Chaos, Solitons and Fractals 211 (2026) 118764

Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

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

Non-monotonic effects of interlayer coupling on cooperation dynamics in higher-order networks ☆

Rongrong Lin [a], Sihang Li [a], Haojie Xu [b], Matjaž Perc [c,d,e,f],*, Luoluo Jiang [g],*, Changwei Huang [a,h],*

[a] School of Computer, Electronics and Information, Guangxi University, Nanning, 530004, China
[b] School of Electronics and Information Engineering, Tiangong University, Tianjin, 300387, China
[c] Faculty of Natural Sciences and Mathematics, University of Maribor, Koroška cesta 160, Maribor, 2000, Slovenia
[d] Community Healthcare Center Dr. Adolf Drolc Maribor, Ulica talcev 9, Maribor, 2000, Slovenia
[e] Department of Physics, Kyung Hee University, Dongdaemun-gu Seoul, 02447, Republic of Korea
[f] University College, Korea University, Seongbuk-gu Seoul, 02841, Republic of Korea
[g] School of Information Engineering and Artificial Intelligence, Zhejiang University of Finance and Economics, Hangzhou, 310018, China
[h] Guangxi Key Laboratory of Multimedia Communications and Network Technology, Guangxi University, Nanning, 530004, China

ARTICLE INFO

Keywords:
Evolutionary game theory
Multilayer coupled networks
Higher-order interactions
Reinforcement learning
Cooperation evolution

ABSTRACT

Given the distinct advantages of higher-order networks over pairwise interaction networks in capturing the intricate interdependencies among individuals, the study of cooperation dynamics based on higher-order interactions has attracted increasing attention. However, little research has explored how the transition from pairwise to higher-order interactions influences the evolution of cooperation. In this study, we develop an evolutionary game model on a two-layer coupled network, wherein agents participate in both pairwise and higher-order interactions. To integrate strategies across layers, we introduce a novel self-regarding Q-learning algorithm. Our results reveal a significant non-monotonic effect of interlayer coupling strength on cooperation, which strongly depends on the intensity of the social dilemma. Moderate interlayer coupling promotes cooperation through payoff synergy under low dilemma intensity, whereas under high dilemma intensity, it inhibits cooperation due to conflicting payoff signals between layers. Micro-level analysis reveals that varying the coupling strength shifts agents' decision-making from pairwise to higher-order interactions, thereby altering evolutionary trajectories and collective outcomes. These findings remain robust across different network topologies, highlighting how environmental pressure and interlayer coupling jointly shape the emergence of cooperation in complex systems.

1. Introduction

Cooperation is a ubiquitous phenomenon in biological and social systems, playing a vital role in individual survival and population prosperity [1–3]. However, altruistic cooperation requires individuals to bear personal costs while benefiting others, which contradicts the instinct of maximizing self-regarding under the classical Darwinian evolutionary framework [4–6]. To explain the emergence of cooperation, evolutionary game theory provides a powerful framework for analyzing the conflicts between individual

☆ Given his role as Section Editor, Matjaz Perc had no involvement in the peer-review of this article and has no access to information regarding its peer-review. Full responsibility for the editorial process for this article was delegated to another journal editor.

* Corresponding authors.
E-mail addresses: matjaz.perc@gmail.com (M. Perc), jiangluoluo@zufe.edu.cn (L. Jiang), cwhuang@gxu.edu.cn (C. Huang).

https://doi.org/10.1016/j.chaos.2026.118764
Received 24 May 2026; Received in revised form 27 June 2026; Accepted 1 July 2026

and collective interests in social dilemmas [7]. Recent studies have further revealed that seemingly disadvantageous behaviors, such as disguise, can serve as a medium for the emergence and stabilization of cooperation [8]. However, cooperative behavior in the real world often transcends simple pairwise interactions, emerging more frequently within complex group interactions among multiple agents [9–11]. Building on this, higher-order network tools such as hypergraphs [12,13] and simplicial complexes [14,15] can effectively represent non-pairwise interactions among agents, thereby offering a novel perspective for the in-depth analysis of the evolutionary dynamics of collective cooperation.

