures are unstable and undergo significant changes, such as in Figs. 10(d)-(f) with $\gamma = 0.8$, where the number of communities at $t = 2,004$, $t = 4,234$, and $t = 5,500$ are 4, 2, and 4, respectively, as well as the positions of the emerging communities continuously change, and the community eventually disappears around $t = 6,000$, after which the value of N_c stabilizes. This observation underscores the influence of the discount factor on the stability of community structures. A larger γ, indicative of individuals placing greater emphasis on future rewards, proves detrimental to the stability of the community.

4.3 Influence of Dimensions in the Two-Dimensional Space on Community Structures

We note that in the preceding numerical simulations, the dimensions of the two-dimensional space were held constantly at 10×10. In this subsection, we introduce variability by treating the number of rows and columns as parameters to study their influence on the community. The number of individuals at each location at $t = 0$ is 4, and the initial configuration implies that the system scale changes as the number of rows and columns varies. With the learning rate, discount factor, exploitation rate, and payoff parameter set to 0.9, 0.2, 0.7, and 0.1, respectively, we show the heat map of N_c concerning the number of rows and columns in Fig. 11(a). Figs. 11(b)-(g) demonstrate the network structure corresponding to (row, col) = (4, 4), (6, 6), (9, 9), (11, 11), (12, 12), and (13, 13). Different colors in the network represent different community structures formed.

According to the results shown in Fig. 11(a), it can be seen that N_c increases steadily with the growth of the number of rows and columns, which illustrates the consistent formation of communities in the system, with the size of the community progressively expanding. Further insights can be gleaned from the networks presented in Figs. 11(b)-(g). Under (row, col) = (4, 4) and (6, 6), only one community is observed, whereas the communities increase to 2, 4, 4, and 5 under (row, col) = (9, 9), (11, 11), (12, 12), and (13, 13). This observation suggests that the number of communities grows proportionally with the number of rows and columns. When the number of rows and columns is small, the limited space constrains individual movement, resulting in a small number of communities formed but with a large number of individuals compared to the total population. Conversely, with a larger number of rows and columns, individuals have more space to choose and move,

leading to the formation of more communities, albeit with fewer individuals in each community.

5 Model Comparisons

In this section, we compare and analyze our proposed model from two main aspects. On the one hand, from the reinforcement learning point of view, we verify how the use of heuristic models affects the results in the absence of reinforcement learning mechanisms. On the other hand, from the perspective of the classical network model, the model proposed in this paper is compared and analyzed with the classical network in terms of network structure, clustering coefficients, and degree distribution.

5.1 Comparison with Heuristic Model

Hereby, we focus on the impact of reinforcement learning on the outcomes and propose an alternative heuristic model. Specifically, each individual begins with random exploration, which can be considered as setting the initial exploration rate to $\epsilon = 1$. The exploration rate then decays by a factor of 0.99 at each time step, and each individual explores its neighboring states with a probability ϵ. If an individual finds a neighboring state that offers greater benefits than its current state, it will transition to the new state; otherwise, it will retain its original state. It is important to note that state and payoff calculations remain the same as in SWOBD, except that a heuristic model is used instead of reinforcement learning.

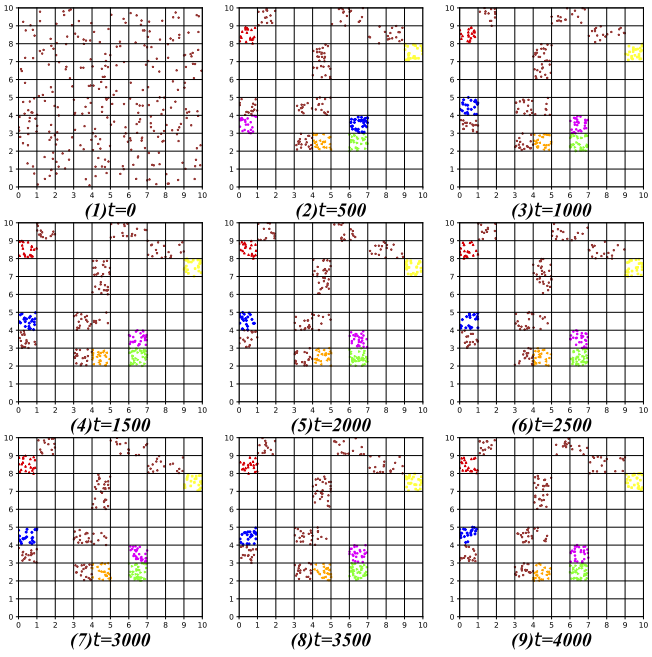


Fig. 12. **Evolutionary snapshots of the heuristic model.** This figure illustrates the snapshots in the two-dimensional space generated based on the heuristic model at different time steps, where the time step is taken from 0 to 4,000 at an interval of 500.

In Fig. 12, we demonstrate evolutionary snapshots of the heuristic model in a two-dimensional space. These snapshots reveal that a community structure begins to emerge around $t = 500$, but remains nearly constant thereafter. Additionally, Fig. 8 of the supplementary material displays the performance metrics of the heuristic model, including degree distributions and state transition ratios. The state transition ratio means the proportion of individuals that change their states at each time step. The degree distribution, shown in Fig. 8(a) of the supplementary material, indicates that the network generated by the heuristic model does not follow typical normal or power-law distributions, but rather exhibits an irregular distribution. Moreover, Fig. 8(b) of the supplementary material shows that the state transition ratio of the heuristic model rapidly decreases to 0, suggesting that individuals maintain their states unchanged after stabilization. This observation contradicts the real-world scenario, where community structures are typically dynamically stabilizing and do not remain constant indefinitely. In contrast, for SWOBD with a reinforcement learning mechanism, the final state transition ratio stabilizes around 0.25, indicating that the community structure evolves dynamically and maintains a dynamic equilibrium, which aligns with real-world observations.

The results above demonstrate that SWOBD with reinforcement learning is more effective in modeling the evolution of complex networks and community structures, and it cannot be replaced by a simple heuristic model. This effectiveness arises from several key differences. In the heuristic model, individuals lack memory and learning capabilities, focusing solely on maximizing immediate benefits. Consequently, individuals in the same state take identical actions. In contrast, the reinforcement learning model provides each individual with a Q-table and the ability to learn, allowing them to consider long-term cumulative payoffs, which contributes to the fact that even individuals in the same state may act differently based on their different experiences. Thus, it can be concluded that reinforcement learning is crucial to the effectiveness of the proposed SWOBD model.

5.2 Comparison with Classic Networks

In this subsection, we perform a comparative analysis between the model proposed in this paper (both SWBD with power-law distribution and SWOBD) with classic networks, including Watts-Strogatz small-world (WS) and Barabási-Albert scale-free (BA) networks. The comparison spans aspects such as network structure, degree distribution, clustering coefficient, etc., and the results are demonstrated in Fig. 13. All networks have 300 nodes (SWBD with power-law distribution also reaches 300 nodes when it evolves to a stable state). In the WS network, each node initially connects to the 2 nearest nodes on its left and right sides, with a subsequent reconnection probability set to 0.3 for each edge, and no repeated edges or self-loops are permitted. The new nodes in the BA network are linked to the 3 existing nodes through the degree-preferential connection mechanism.

In complex networks, the clustering coefficient of node i with degree k_i is defined as the ratio of the number of connected edges between neighboring nodes of node i to the maximum possible number of connected edges between these neighboring nodes [63], which can be calculated using the following formula: $C_i = \frac{2E_i}{k_i(k_i-1)}$, where E_i represents the actual number of edges between the k_i neighbors of node i. If node i has only one neighbor or no neighbors, we denote $C_i = 0$. The clustering coefficient of the network is the average of the clustering coefficient of all the nodes in the network, which reflects the degree of node aggregation in the network, with a higher clustering coefficient indicating a greater node aggregation, and vice versa. By plotting the distribution of node clustering coefficients in the networks (as shown in the second row of Figs. 13(e)-(h)) and calculating the clustering coefficients of the networks, we get that most nodes in the WS and BA networks exhibit low clustering coefficients, with overall network clustering coefficients of 0.201 and 0.079, respectively, whereas the majority of nodes in the networks of SWOBD and SWBD with power-law distribution display high clustering coefficients, with overall network clustering coefficients of 0.751 and 0.714, respectively. This phenomenon suggests that the degree of aggregation in BA and WS networks is low, while in SWOBD and SWBD with power-law distribution, it is high.