In recent years, research based on complex networks has continuously expanded our understanding of the mechanisms behind the emergence of collective cooperation, with network structures [16,17] and strategy updating rules [18] serving as two core dimensions. Regarding network structure, research perspectives have expanded from single-layer networks to multilayer networks [19–21] and higher-order networks [22], frameworks that can more realistically describe the diverse interaction patterns in real-world systems [23,24]. Meanwhile, advances in network reliability assessment methods have provided new tools for understanding the robustness and dynamics of social networks [25]. For instance, studies show that in multilayer networks, interlayer coupling relationships can effectively enhance the cooperation of the system [26–28]; meanwhile, higher-order interaction network structures have been proven to maintain or even promote cooperation, and potentially trigger an explosive growth in cooperation frequencies [29], even in scenarios where pairwise interactions would otherwise lead to complete defection [30,31]. In parallel, research on strategy updating rules has also evolved from traditional evolutionary mechanisms like the replicator equation [32] and simple imitation [33] to more adaptive Q-learning algorithms for strategy updating [34,35]. Consequently, applying reinforcement learning to higher-order networks such as hypergraphs has become a current research hotspot; related studies have elucidated how agents achieve the emergence of collective cooperation in complex higher-order environments through dynamic learning processes [36]. This method provides a powerful computational analysis framework for revealing non-monotonic evolutionary phenomena, such as asymmetric responses of cooperation frequencies to varying dilemma intensities [33] and punishment mechanisms [37]. Therefore, combining the Q-learning strategy updating rule with more realistic multilayer or higher-order network structures has become a crucial trend for deeply uncovering the mechanisms of cooperation emergence under higher-order interactions.

While the aforementioned studies have greatly advanced our understanding of the evolution of cooperation, and evolutionary game and network approaches have been extended to a wide range of prosocial behaviors and practical governance scenarios, demonstrating strong generalizability across domains [38–40], two critical limitations remain to be addressed. First, prior research on multilayer network coupling predominantly focuses on systems composed of pairwise networks, exploring phenomena such as the incentive effects between regular and scale-free networks [41], the guiding role of neutral populations [42], and public cooperation under environmental asymmetry [43]. These studies generally lack a systematic examination of scenarios where pairwise and higher-order interactions coexist [44]. Second, existing models involving higher-order interactions universally assume that agents adopt identical strategies across interactions of different scales. For example, whether by adjusting the proportion of motifs to analyze its impact on cooperation frequency [45] or by introducing parameters to adjust the payoff weights of different-order interactions [15], most of these models ignore the complexity of context-dependent strategy adjustment. In reality, behavioral economics suggests that individuals flexibly adjust their decisions based on interaction structures and group sizes. Prospect theory, for instance, posits that decision-making is driven by changes in gains and losses relative to a reference point, exhibiting distinct loss aversion [46,47]. In a two-player game, an agent's reference point is usually a single opponent; in a multi-player game, however, it shifts to a comparison with the group average or the highest earner. This shift can fundamentally alter value judgments and risk preferences—an individual willing to take risks for higher payoffs in a two-player game might become conservative in a multi-player setting to avoid severe losses. Yet, many higher-order interaction models overlook this context-dependent strategy adjustment.

To more accurately depict the mechanisms driving cooperative behavior in social dilemmas, it is imperative to develop a new model that integrates pairwise and higher-order network coupling structures while allowing agents to adopt differentiated strategies based on interaction scales. Therefore, this study proposes a two-layer coupled network model that integrates pairwise interactions and higher-order group interactions, combined with a self-regarding Q-learning algorithm. In this model, agents are placed on a coupled two-layer square lattice network: the upper layer involves a two-player snowdrift game, while the lower layer features a three-player snowdrift game. These two layers are coupled via a one-to-one correspondence of agents, meaning their evolution is jointly driven by cumulative payoffs from both layers. This design overcomes the limiting assumption of adopting a single strategy in traditional higher-order models, allowing agents to adapt their joint strategies based on the combined payoff from both layers. For strategy updating, each agent employs a self-regarding Q-learning algorithm to learn and optimize their strategy combinations across both layers through a shared Q-table. This approach not only captures the agents' ability to accumulate experience through trial and error but also reflects their trade-off process when integrating payoffs from cross-scale interactions.