Figs. 13(i)-(l) present the degree distributions of the four networks. It is evident that the degree distributions of WS and BA networks follow normal and power-law distributions, respectively, while the degree distributions of SWBD with power-law

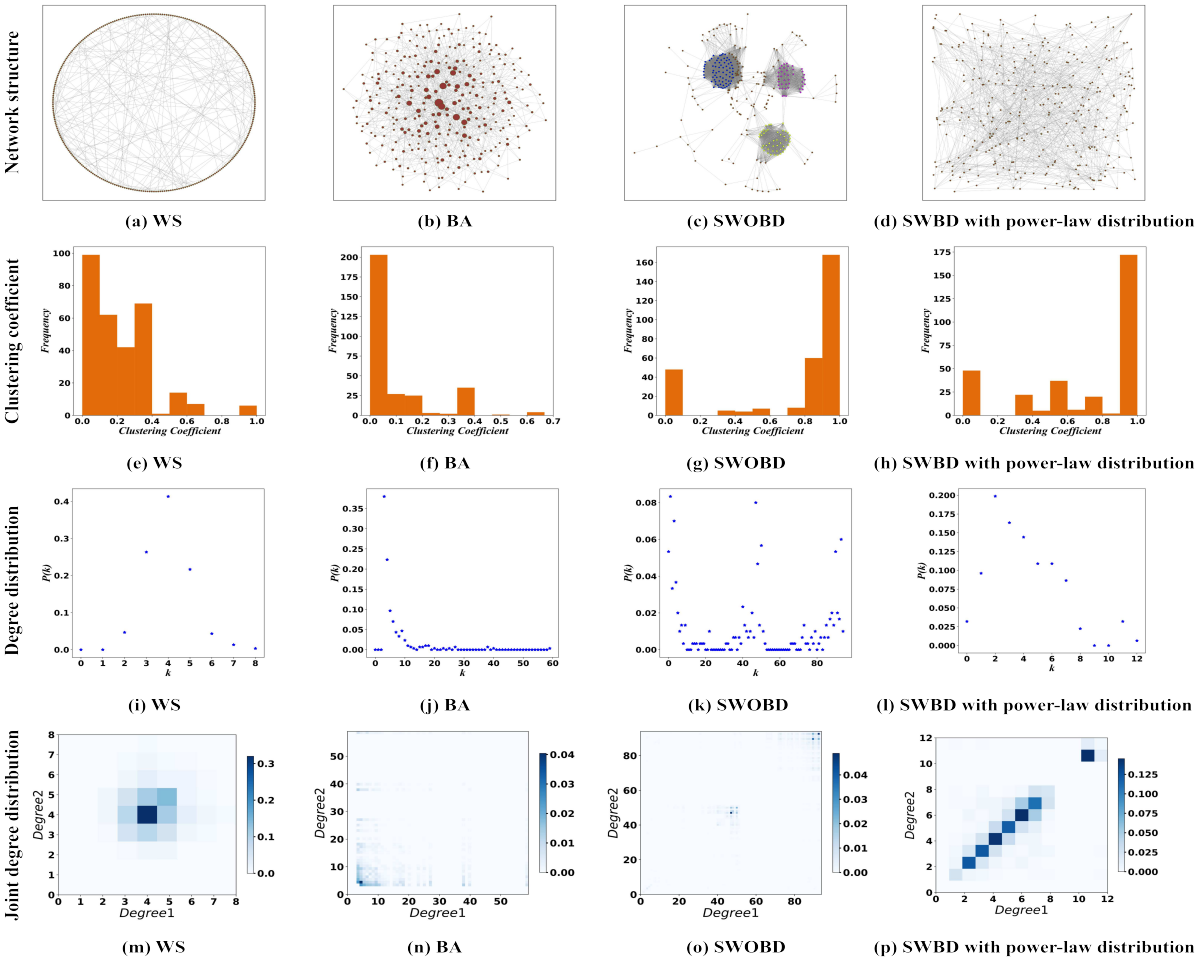


Fig. 13. **Comparison of the networks proposed in this paper with classical networks.** In this figure, we respectively compare our SWOBD network (third column) and the network of SWBD with power-law distribution (fourth column) with the WS small-world network (first column) and the BA scale-free network (second column) in terms of network structure in subplots (a)-(d), clustering coefficient in subplots (e)-(h), degree distribution in subplots (i)-(l), and joint degree distribution in subplots (m)-(p).

distribution and SWOBD networks exhibit irregular patterns due to parameter settings of the model, and we emphasize that the degree distributions of SWOBD and SWBD with power-law distribution can also conform power-law and normal distributions, as demonstrated in Fig. 7 of the supplementary material. Degree correlation is another metric used to assess the connection between nodes with similar or dissimilar degrees. A network is considered degree-degree positively correlated or assortative if nodes with large degrees tend to connect with other nodes with large degrees, and vice versa for degree-degree negatively correlated or disassortative. In this paper, we employ the joint degree distribution to characterize the degree correlation. The joint probability $P(j,k)$ is defined as the probability that the two endpoints of a randomly selected edge in the network have degrees j and k, respectively [64]. It represents the proportion of the number of edges that exist between nodes with degrees j and k in the network to the total number of edges and can be expressed as: $P(j,k) = \frac{m(j,k)\mu(j,k)}{2M}$, where M is the number of edges in the network, $m(j,k)$ indicates the number of connected edges between nodes with degree j and nodes with degree k. If $j = k$, then $\mu(j,k) = 2$, otherwise, $\mu(j,k) = 1$.

By observing the joint degree distribution in the fourth row of Fig. 13 and calculating the assortativity coefficient of the network, we get that most edges in the BA network (assortativity coefficient of -0.082) are connected between small-degree nodes and small-degree nodes. Besides, due to the presence of hub nodes, many small-degree nodes are connected to large-degree nodes, indicating that the BA network is disassortative. However, the WS network (assortativity coefficient of -0.035) shows most connected edges from moderate-degree nodes to other moderate-degree nodes, suggesting a neutral network. Concerning SWOBD (assortativity coefficient of 0.807) and SWBD with power-law distribution (assortativity coefficient of 0.877) in Figs. 13(o) and (p), the values in and around the diagonal in the heat maps of the joint degree distribution are relatively large, suggesting that both networks are assortative and the degree correlation is high. In addition, the network structure, clustering coefficient, degree distribution, and joint degree distribution of networks generated by the death process following uniform, exponential, and lognormal distributions are studied, and the interested reader can refer to subsection 3.6 of the supplementary material for more details.

6 Conclusions and Outlooks

Complex network modeling has always played a significant role in network science, which not only reveals the formation mechanisms in real networks but is also crucial in studying structural dynamical systems. In this paper, we propose a novel complex network model based on the game between individuals with Q-learning in reinforcement learning and take the lifetime of individuals into account, which follows

an arbitrary distribution. Besides, we derive the probability distribution of the system size when it evolves to a steady state and the expected scale of the system. In the numerical simulations, we initially study the stable distribution of the number of individuals under various death processes governed by different distributions, find that the numerical simulation results coincide with the theoretical derivation, and verify the correctness of the theoretical derivation by typical statistical indices, such as relative error and Kullback-Leibler divergence. Subsequently, we observe the evolution of network cooperative behaviors, reveal the emergence and evolution of community structure, and investigate the effects of the payoff parameter and exploitation rate on community structure on both SWBD and SWOBD. We discover that the exploitation rate and the presence of the birth-death process have no significant effect on cooperators, while decreasing the payoff parameter can effectively promote the emergence of cooperators. Besides, we gain that a smaller payoff parameter and a larger exploitation rate will be more conducive to the evolution of community structure on SWOBD and SWBD with lognormal distribution, while on SWBD with power-law, uniform, and exponential distributions, it is difficult for a community structure to emerge regardless of the values of the payoff parameter and the exploitation rate. Furthermore, we apply our model to real situations, namely, fit i) the population evolution of four distinct countries and ii) the degree distributions of six different real networks. Whether in fitting population sizes using the birth-death process or modeling degree distributions in complex networks generated through reinforcement learning, the model proposed in this paper performs well in both cases, demonstrating its excellent practical values. Next, we perform an in-depth analysis of the community structure shaped by SWOBD and reach the conclusion that the learning rate determines the speed of community formation, the discount factor governs the stability of the community, and the dimension of the two-dimensional space impacts the size of the community. Then, we demonstrate the critical role of reinforcement learning in the evolution of complex networks by comparing it with a heuristic model as an alternative approach. Furthermore, through the comparative assessment of our model with classic BA and WS networks concerning network structure, degree distribution, clustering coefficients, etc., we find that our model is capable of generating networks that conform to both power-law and normal distributions, and the networks exhibit high clustering coefficients and assortativity characteristics than BA and WS networks.

The model presented in this paper provides a fresh perspective on the formation and development of community structures in the real world. It also furnishes a valuable framework for studying the dynamic behavior of populations, encompassing phenomena, such as the spread of diseases, the evolution of cooperative behavior, and the synchronization of individuals. Additionally, it is worth emphasizing that the snowdrift game utilized in this paper is not restrictive. The proposed framework is broadly applicable and can be extended to other games by appropriately modifying the payoff matrix. However, there are still some areas for improvement. In particular, beyond the Fermi function utilized in this paper to update individual strategies, other update rules, such as best-take-over and the Moran process, are also worth considering for strategy updates. Besides, expanding the movement of individuals in the two-dimensional plane to the three-dimensional space is also an interesting point. In addition to the evolution of cooperative behavior explored in this paper, the network model developed herein can also serve as a foundational structure for investigating other forms of prosocial behavior on complex networks, including fairness [65], trust [66], the safe development of artificial intelligence [67], technological coordination [68], and open data management [69]. In the near future, these issues will be the focus of our studies.

REFERENCES

[1] J.-P. Onnela, J. Saramäki, J. Hyvönen, G. Szabó, D. Lazer, K. Kaski, J. Kertész, and A.-L. Barabási, "Structure and tie strengths in mobile communication networks," *Proceedings of the National Academy of Sciences*, vol. 104, no. 18, pp. 7332–7336, 2007.

[2] S. Ramanathan and M. Steenstrup, "A survey of routing techniques for mobile communications networks," *Mobile Networks and Applications*, vol. 1, pp. 89–104, 1996.

[3] C. A. Heaney and B. A. Israel, "Social networks and social support," *Health Behavior and Health Education: Theory, Research, and Practice*, vol. 4, pp. 189–210, 2008.

[4] L. F. Berkman, "The relationship of social networks and social support to morbidity and mortality.," *Social Support and Health*, pp. 241–262, 1985.

[5] M. Girvan and M. E. Newman, "Community structure in social and biological networks," *Proceedings of the National Academy of Sciences*, vol. 99, no. 12, pp. 7821–7826, 2002.

[6] G. A. Pavlopoulos, M. Secrier, C. N. Moschopoulos, T. G. Soldatos, S. Kossida, J. Aerts, R. Schneider, and P. G. Bagos, "Using graph theory to analyze biological networks," *BioData Mining*, vol. 4, pp. 1–27, 2011.

[7] J. R. Banavar, A. Maritan, and A. Rinaldo, "Size and form in efficient transportation networks," *Nature*, vol. 399, no. 6732, pp. 130–132, 1999.