The main contributions of this study are as follows:

- Constructing a network game framework that couples interactions of pairwise and high-orders. The proposed framework expands upon existing research limited to coupling networks of identical interaction orders, including purely pairwise or purely high-order structures. This advancement enables a more realistic simulation of complex systems where multiple interaction types coexist in real societies.
- Introducing an updating mechanism that allows agents to make differentiated decisions based on the interaction scale. This circumvents the single strategy assumption conventionally found in traditional models of high-order interactions.
- Revealing the nonlinear regulatory effect of interlayer coupling on cooperative behavior, demonstrating a dual effect of promoting or inhibiting cooperation under different dilemma intensities. This mechanism exhibits strong robustness across both regular and heterogeneous network topologies.

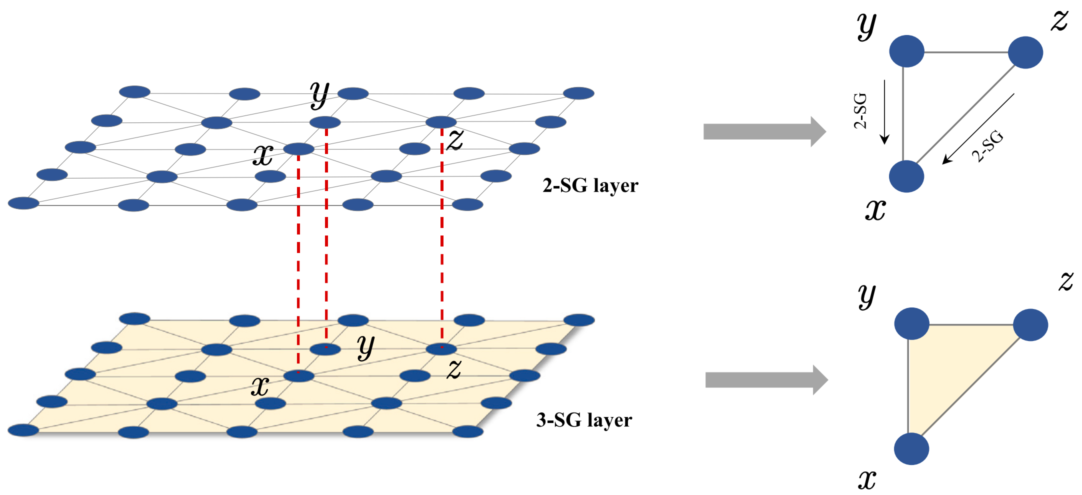

Fig. 1. Schematic diagram of the two-layer coupled network model. The model consists of two structurally identical square lattice network layers with periodic boundary conditions (the upper layer is the 2-SG layer, and the lower layer is the 3-SG layer). The nodes in each layer are occupied by agents, and intra-layer connections are established via the Moore neighborhood. Each agent simultaneously participates in the two-player snowdrift game in the upper layer and the three-player snowdrift game in the lower layer. The two layers are coupled through a one-to-one positional correspondence, allowing the game payoffs from both layers to jointly influence the agent's decision-making.

The structure of this paper is organized as follows: Section 2 introduces the model construction, including the two-layer network structure, the payoff calculation for the multi-player snowdrift game, and the specific design of the Q-learning algorithm. Section 3 presents and analyzes the numerical simulation results, exploring the joint impact of interlayer coupling and social dilemmas on the evolution of cooperation alongside its microscopic mechanisms. Finally, Section 4 summarizes the paper and provides a conclusion.