[8] H. Bast, D. Delling, A. Goldberg, M. Müller-Hannemann, T. Pajor, P. Sanders, D. Wagner, and R. F. Werneck, "Route planning in transportation networks," *Algorithm Engineering: Selected Results and Surveys*, pp. 19–80, 2016.

[9] Z. Lu, C. Xu, B. Du, T. Ishida, L. Zhang, and M. Sugiyama, "Localdrop: A hybrid regularization for deep neural networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 7, pp. 3590–3601, 2021.

[10] B. Schwikowski, P. Uetz, and S. Fields, "A network of protein–protein interactions in yeast," *Nature Biotechnology*, vol. 18, no. 12, pp. 1257–1261, 2000.

[11] L. Zhang, C. Bao, and K. Ma, "Self-distillation: Towards efficient and compact neural networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 8, pp. 4388–4403, 2021.

[12] P. Erdős, A. Rényi, *et al.*, "On the evolution of random graphs," *Publication of the Mathematical Institute of the Hungarian Academy of Sciences*, vol. 5, no. 1, pp. 17–60, 1960.

[13] S. Milgram, "The small world problem," *Psychology Today*, vol. 2, no. 1, pp. 60–67, 1967.

[14] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world'networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.

[15] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.

[16] P. Sen, S. Dasgupta, A. Chatterjee, P. Sreeram, G. Mukherjee, and S. Manna, "Small-world properties of the indian railway network," *Physical Review E*, vol. 67, no. 3, p. 036106, 2003.

[17] M. Rubinov, S. A. Knock, C. J. Stam, S. Micheloyannis, A. W. Harris, L. M. Williams, and M. Breakspear, "Small-world properties of nonlinear brain activity in schizophrenia," *Human Brain Mapping*, vol. 30, no. 2, pp. 403–416, 2009.

[18] A. Clauset, C. R. Shalizi, and M. E. Newman, "Power-law distributions in empirical data," *SIAM Review*, vol. 51, no. 4, pp. 661–703, 2009.

[19] L. A. Adamic and B. A. Huberman, "Power-law distribution of the world wide web," *Science*, vol. 287, no. 5461, pp. 2115–2115, 2000.

[20] W. Lu and T. Chen, "Quad-condition, synchronization, consensus of multiagents, and anti-synchronization of

complex networks," *IEEE Transactions on Cybernetics*, vol. 51, no. 6, pp. 3384–3388, 2019.
[21] Q. Wang, M. Perc, Z. Duan, and G. Chen, "Synchronization transitions on scale-free neuronal networks due to finite information transmission delays," *Physical Review E*, vol. 80, no. 2, p. 026206, 2009.
[22] B. Pi, Z. Zeng, M. Feng, and J. Kurths, "Evolutionary multigame with conformists and profiteers based on dynamic complex networks," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 32, no. 2, 2022.
[23] Z. Wang, L. Wang, A. Szolnoki, and M. Perc, "Evolutionary games on multilayer networks: a colloquium," *The European Physical Journal B*, vol. 88, pp. 1–15, 2015.
[24] Y. Li, B. Pi, and M. Feng, "Limited resource network modeling and its opinion diffusion dynamics," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 32, no. 4, 2022.
[25] C. Hens, U. Harush, S. Haber, R. Cohen, and B. Barzel, "Spatiotemporal signal propagation in complex networks," *Nature Physics*, vol. 15, no. 4, pp. 403–412, 2019.
[26] L. Lü and T. Zhou, "Link prediction in complex networks: A survey," *Physica A: Statistical Mechanics and its Applications*, vol. 390, no. 6, pp. 1150–1170, 2011.
[27] N. Du, B. Wang, and B. Wu, "Community detection in complex networks," *Journal of Computer Science and Technology*, vol. 23, no. 4, pp. 672–683, 2008.
[28] B. A. Travençolo, M. P. Viana, and L. da Fontoura Costa, "Border detection in complex networks," *New Journal of Physics*, vol. 11, no. 6, p. 063019, 2009.
[29] M. A. Nowak and R. M. May, "Evolutionary games and spatial chaos," *Nature*, vol. 359, no. 6398, pp. 826–829, 1992.
[30] D. Challet and Y.-C. Zhang, "Emergence of cooperation and organization in an evolutionary game," *Physica A: Statistical Mechanics and its Applications*, vol. 246, no. 3-4, pp. 407–418, 1997.
[31] H. Ebel and S. Bornholdt, "Coevolutionary games on networks," *Physical Review E*, vol. 66, no. 5, p. 056118, 2002.
[32] C. Hauert and M. Doebeli, "Spatial structure often inhibits the evolution of cooperation in the snowdrift game," *Nature*, vol. 428, no. 6983, pp. 643–646, 2004.
[33] F. C. Santos and J. M. Pacheco, "Scale-free networks provide a unifying framework for the emergence of cooperation," *Physical Review Letters*, vol. 95, no. 9, p. 098104, 2005.
[34] M. A. Nowak, "Five rules for the evolution of cooperation," *Science*, vol. 314, no. 5805, pp. 1560–1563, 2006.
[35] W. Han, Z. Zhang, J. Sun, and C. Xia, "Role of reputation constraints in the spatial public goods game with second-order reputation evaluation," *Chaos, Solitons & Fractals*, vol. 161, p. 112385, 2022.
[36] M. Perc, J. J. Jordan, D. G. Rand, Z. Wang, S. Boccaletti, and A. Szolnoki, "Statistical physics of human cooperation," *Physics Reports*, vol. 687, pp. 1–51, 2017.
[37] M. Chica, J. M. Hernández, and M. Perc, "Rewarding policies in an asymmetric game for sustainable tourism," *Applied Mathematics and Computation*, vol. 457, p. 128183, 2023.
[38] Y. Zhang, Z. Zeng, B. Pi, and M. Feng, "An evolutionary game with revengers and sufferers on complex networks," *Applied Mathematics and Computation*, vol. 457, p. 128168, 2023.
[39] Q. Su, A. McAvoy, L. Wang, and M. A. Nowak, "Evolutionary dynamics with game transitions," *Proceedings of the National Academy of Sciences*, vol. 116, no. 51, pp. 25398–25404, 2019.
[40] M. Feng, B. Pi, L.-J. Deng, and J. Kurths, "An evolutionary game with the game transitions based on the markov process," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 54, no. 1, pp. 609–621, 2023.
[41] T. Wen, J. Cao, and K. H. Cheong, "Gravity-based community vulnerability evaluation model in social networks: Gbcve," *IEEE Transactions on Cybernetics*, vol. 53, no. 4, pp. 2467–2479, 2021.
[42] Z. Zhang and B. Wu, "Topological properties, spectra analysis, and consensus problems for a class of network models based on m-fission operation," *IEEE Transactions on Cybernetics*, vol. 52, no. 9, pp. 9905–9921, 2021.
[43] D. Li, W. Song, and J. Liu, "Complex network evolution model based on turing pattern dynamics," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 4, pp. 4229–4244, 2022.
[44] M. Feng, Y. Li, F. Chen, and J. Kurths, "Heritable deleting strategies for birth and death evolving networks from a queueing system perspective," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 52, no. 10, pp. 6662–6673, 2022.
[45] F. Battiston, G. Cencetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, J.-G. Young, and G. Petri, "Networks beyond pairwise interactions: Structure and dynamics," *Physics Reports*, vol. 874, pp. 1–92, 2020.
[46] S. A. West, S. P. Diggle, A. Buckling, A. Gardner, and A. S. Griffin, "The social lives of microbes," *Annual Review of Ecology, Evolution, and Systematics*, vol. 38, no. 1, pp. 53–77, 2007.
[47] R. Axelrod and R. O. Keohane, "Achieving cooperation under anarchy: Strategies and institutions," *World Politics*, vol. 38, no. 1, pp. 226–254, 1985.
[48] S. Zhang, R.-S. Wang, and X.-S. Zhang, "Uncovering fuzzy community structure in complex networks," *Physical Review E*, vol. 76, no. 4, p. 046103, 2007.
[49] J. A. Dunne, R. J. Williams, and N. D. Martinez, "Network structure and biodiversity loss in food webs: robustness increases with connectance," *Ecology Letters*, vol. 5, no. 4, pp. 558–567, 2002.
[50] C. He, X. Fei, Q. Cheng, H. Li, Z. Hu, and Y. Tang, "A survey of community detection in complex networks using nonnegative matrix factorization," *IEEE Transactions on Computational Social Systems*, vol. 9, no. 2, pp. 440–457, 2021.
[51] N. T. Bailey, *The elements of stochastic processes with applications to the natural sciences*, vol. 25. John Wiley & Sons, 1991.
[52] P. Chakrabarti, B. Satpathy, H. A. Shankar, R. Bapat, H. R. Sherazi, S. Poddar, and J. A. Ware, "Statistical analysis of strategic market management based on neuro-fuzzy model of human nature, poisson process and renewal theory," *Elementary Education Online*, vol. 19, no. 4, pp. 7146–7159, 2020.
[53] J. Brown, "Some tests of the decay theory of immediate memory," *Quarterly Journal of Experimental Psychology*, vol. 10, no. 1, pp. 12–21, 1958.
[54] M. E. Newman, "The structure and function of complex networks," *SIAM Review*, vol. 45, no. 2, pp. 167–256, 2003.
[55] X. Hu and X. Liu, "Unfixed-neighbor-mechanism promotes cooperation in evolutionary snowdrift game on lattice," *Physica A: Statistical Mechanics and its Applications*, vol. 572, p. 125910, 2021.
[56] H. Chiba-Okabe and J. B. Plotkin, "Social learning with complex contagion," *Proceedings of the National Academy of Sciences*, vol. 121, no. 49, p. e2414291121, 2024.
[57] X. Gabaix and Y. M. Ioannides, "The evolution of city size distributions," in *Handbook of Regional and Urban Economics*, vol. 4, pp. 2341–2378, Elsevier, 2004.
[58] F. Olken and D. Rotem, "Random sampling from databases: a survey," *Statistics and Computing*, vol. 5, pp. 25–42, 1995.
[59] J. Bisquert, F. Fabregat-Santiago, I. Mora-Sero, G. Garcia-Belmonte, and S. Gimenez, "Electron lifetime in dye-

sensitized solar cells: theory and interpretation of measurements," *The Journal of Physical Chemistry C*, vol. 113, no. 40, pp. 17278–17290, 2009.
[60] E. Limpert, W. A. Stahel, and M. Abbt, "Log-normal distributions across the sciences: keys and clues: on the charms of statistics, and how mechanical models resembling gambling machines offer a link to a handy way to characterize log-normal distributions, which can provide deeper insight into variability and probability—normal or log-normal: that is the question," *BioScience*, vol. 51, no. 5, pp. 341–352, 2001.
[61] M. E. Newman, "Power laws, pareto distributions and zipf's law," *Contemporary Physics*, vol. 46, no. 5, pp. 323–351, 2005.
[62] R. A. Rossi and N. K. Ahmed, "The network data repository with interactive graph analytics and visualization," in *Association for the Advancement of Artificial Intelligence (AAAI)*, 2015.
[63] T. Zhou, G. Yan, and B.-H. Wang, "Maximal planar networks with large clustering coefficient and power-law degree distribution," *Physical Review E*, vol. 71, no. 4, p. 046141, 2005.
[64] M. E. Newman, "Assortative mixing in networks," *Physical Review Letters*, vol. 89, no. 20, p. 208701, 2002.
[65] T. Cimpeanu, A. Di Stefano, C. Perret, and T. A. Han, "Social diversity reduces the complexity and cost of fostering fairness," *Chaos, Solitons & Fractals*, vol. 167, p. 113051, 2023.
[66] A. Kumar, V. Capraro, and M. Perc, "The evolution of trust and trustworthiness," *Journal of the Royal Society Interface*, vol. 17, no. 169, p. 20200491, 2020.
[67] T. Cimpeanu, F. C. Santos, L. M. Pereira, T. Lenaerts, and T. A. Han, "Artificial intelligence development races in heterogeneous settings," *Scientific Reports*, vol. 12, no. 1, p. 1723, 2022.
[68] N. B. Ogbo, T. Cimpeanu, A. Di Stefano, and T. A. Han, "Shake on it: The role of commitments and the evolution of coordination in networks of technology firms," in *Artificial Life Conference Proceedings 34*, vol. 2022, p. 41, 2022.
[69] T. P. Benko, B. Pi, Q. Li, M. Feng, M. Perc, and H. B. Vošner, "Evolutionary games for cooperation in open data management," *Applied Mathematics and Computation*, vol. 496, p. 129364, 2025.



Bin Pi (Student Member, IEEE) received the B.E. degree in data science and big data technology from the College of Artificial Intelligence, Southwest University, Chongqing, China, in 2023. He is currently pursuing the M.S. degree in mathematics with the School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, China. His research interests include complex networks, evolutionary games, stochastic processes, and reinforcement learning.



Liang-Jian Deng (Senior Member, IEEE) received the B.S. and Ph.D. degrees in applied mathematics from the School of Mathematical Sciences, University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2010 and 2016, respectively. From 2013 to 2014, he was a joint-training Ph.D. Student with Case Western Reserve University, Cleveland, OH, USA. In 2017, he was a Post-Doctoral Researcher with Hong Kong Baptist University (HKBU), Hong Kong. In addition, he has stayed at the Isaac Newton Institute for Mathematical Sciences, University of Cambridge, Cambridge, U.K., and HKBU, for short visits. He is currently a Professor with the School of Mathematical Sciences, UESTC. His research interests include use of optimization modeling, deep learning and numerical PDEs, to address several tasks in image processing and computer vision, e.g., resolution enhancement and restoration. Please visit his homepage for more info.: https://liangjiandeng.github.io/.



Minyu Feng (Senior Member, IEEE) received his Ph.D. degree in Computer Science from a joint program between University of Electronic Science and Technology of China, Chengdu, China, and Humboldt University of Berlin, Berlin, Germany in 2018. Since 2019, he has been an associate professor at the College of Artificial Intelligence, Southwest University, Chongqing, China. Dr. Feng is a Senior Member of IEEE, China Computer Federation (CCF), and Chinese Association of Automation (CAA). Currently, he serves as a Subject Editor for Applied Mathematical Modelling, and an Editorial Board Member for Humanities & Social Sciences Communications, Scientific Reports, and International Journal of Mathematics for Industry. Besides, he is a Reviewer for Mathematical Reviews of the American Mathematical Society. Dr. Feng's research interests include Complex Systems, Evolutionary Game Theory, Computational Social Science, and Mathematical Epidemiology.



Matjaž Perc (Member, IEEE) received his Ph.D. in physics from the University of Maribor in 2006. He is currently Professor of physics at the University of Maribor, staff researcher at the Community Healthcare Center Dr. Adolf Drolc Maribor, Adjunct Professor at Kyung Hee University and Korea University, and External faculty member at the Complexity Science Hub in Vienna. He is a member of Academia Europaea and the European Academy of Sciences and Arts, and among top 1% most cited physicists according to 2020, 2021, 2022, 2023, and 2024 Clarivate Analytics data. He is also the 2015 recipient of the Young Scientist Award for Socio and Econophysics from the German Physical Society, and the 2017 USERN Laureate. In 2018 he received the Zois Award, which is the highest national research award in Slovenia. In 2019 he became Fellow of the American Physical Society. Since 2021 he is also Vice-Dean of Natural Sciences at the European Academy of Sciences and Arts.



Jürgen Kurths received the B.S. degree in mathematics from the University of Rostock, Rostock, Germany, the Ph.D. degree from the Academy of Sciences, German Democratic Republic, Berlin, Germany, in 1983, the Honorary degree from N.I.Lobachevsky State University, Nizhny Novgorod, Russia in 2008, and the Honorary degree from Saratow State University, Saratov, Russia, in 2012. From 1994 to 2008, he was a Full Professor with the University of Potsdam, Potsdam, Germany. Since 2008, he has been a Professor of nonlinear dynamics with the Humboldt University of Berlin, Berlin, Germany, and the Chair of the Research Domain Complexity Science with the Potsdam Institute for Climate Impact Research, Potsdam, Germany. He has authored more than 700 papers, which are cited more than 60000 times (H-index: 111). His main research interests include synchronization, complex networks, time series analysis, and their applications. Dr. Kurths was the recipient of the Alexander von Humboldt Research Award from India, in 2005, and from Poland in 2021, the Richardson Medal of the European Geophysical Union in 2013, and the Eight Honorary Doctorates. He is a Highly Cited Researcher in Engineering. He is a member of the Academia 1024 Europaea. He is an Editor-in-Chief of CHAOS and on the Editorial Boards of more than ten journals. He is a Fellow of the American Physical Society, the Royal Society of 1023 Edinburgh, and the Network Science Society.

Dynamic Evolution of Complex Networks: A Reinforcement Learning Approach Applying Evolutionary Games to Community Structure (Supplementary Material)

Bin Pi, *Student Member, IEEE,* Liang-Jian Deng, *Senior Member, IEEE,* Minyu Feng, *Senior Member, IEEE,* Matjaž Perc, *Member, IEEE,* and Jürgen Kurths

Abstract—This supplementary material provides the proofs for Theorems 1 and 2 from the main manuscript, as well as additional figures that support and complement the main findings. Additionally, further simulation results and analyses related to the SWBD model with uniform, exponential, and lognormal distributions are also provided for interested readers.

Index Terms—Complex networks, Evolutionary games, Reinforcement learning, Community structure, Stochastic process.

1 Proof of Theorem 1

Theorem 1. *For the continuous-time Markov chain $N(t)$ with the state space E, assume that the expectation of death process $\{G(t), t \geq 0\}$ exists, its limiting probability $\{\pi_i, i = 1, 2, \cdots\}$ exists and follows*

$$\begin{aligned}\pi_i &= \lim_{t\to\infty} p_i(t) = \lim_{t\to\infty} P\{N(t) = i\} \\ &= \frac{\{\lambda E[G(t)]\}^i}{i!} e^{-\lambda E[G(t)]}.\end{aligned} \tag{1}$$

Proof. As we stated, new individuals enter the system at time intervals following an exponential distribution with parameter λ and have a lifetime that obeys a general distribution. We yield

$$\begin{aligned}p_i(t) = P\{N(t) = i\} &= \sum_{n=i}^{\infty} P_{n,i}(N \mid A, t) P[A(t) = n] \\ &= \sum_{n=i}^{\infty} \frac{n![X(t)]^i [1 - X(t)]^{n-i}}{i!(n-i)!} \frac{(\lambda t)^n}{n!} e^{-\lambda t} \\ &= \frac{[\lambda t X(t)]^i}{i!} e^{-\lambda t} \sum_{n=i}^{\infty} \frac{[\lambda t - \lambda t X(t)]^{n-i}}{(n-i)!} \\ &= \frac{[\lambda t X(t)]^i}{i!} e^{-\lambda t X(t)}.\end{aligned} \tag{2}$$

Obviously, the probability $p_i(t)$ of the number of individuals in the system at time t is a non-homogeneous Poisson process. Therefore, under the condition that the expectation of the general distribution $\{G(t), t \geq 0\}$ exists, its limiting probability $\{\pi_i, i = 1, 2, \cdots\}$ exists and follows

$$\begin{aligned}\pi_i &= \lim_{t\to\infty} p_i(t) = \lim_{t\to\infty} \frac{[\lambda t X(t)]^i}{i!} e^{-\lambda t X(t)} \\ &= \frac{\lambda^i}{i!} \lim_{t\to\infty} [tX(t)]^i e^{-\lambda t X(t)} \\ &= \frac{\lambda^i}{i!} \left\{ \int_0^{\infty} [1 - G(t-x)] dx \right\}^i e^{-\lambda \int_0^{\infty} [1 - G(t-x)] dx} \\ &= \frac{\{\lambda E[G(t)]\}^i}{i!} e^{-\lambda E[G(t)]}.\end{aligned} \tag{3}$$

The results follow. □

2 Proof of Theorem 2

Theorem 2. *The average scale of the system is*

$$E[N(t)] = \lambda E[G(t)], \tag{4}$$

B. Pi, L.-J. Deng, and M. Feng were supported in part by the National Nature Science Foundation of China (NSFC) under Grant Nos. 12271083, 62206230, and 62273077, in part by the Sichuan Province's Science and Technology Empowerment for Disaster Prevention, Mitigation, and Relief Project under Grant No. 2025YFNH0001, and in part by the Natural Science Foundation of Chongqing under Grant No. CSTB2023NSCQ-MSX0064. M.P. was supported by the Slovenian Research and Innovation Agency (Javna agencija za znanstvenoraziskovalno in inovacijsko dejavnost Republike Slovenije) under Grant Nos. P1-0403 and N1-0232.