2. Model

This study constructs a two-layer coupled network as the interactive environment for the agents, as illustrated in Fig. 1. Both layers possess identical topological structures, utilizing a square lattice with periodic boundary conditions. The lattice size L is set to 50, meaning each layer contains a total of $N = L \times L$ nodes. Each node is occupied by an agent, and agents interact within their respective layers using a Moore neighborhood (comprising 4 orthogonal and 4 diagonal directions, totaling 8 nearest neighbors).

Each agent is simultaneously embedded within this two-layer structure. Specifically, in 2-SG layer, agents engage in a two-player snowdrift game with neighbors in their Moore neighborhood. In 3-SG layer, they participate in a three-player snowdrift game with their Moore neighbors. The two layers are interconnected through a one-to-one positional coupling, ensuring that the game payoffs of an agent in both layers influence each other and jointly constitute the overall payoff feedback.

The Multi-player Snowdrift Game describes a classic M-player social dilemma scenario, where M denotes the interaction group size. In this game, a group of M agents collectively faces a snowdrift (obstacle) that needs to be cleared. The total cost of clearing the snowdrift is c. Once the path is cleared, each agent receives a benefit b, regardless of whether they actively participated in the snow-clearing effort.

Agents have two strategy options: to become a cooperator, who participates in clearing the snow and shares the cost, or to become a defector, who does not participate. The payoff rules of the snowdrift game are defined as follows: if the number of cooperators in the group is $n_C \geq 1$, the total snow-clearing cost c is equally divided among these cooperators. The cost borne by each cooperator is c/n_C, yielding a net payoff of $b - c/n_C$. Defectors, on the other hand, do not incur any cost but equally enjoy the benefit b of the cleared path, resulting in a net payoff of b. If no one in the group clears the snow (i.e., $n_C = 0$), the snowdrift remains uncleared, and all agents fail to obtain any benefit, yielding a payoff of 0 for everyone. Based on these rules, the payoffs of the agents are calculated as

$$\bar{\pi}_C\left(n_C\right) = b - \frac{c}{n_C} \quad \text{for} \quad n_C \in [1, M],$$
$$\bar{\pi}_D\left(n_C\right) = \begin{cases} 0 & \text{for} \quad n_C = 0 \\ b & \text{for} \quad n_C \in [1, M-1]. \end{cases} \tag{1}$$

In the analysis of the multi-player snowdrift game, to simplify the model and highlight the critical relationship between cost and benefit, we introduce the cost-to-benefit ratio $r = c/b$ to normalize the payoffs. Through this normalization, the basic parameters of the game are reduced to a single variable r, where $r \in [0,1]$. The normalized payoffs of the agents are given as

$$\begin{aligned}
&\pi_C\left(n_C\right) = 1 - \frac{r}{n_C} \quad \text{for} \quad n_C \in [1, M], \\
&\pi_D\left(n_C\right) = \begin{cases} 0 & \text{for} \quad n_C = 0 \\ 1 & \text{for} \quad n_C \in [1, M-1]. \end{cases}
\end{aligned} \tag{2}$$

In the first layer (2-SG), an agent plays the two-player snowdrift game with all its neighbors, and its normalized payoff is calculated as

$$\pi_{x,y} = \begin{cases} 1 - a_x \cdot \frac{r}{n_C^{/}} & \text{if} \quad n_C^{/} > 0 \\ 0 & \text{if} \quad n_C^{/} = 0 \end{cases}, \tag{3}$$

where $r \in [0,1]$ is used to quantitatively characterize the dilemma intensity of the snowdrift game, with a larger r indicating a stronger dilemma; $\pi_{x,y}$ denotes the payoff obtained by agent x along the link formed with its neighbor y; a_x represents the strategy of agent x, where $a_x = 1$ indicates cooperation and $a_x = 0$ indicates defection; and $n_C^{/} \in \{0,1,2\}$ denotes the number of cooperators among the two agents (agent x and its neighbor y) on the link.