Bin Pi and Liang-Jian Deng are with the School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: liangjian.deng@uestc.edu.cn).

Minyu Feng is with the College of Artificial Intelligence, Southwest University, Chongqing 400715, China (e-mail: myfeng@swu.edu.cn).

Matjaž Perc is with Faculty of Natural Sciences and Mathematics, University of Maribor, Koroška cesta 160, 2000 Maribor, Slovenia, with Community Healthcare Center Dr. Adolf Drolc Maribor, Ulica talcev 9, 2000 Maribor, Slovenia, with Department of Physics, Kyung Hee University, 26 Kyungheedae-ro, Dongdaemun-gu, Seoul 02447, Republic of Korea, with Complexity Science Hub, Metternichgasse 8, 1030 Vienna, Austria, and with University College, Korea University, 145 Anam-ro, Seongbuk-gu, Seoul 02841, Republic of Korea.

Jürgen Kurths is with the Department of Complexity Science, Potsdam Institute for Climate Impact Research, 14473 Potsdam, Germany, and also with the Institute of Physics, Humboldt University of Berlin, 12489 Berlin, Germany.

Corresponding authors: Liang-Jian Deng and Minyu Feng.

the variance of the scale is

$$D[N(t)] = \lambda E[G(t)], \tag{5}$$

and the average staying time of each individual is

$$E(T) = E[G(t)]. \tag{6}$$

Proof. The average scale of the system is the expectation of the result in Thm. 1, which can be denoted as

$$\begin{aligned} E[N(t)] &= \sum_{i=0}^{\infty} i\pi_i = e^{-\lambda E[G(t)]} \sum_{i=1}^{\infty} \frac{\{\lambda E[G(t)]\}^i}{(i-1)!} \\ &= \lambda E[G(t)], \end{aligned} \tag{7}$$

the variance is expressed as

$$\begin{aligned} D[N(t)] &= E\left[N^2(t)\right] - E^2[N(t)] \\ &= \sum_{i=0}^{\infty} i^2\pi_i - \{\lambda E[G(t)]\}^2 \\ &= e^{-\lambda E[G(t)]} \sum_{i=1}^{\infty} i\frac{\{\lambda E[G(t)]\}^i}{(i-1)!} - \{\lambda E[G(t)]\}^2 \\ &= \lambda E[G(t)], \end{aligned} \tag{8}$$

and the average staying time of each individual is calculated by

$$E(T) = \frac{E[N(t)]}{\lambda} = E[G(t)]. \tag{9}$$

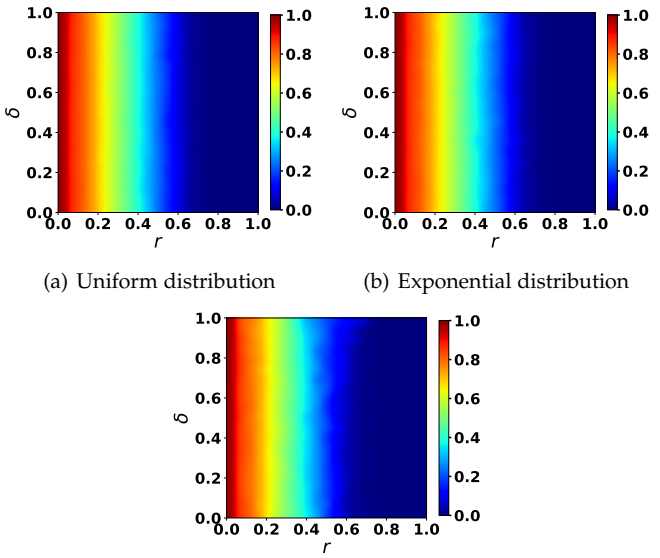
The results follow. □

3 Further Simulation Results

In this section, we present the evolution of cooperative behaviors and communities, the effect of payoff parameter and exploitation ratio on community structures, and the network structure and metrics in systems with the birth-death process (SWBD), where the death process obeys different distributions, including uniform (SWBD with uniform distribution), exponential (SWBD with exponential distribution), and lognormal (SWBD with lognormal distribution) distributions. Moreover, this section includes supplementary figures that provide additional support and clarification for the key results discussed in the main manuscript. We emphasize that in order to ensure comparability, all parameter settings in each simulation are the same as those in the corresponding simulation in the main manuscript, with the exception of the death process distribution. Furthermore, based on the theoretical results derived in subsection 2.1 of the main manuscript, we adjust the parameters for different death distributions to ensure that the system scales remain similar when the evolution reaches stability.

3.1 Further Results for Evolution of Network Cooperation Behaviors

In this subsection, we illustrate the evolution of network cooperation behaviors under the death process following uniform, exponential, and lognormal distributions, and the results are shown in Fig. 1. The learning rate, discount factor, and weight fading factor of the edge are set to 0.7, 0.3, and 2, respectively, consistent with the settings in the main manuscript. In addition, we set parameters for the different death distributions to ensure that the systems are similar in scale when evolutionary stabilization occurs. In particular, for the uniform distribution, we set the birth rate to $\lambda = 4$ and the starting and ending values to $a = 120$ and $b = 135$. The birth and death rates of the exponential distribution are set to $\lambda = 5.1$ and $\kappa = 0.01$. In the case of the lognormal distribution, the birth rate, mean, and standard deviation are set as $\lambda = 4, \upsilon = 2.85$, and $\phi = 2$, respectively. With these settings, the scale of the system at a steady state in all three cases is approximately 510, which aligns with the power-law distribution scenario discussed in the main manuscript.



(a) Uniform distribution (b) Exponential distribution

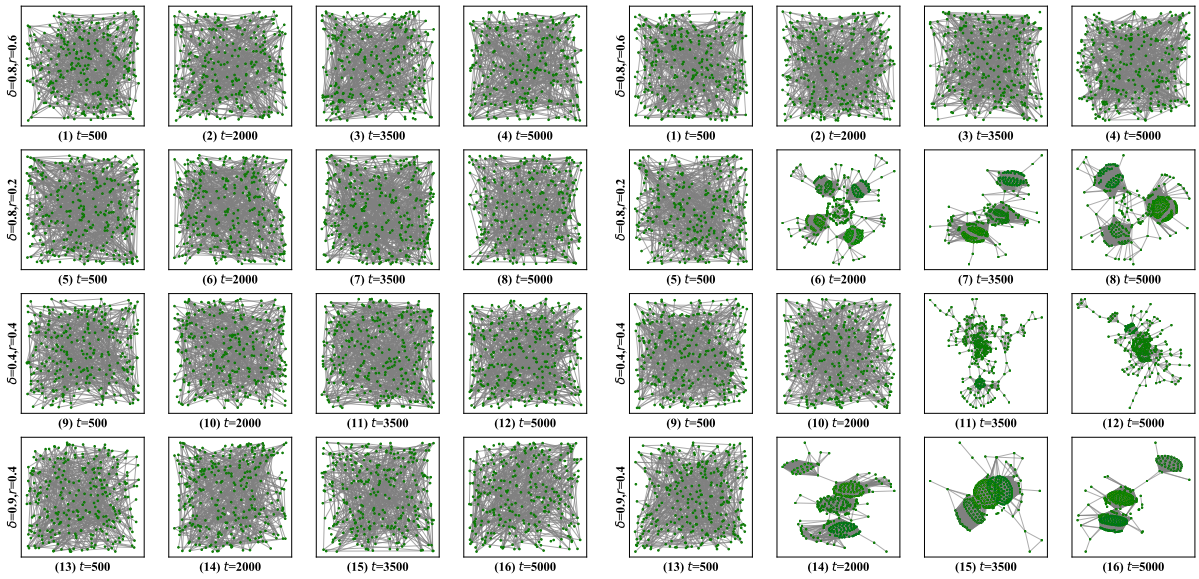
(c) Lognormal distribution

Fig. 1. **Heat maps of cooperation fraction regarding payoff parameter r and exploitation rate δ under different death processes.** The figure demonstrates the effect of r and δ on cooperative behaviors when the death process follows uniform (in subplot (a)), exponential (in subplot (b)), and lognormal (in subplot (c)) distributions. The ranges of both the x-axis and y-axis are set to [0, 1], with the x-axis and y-axis respectively representing the payoff parameter r and exploitation rate δ.