The payoff obtained by an agent in the two-player snowdrift game layer is calculated as

$$\pi_x^{/} = \sum_{y=1}^{k_x^{/}} \pi_{x,y}, \tag{4}$$

where $k_x^{/}$ denotes the number of two-player games in which agent x participates (in a regular Moore neighborhood, $k_x^{/} = 8$), and $\pi_x^{/}$ represents the cumulative payoff obtained by agent x from all links formed with its neighbors.

When calculating the payoffs of the agents, we use the average payoff for comparison. The average payoff per game for an agent in the two-player snowdrift game layer is calculated as

$$\bar{\pi}_x^{/} = \frac{\pi_x^{/}}{k_x^{/}}. \tag{5}$$

In the second layer, agent x plays the three-player snowdrift game with all its neighbors. For the three-player snowdrift game (i.e., $M = 3$), we adopt the standard payoff model for the multi-player snowdrift game:

$$\pi_{x,h} = \begin{cases} 1 - a_x \cdot \frac{r}{n_C^{\triangle}} & \text{if} \quad n_C^{\triangle} > 0 \\ 0 & \text{if} \quad n_C^{\triangle} = 0 \end{cases}, \tag{6}$$

where $\pi_{x,h}$ denotes the payoff obtained by agent x on the hth triangle; a_x represents the strategy of agent x, where $a_x = 1$ indicates cooperation and $a_x = 0$ indicates defection; and $n_C^{\triangle}$ denotes the number of cooperators in the hth triangle in which agent x participates.

The payoff obtained by agent x in the three-player snowdrift game layer is calculated as the sum of its payoffs across all the triangular groups it participates in, as shown in Eq. (7). Here, $k_x^{\triangle}$ denotes the number of three-player games in which agent x participates.

$$\pi_x^{\triangle} = \sum_{h=1}^{k_x^{\triangle}} \pi_{x,h}. \tag{7}$$

In this model, the three-player games are conducted within specific triangular groups formed by a central agent and its Moore neighborhood. Specifically, the eight neighbors of agent x are labeled as y_j in clockwise order, where the index $j \in \{0,1,2,\ldots,7\}$. Agent x consistently participates in eight such triangular groups, defined as:

$$T_j = (x, y_j, y_{(j+1) \bmod 8}) \qquad \text{if} \quad j = 0,1,\ldots,7. \tag{8}$$

Each group T_j is defined by a triplet consisting of the central agent and two of its adjacent neighbors. The use of modulo 8 for the index $j+1$ ensures the closure of the interaction ring. Consequently, for every agent on the square lattice network, the number of overlapping triangular interaction groups is a fixed value, $k_x^{\triangle} = 8$.

Similarly, we calculate the average payoff. The average payoff per game for an agent in the three-player snowdrift game layer is calculated as

$$\bar{\pi}_x^{\triangle} = \frac{\pi_x^{\triangle}}{k_x^{\triangle}}. \tag{9}$$

In our research, we consider the comprehensive payoff f_x of an agent in the two-layer network. The comprehensive payoff of each agent, calculated as a linear combination of the payoffs from the two game layers, is given by:

$$f_x = (1-\alpha)\,\bar{\pi}_x^{/} + \alpha \bar{\pi}_x^{\triangle}, \tag{10}$$

where $\alpha \in [0,1]$ represents the coupling strength, characterizing the weight of the payoffs obtained from the two layers. An increase in the value of α leads to a higher proportion of the payoff coming from the three-player snowdrift game layer, while a decrease in α results in a higher proportion from the two-player snowdrift game layer. Specifically, $\alpha = 0$ indicates that the comprehensive payoff f_x of the agent depends entirely on its average payoff in the 2-SG layer, and $\alpha = 1$ indicates that f_x depends entirely on its average payoff in the 3-SG layer.

Furthermore, we employ the self-regarding Q-learning algorithm for the strategy updating of agents. Unlike conventional self-regarding Q-learning, in the two-layer network of this study, each agent is defined as a unified decision-making entity. It governs its corresponding behaviors in both the 2-SG and 3-SG layers and conducts learning and decision-making based on a shared Q-table.