From Fig. 1, we observe that the evolution of the cooperation ratio under the three death distributions is similar. Specifically, the cooperation ratio decreases with the increase of the payoff parameter r, while the exploitation rate δ has almost no effect on the survival of cooperators. The phenomenon is the same as the results in the main manuscript, where the death process obeys a power-law distribution.

3.2 Further Results for Emergence and Evolution of Communities

Fig. 2 illustrates the corresponding network structures for the evolutionary snapshots in the two-dimensional space



(a) Network structure of SWBD with power-law distribution (b) Network structure of SWOBD

Fig. 2. **Network structures corresponding to the evolutionary snapshots in the two-dimensional space shown in Fig. 6 of the main manuscript.** The parameter pairs (δ, r) are set to (0.8, 0.6), (0.8, 0.2), (0.4, 0.4), and (0.9, 0.4) from top to bottom. The time steps are fixed at 500, 2000, 3500, and 5000 from left to right.

depicted in Fig. 6 of the main manuscript, enabling a direct comparison between individual movements and the resulting network formations.

To explore the distribution of individuals in two-dimensional space and their corresponding network structure under the death process obeying other distributions, we select the same four sets of parameters (δ, r) as in the main manuscript, i.e., (0.8, 0.6), (0.8, 0.2), (0.4, 0.4), and (0.9, 0.4). The evolutionary snapshots in the two-dimensional space and their corresponding network structures under different parameter pairs (δ, r) and death processes are demonstrated in Fig. 3.

From these snapshots of Fig. 3, it is evident that, regardless of the distribution followed by the death process, no distinct community structure emerges in the system. The distribution of individuals in the two-dimensional space remains random, which is consistent with the results observed when the death process follows a power-law distribution as presented in the main manuscript.

In addition to the snowdrift game, we further investigate the emergence and evolution of communities in SWOBD within the context of the stag hunt game (SHG), where the payoff matrix is defined as:

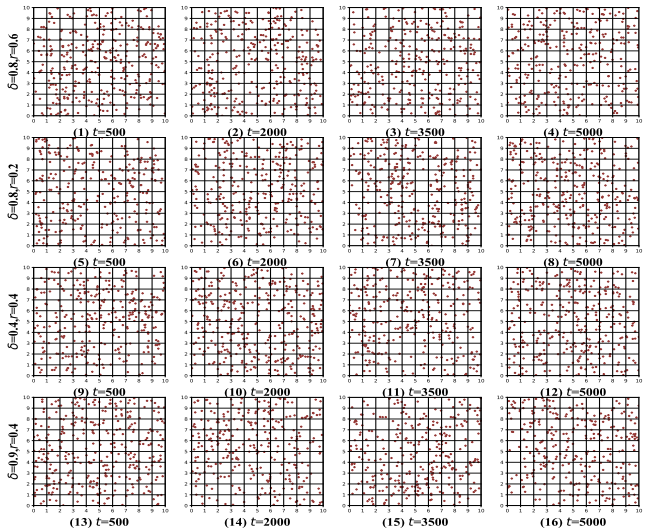
$$M' = \begin{pmatrix} 1 & -r' \\ r' & 0 \end{pmatrix}, \tag{10}$$

where $r' \in [0, 1]$ represents a tunable payoff parameter. Figs. 4(a) and 4(b) depict the evolutionary snapshots in the two-dimensional space and the corresponding network structures of SWOBD under different combinations of the exploitation rate δ and the payoff parameter r' in the SHG setting, respectively. The results show that higher values of δ and lower values of r' favor the rapid formation and sustained evolution of communities, consistent with the findings observed in the SWOBD within the snowdrift game scenario. These findings confirm that the snowdrift game employed in this study serves as a special case, and the proposed framework can be extended to other game scenarios by substituting the corresponding payoff matrix.

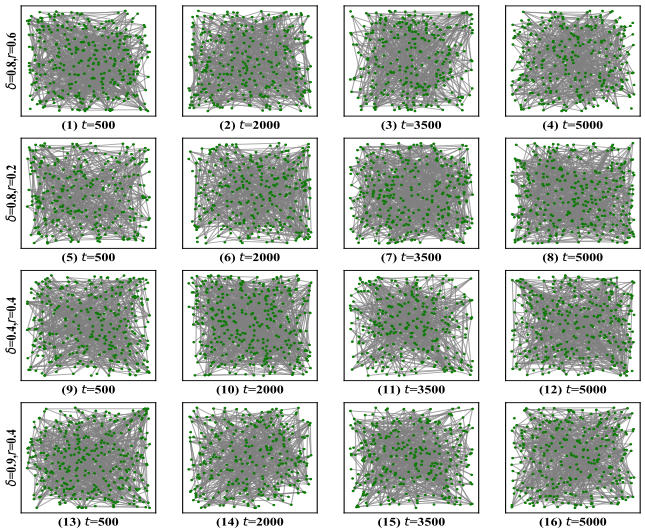
3.3 Further Results for Effect of Payoff Parameter and Exploitation Ratio on Community Structures

Although the results shown in subsection 3.2 indicate no community structure formation for four specific parameter combinations, this observation may not generalize across all parameter combinations. To further investigate the effect of the payoff parameter and the exploitation rate on community structures, we plot the heat maps of N_c (the sum of the number of individuals in the four locations with the highest populations throughout the evolutionary process) with respect to r and δ under three different death distributions, and the results are demonstrated in Fig. 5.

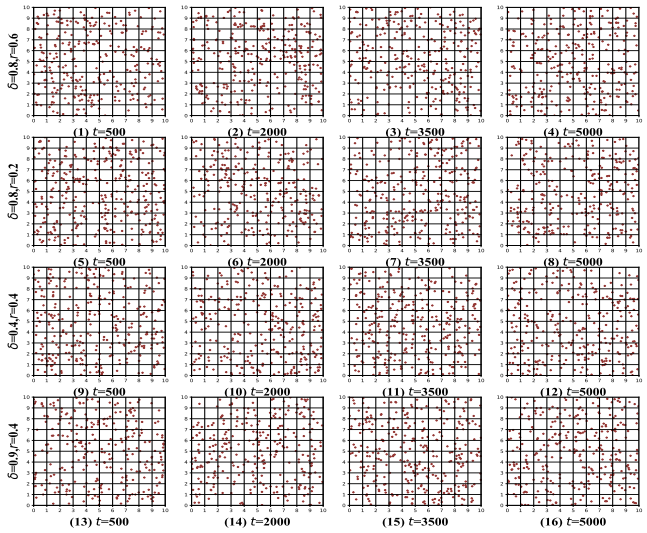
From Figs. 5(a) and 5(b), we can get that the value of N_c is relatively small when the death process obeys both uniform and exponential distributions, indicating a lack of community structure in these cases. Furthermore, the heat maps exhibit a highly uneven distribution of N_c, suggesting that community formation is not significantly influenced by the payoff parameter r or the exploitation rate δ. These findings align with the results obtained with the power-law distribution for the death process, as discussed in the main manuscript. In contrast to the results for other death distributions, the case where the death process follows a lognormal distribution shows a notable emergence of community structures. As illustrated in Fig. 5(c), when the exploitation rate δ is large and the payoff parameter r is small, N_c can reach approximately 100, indicating the formation of community structures within the system.



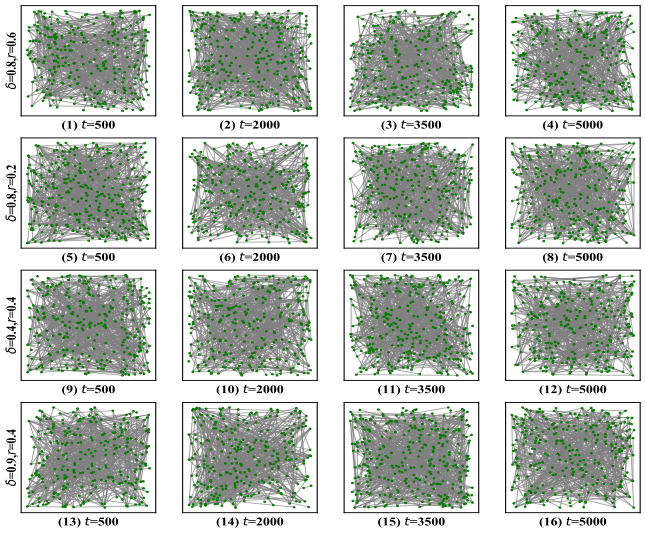
(a) Individual distribution under uniform distribution



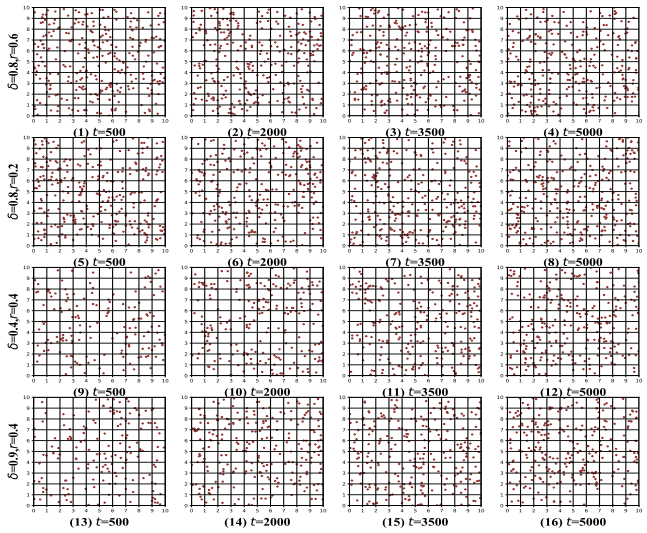
(b) Network structure under uniform distribution