To this end, we define a finite state space S and an action space A, such that $S = A = \{\mathrm{DD}, \mathrm{DC}, \mathrm{CD}, \mathrm{CC}\}$. Here, the state $s_t \in S$ represents the joint strategy adopted by the agent in the previous round (round $t-1$) across both game layers, which encodes its historical experience. The action $a_t \in A$, on the other hand, represents the new joint strategy the agent plans to execute in both layers in the current round (round t). The specific definitions of each composite symbol are as follows: DD indicates defection in both layers; DC indicates defection in the 2-SG layer and cooperation in the 3-SG layer; CD indicates cooperation in the 2-SG layer and defection in the 3-SG layer; CC indicates cooperation in both layers.

The Q-values for all possible state–action pairs are stored in the agent's shared Q-table, where the row index corresponds to the state and the column index corresponds to the action. The Q-table of agent x at time step t can be expressed as the following 4×4 matrix:

$$Q^x_{s,a}(t) = \begin{bmatrix} Q^x_{\mathrm{DD,DD}}(t) & Q^x_{\mathrm{DD,DC}}(t) & Q^x_{\mathrm{DD,CD}}(t) & Q^x_{\mathrm{DD,CC}}(t) \\ Q^x_{\mathrm{DC,DD}}(t) & Q^x_{\mathrm{DC,DC}}(t) & Q^x_{\mathrm{DC,CD}}(t) & Q^x_{\mathrm{DC,CC}}(t) \\ Q^x_{\mathrm{CD,DD}}(t) & Q^x_{\mathrm{CD,DC}}(t) & Q^x_{\mathrm{CD,CD}}(t) & Q^x_{\mathrm{CD,CC}}(t) \\ Q^x_{\mathrm{CC,DD}}(t) & Q^x_{\mathrm{CC,DC}}(t) & Q^x_{\mathrm{CC,CD}}(t) & Q^x_{\mathrm{CC,CC}}(t) \end{bmatrix}, \tag{11}$$

where $Q^x_{s,a}(t)$ denotes the Q-value corresponding to the selection of action a when agent x is in state s at time step t. Initially, each agent is randomly assigned a state from the state space, and the Q-tables of all agents are initialized to 0. At round t, the agent selects an action based on its current state s_t, utilizing the ε-greedy algorithm. Specifically, with probability $\varepsilon \in (0,1]$, the agent randomly chooses an action a_t from the action space; with probability $1-\varepsilon$, it selects the action a_t associated with the highest Q-value for the current state. The agent participates in the game by executing action a_t and obtains its comprehensive payoff $f_x(t)$ calculated based on Eq. (10), which serves as the immediate reward. In this model, the action a_t executed by the agent directly becomes the state s_{t+1} it observes in the subsequent round. Subsequently, in accordance with the Q-learning framework, the Q-value in the shared Q-table is updated via the following Bellman equation, utilizing the current state s_t, the executed action a_t, the immediate reward $f_x(t)$, and the next state s_{t+1}:

$$Q^x_{s,a}(t+1) = (1-\delta)Q^x_{s,a}(t) + \delta\left[f_x(t) + \gamma \max_{a' \in A} Q^x_{s_{t+1},a'}(t)\right], \tag{12}$$

where $\delta \in (0,1]$ is the learning rate, which controls the speed at which new information overrides old Q-values; $\gamma \in [0,1)$ is the discount factor, determining the importance of future rewards; and $\max_{a' \in A} Q^x_{s_{t+1},a'}(t)$ represents the maximum estimated Q-value among all possible actions in the next state s_{t+1}. In this study, we set the parameters to $\delta = 0.5$, $\gamma = 0.9$, and $\varepsilon = 0.1$.