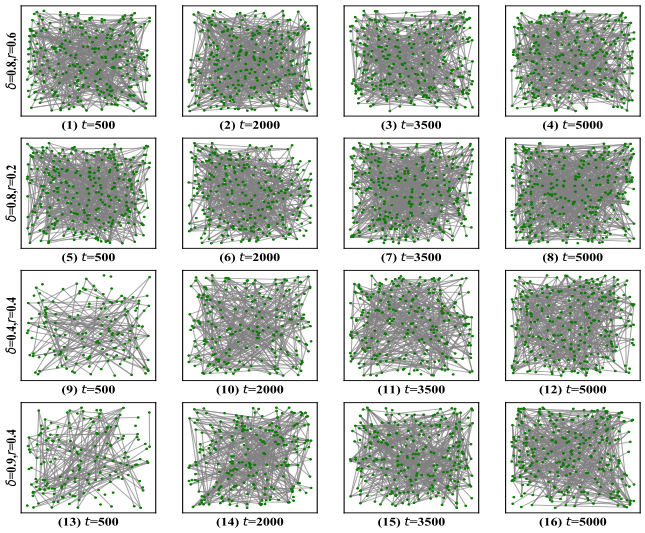
(c) Individual distribution under exponential distribution



(d) Network structure under exponential distribution

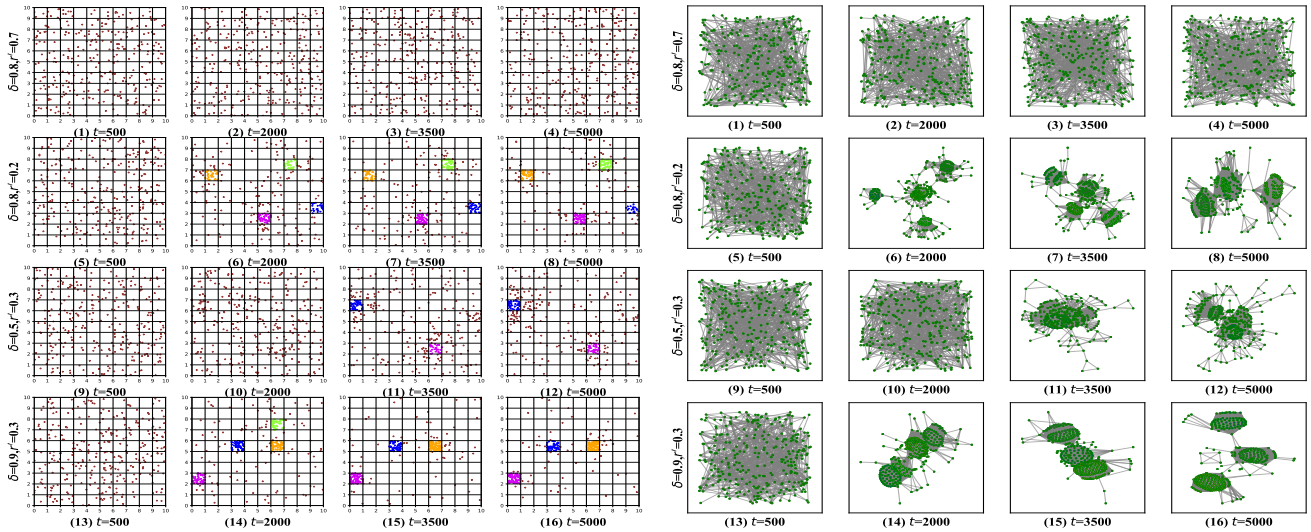


(e) Individual distribution under lognormal distribution



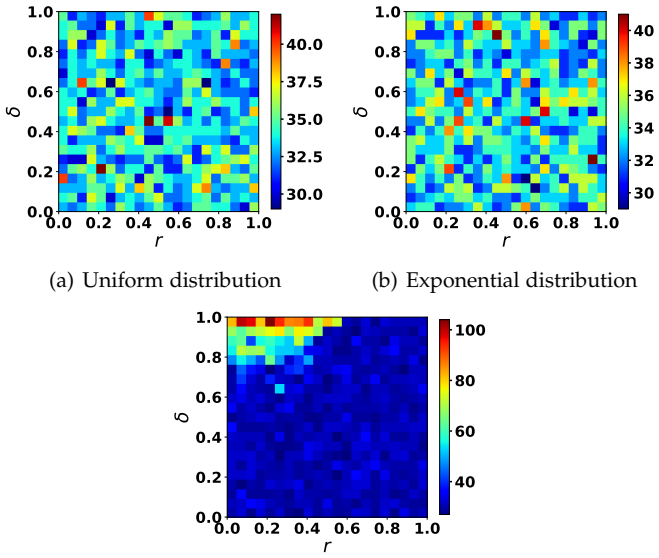
(f) Network structure under lognormal distribution

Fig. 3. **Evolutionary snapshots in the two-dimensional space and their corresponding network structures under different parameter pairs (δ, r) and death processes.** By setting the parameter pairs (δ, r) as (0.8, 0.6), (0.8, 0.2), (0.4, 0.4), and (0.9, 0.4) from top to bottom and fixing the time step at 500, 2000, 3500, and 5000 from left to right, we present the evolutionary snapshots of individual movement in the two-dimensional space on the system with the birth-death process, where the death process follows uniform (in panel (a)), exponential (in panel (c)), and lognormal (in panel (e)) distributions, respectively. The corresponding network structures are displayed in panels (b), (d), and (f). All other parameters in these subplots remain consistent with the main manuscript for comparison.



(a) Individual distribution of SWOBD within SHG (b) Network structure of SWOBD within SHG

Fig. 4. **Evolutionary snapshots in the two-dimensional space and corresponding network structures of SWOBD under different parameter pairs (δ, r') within the context of the stag hunt game.** Suplot (a) presents the evolutionary snapshots of individual movement in the two-dimensional space on the system without the birth-death process, under four different combinations of (δ, r'): (0.8, 0.7), (0.8, 0.2), (0.5, 0.3), and (0.9, 0.3), arranged from top to bottom. For each combination, the system state is shown at time steps of 500, 2000, 3500, and 5000, from left to right. Subplot (b) displays the corresponding network structures. All other parameters are kept consistent with those used in the main manuscript to facilitate direct comparison.



(a) Uniform distribution (b) Exponential distribution

(c) Lognormal distribution

Fig. 5. **Heat maps of N_c with respect to payoff parameter r and exploitation rate δ under various death processes.** By setting the y-axis as the exploitation rate δ with a range [0, 1] and the x-axis as the payoff parameter r with a range [0, 1], we demonstrate the heat maps of N_c about payoff parameter r and exploitation rate δ under the death process obeys uniform (in subplot (a)), exponential (in subplot (b)), and lognormal (in subplot (c)) distributions, respectively.

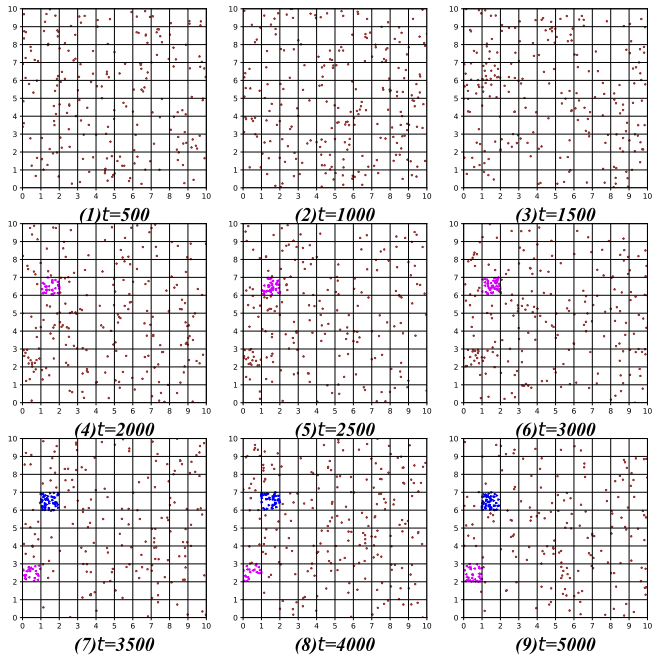


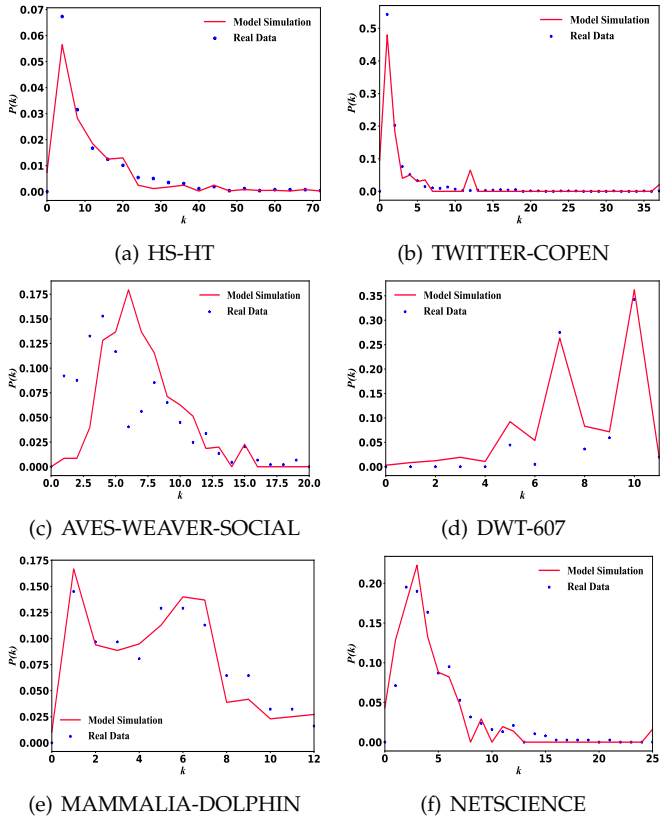
Fig. 6. **Evolutionary snapshots in the two-dimensional space for parameter pair (δ, r) = (1.0, 0.1) with a lognormal death process.** The figure depicts the individual distribution in the two-dimensional space over time, where the death process follows a lognormal distribution with the parameters $\delta = 1.0$ and $r = 0.1$. The snapshots reveal the emergence of distinct community structures as time progresses.