The simulation consists of T Monte Carlo steps, where each full MC step comprises N sub-steps. During each sub-step, an agent is randomly selected to perform action selection, state transition, and Q-table updating, ensuring that, on average, every agent receives one opportunity to update its strategy per full MC step. Once the system undergoes a sufficiently long simulation ($T = 3 \times 10^5$) and reaches a statistically stationary state, we observe the strategy distribution on the two-layer network over the final $10,000$ steps and calculate the average cooperation frequency of the system to analyze its stable behaviors. The detailed architecture of the entire learning and evolutionary process is outlined in Algorithm 1. Finally, the ultimate results are obtained by averaging over 10 independent runs.

3. Simulation results and analysis

In this section, we systematically investigate the joint effects of the interlayer coupling strength and the social dilemma on the evolution of cooperation through numerical simulations. We first reveal the overall patterns of the macroscopic cooperation frequency as the parameters vary. Subsequently, we elucidate the underlying microscopic mechanisms by analyzing the dynamic evolution and spatial distribution of interlayer strategy combinations. Finally, we verify the universality of our conclusions across different network topologies.

3.1. Non-monotonic regulation of the cooperation frequency by interlayer coupling

In our model, the payoffs obtained by agents in the two network layers (the 2-SG and 3-SG layers) interact through a coupling mechanism. We introduce the parameter of interlayer coupling strength, α, to quantitatively characterize the weight of this interaction. Adjusting α directly alters the comprehensive payoff of the agents, thereby regulating their cooperative behaviors in the two-layer environment. To this end, we first investigate how the frequency of cooperators, f_C, varies with the interlayer coupling strength α and the dilemma intensity r.

As shown in Fig. 2(a), the variation of the frequency of cooperators f_C with the interlayer coupling strength α exhibits a pronounced non-monotonicity, and this phenomenon strongly depends on the dilemma intensity r. Overall, the declining trend of f_C with increasing r remains consistent across all values of α, indicating that the suppressive effect of the dilemma on cooperation is

Algorithm 1 Evolutionary cooperation dynamics based on Q-learning in two-layer networks

Input: Total number of agents N, State set S, Action set A, Total Monte Carlo steps T.
Output: Frequency of cooperators in 2-SG and 3-SG layers.

Build a two-layer square lattice network with periodic boundary conditions;
Initialize Q-tables of all agents to 0;
for each agent $x \in \{1, 2, \dots, N\}$ **do**
  Randomly initialize the state s of x from S;
end for
for each Monte Carlo step $t \in [1, T]$ **do**
  for each round $i \in [1, N]$ **do**
    Randomly select an agent x;
    Generate a random number $prob$ between 0 and 1;
    if $prob < \epsilon$ **then**
      Randomly choose an action a from A;
    else
      Choose the action a with the highest Q-value in current state s;
      If multiple actions have the same maximum Q-value, select one randomly;
    end if
    Calculate x's average payoffs in 2-SG and 3-SG layers;
    Calculate x's comprehensive payoff according to Eq. (10);
    Update x's Q-table according to Eq. (12);
    Update x's next state s' to current action a;
    Transfer x's state from s to s';
  end for
end for