To provide a more intuitive understanding, Fig. 6 presents the distribution of individuals in the two-dimensional space over time, with the death process following a lognormal distribution and parameters set to $\delta = 1.0$ and $r = 0.1$, respectively. The snapshots reveal that a community structure begins to form around $t = 2,000$. Another community structure emerges around $t = 3,500$ and subsequently stabilizes, which is consistent with the results depicted in Fig. 5(c). However, it is important to note that although community structures are present in this scenario, the value of N_c remains smaller compared to systems without the birth-death process. Specifically, the number

and size of community structures in systems without the birth-death process are larger, as demonstrated in Figs. 6(b) and 7(b) of the main manuscript.

3.4 Further Results for Fit of the Degree Distribution using Reinforcement Learning on Real Networks

In this subsection, we evaluate the fit of the degree distribution on six different real networks, and the results are shown in Fig. 7.



(a) HS-HT (b) TWITTER-COPEN

(c) AVES-WEAVER-SOCIAL (d) DWT-607

(e) MAMMALIA-DOLPHIN (f) NETSCIENCE

Fig. 7. **Fit of the degree distribution on six different real networks.** The figure illustrates the degree distributions of networks generated by the SWOBD model, which are used to fit the degree distributions of the HS-HT network (panel (a)) and the TWITTER-COPEN network (panel (b)). Additionally, we fit the degree distributions of real networks using networks generated by the SWBD model under various death processes. Subplots (c) and (d) show the fitting results for the AVES-WEAVER-SOCIAL and DWT-607 networks, where the death processes follow power-law and uniform distributions. Subplots (e) and (f) present the fitting results for the MAMMALIA-DOLPHIN and NETSCIENCE networks, where the death processes follow exponential and lognormal distributions, respectively. The x-axis represents the degree of the network, while the y-axis indicates the probability of each degree.

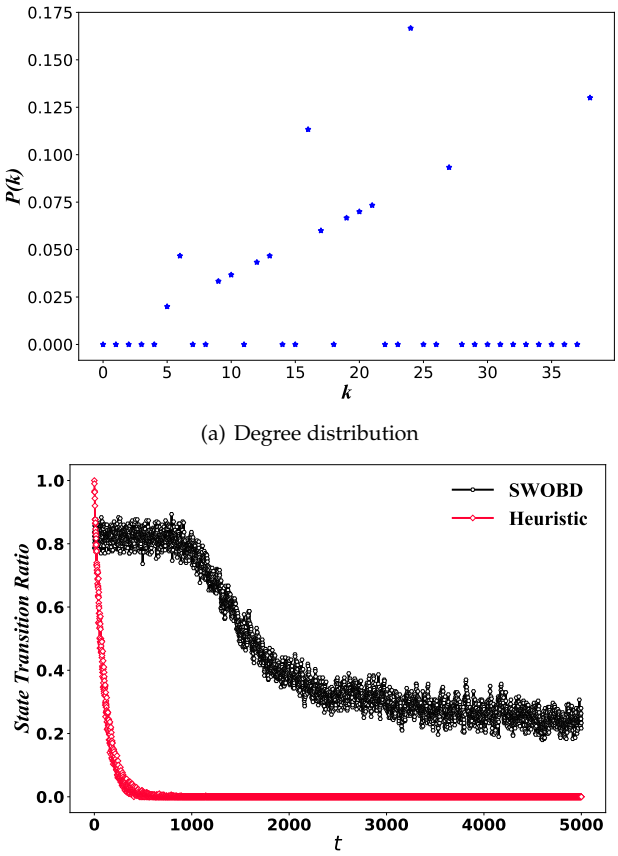
3.5 Further Results for Comparison with Heuristic Model

This subsection presents the performances of the heuristic model, and the results are illustrated in Fig. 8.

3.6 Further Results for Network Structure and Metrics

In this subsection, we study the network structure, clustering coefficient, degree distribution, and joint degree distribution of networks generated by the death process following uniform, exponential, and lognormal distributions, and the results are displayed in Fig. 9.

From Figs. 9(d)-(f), it is evident that most nodes in the networks generated by these three death processes exhibit



(a) Degree distribution

(b) State transition ratio

Fig. 8. **The performances of the heuristic model.** (a) is the degree distribution of the network generated by the heuristic model. (b) shows the evolutionary curves of the state transition ratio as time progresses for the SWOBD and heuristic models.

high clustering coefficients, with values of 0.701, 0.676, and 0.671 for the uniform, exponential, and lognormal distributions, respectively. This suggests a significant level of aggregation among nodes in all three networks. The degree distributions illustrated in Figs. 9(g)-(i) reveal that while most nodes have relatively small degrees, there are also a few nodes with relatively larger degrees. Notably, the degree distributions for the death process obeying uniform and lognormal distributions are similar, both approximately following a power-law distribution. By plotting the joint degree distributions of the three networks, we observe from Figs. 9(j)-(l) that the numbers around and on the diagonal line in the heat maps are substantial. This indicates that all three networks exhibit assortative mixing and high degree correlation, consistent with findings obtained in the main manuscript where the death process adheres to a power-law distribution. Furthermore, the assortativity coefficients for networks generated by the uniform, exponential, and lognormal death processes are 0.874, 0.755, and 0.871, respectively. These relatively high values corroborate the observations from the joint degree distributions and align with our previous analysis.

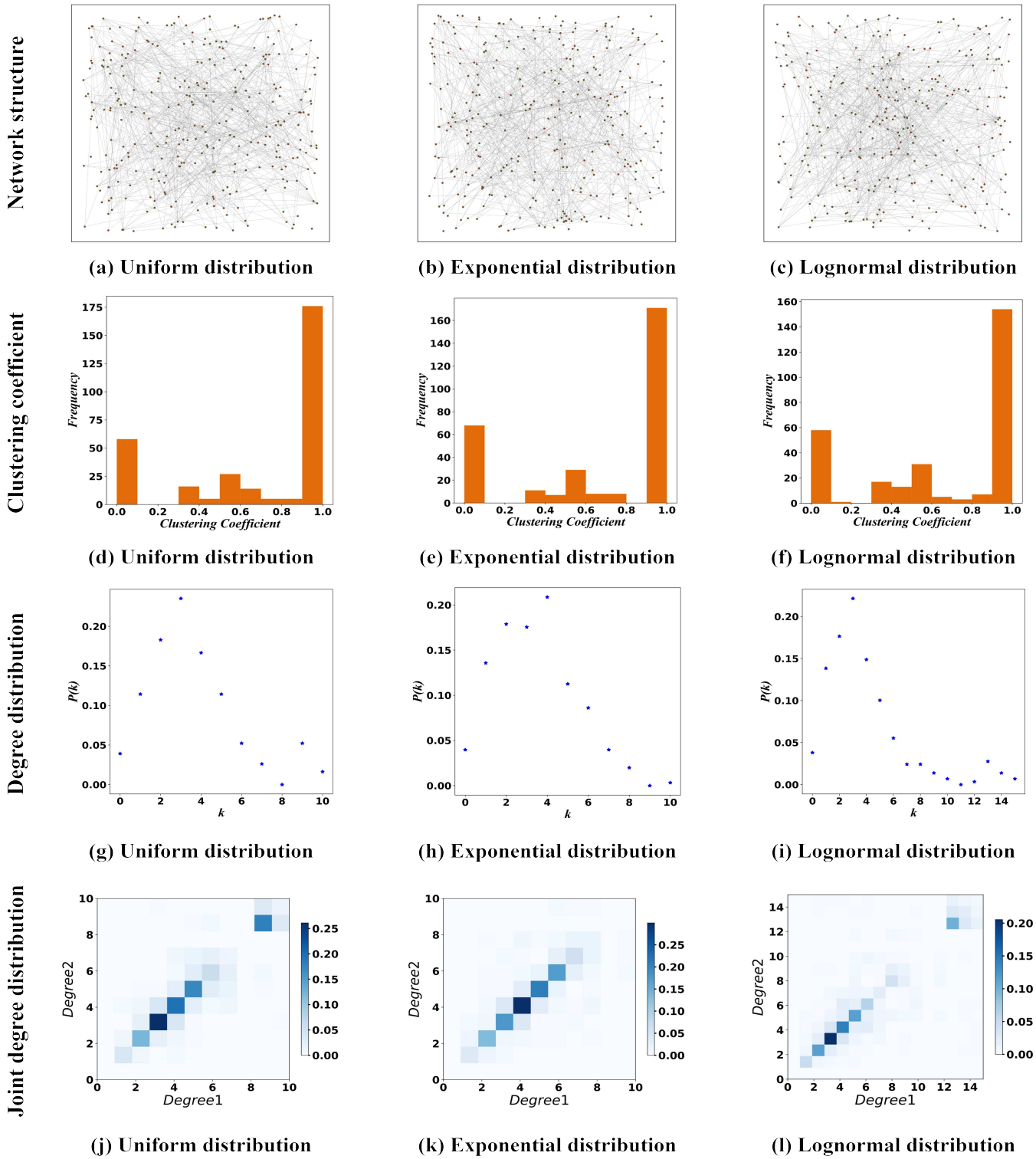


Fig. 9. **Some topological metrics of networks generated by SWBD with uniform, exponential, and lognormal distributions.** In this figure, we respectively demonstrate the networks and their corresponding topological metrics of SWBD with uniform (first column), exponential (second column), and lognormal (third column) distributions. The network structures are shown in subplots (a)-(c), clustering coefficients in subplots (d)-(f), degree distributions in subplots (g)-(i), and joint degree distributions in subplots (j)-(l).