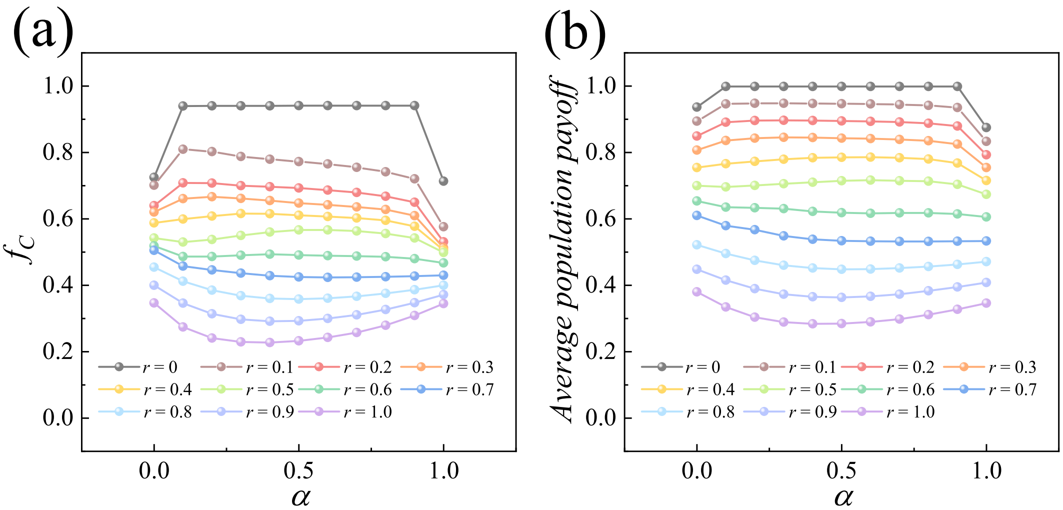

Fig. 2. The frequency of cooperators f_C (a) and the average population payoff (b) as functions of the coupling strength α under different dilemma intensities r. The average population payoff is computed as $\frac{1}{N}\sum_{x=1}^{N}\frac{\bar{\pi}_x^{/}+\bar{\pi}_x^{\triangle}}{2}$ during the steady state.

not altered by the interlayer coupling. However, the effect of the coupling strength α displays significant interval characteristics. For lower dilemma intensities, f_C initially increases with α, peaks at $\alpha \approx 0.1$, and subsequently decreases. This suggests that moderate interlayer payoff coupling optimally promotes cooperation. Conversely, under higher dilemma intensities, f_C initially decreases with α, reaches a minimum at $\alpha \approx 0.4$, and subsequently increases. This suggests that intermediate interlayer payoff coupling most strongly suppresses cooperation in this scenario. The curves for intermediate r are more complex, exhibiting behaviors that lie between the two limiting cases described above despite some fluctuations in their detailed shapes. As shown in Fig. 2(b), the average population payoff (social welfare) exhibits an overall trend consistent with the cooperation frequency, yet with notable local discrepancies. Under low dilemma intensity, the decline in welfare lags significantly behind the drop in cooperation frequency. This pattern is consistent with findings from previous studies, which have demonstrated that heterogeneous game structures and strategy learning costs can lead to a divergence between cooperation level and social welfare [48,49].

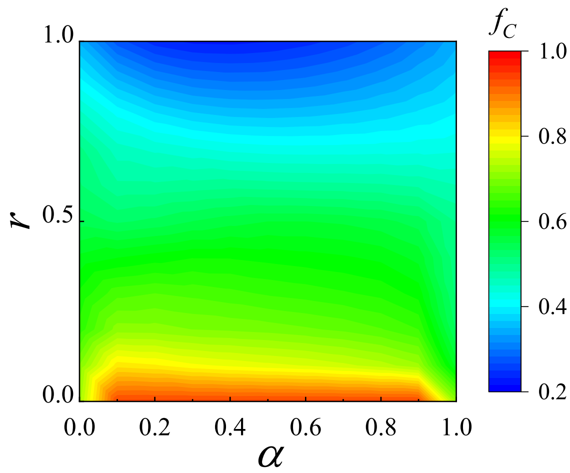

Fig. 3. Phase diagram of the cooperation frequency f_C in the $(\alpha - r)$ parameter space.

This non-monotonic phenomenon stems from the fact that changes in the interlayer coupling strength α alter the payoff driving the strategy updates of agents. When α is small, the agents' decisions are primarily driven by the payoff from the two-player snowdrift game (2-SG), a game structure that favors the formation and stability of reciprocal cooperation. When α is large, decision-making is instead dominated by the three-player snowdrift game (3-SG) payoffs, which entail a stronger "free-riding" temptation for defectors and thus more strongly erode cooperative clusters. Therefore, the variation of α characterizes the shift in the payoff structure between heterogeneous game layers. More precisely, within self-regarding Q-learning, α serves as a weighting factor that balances the contributions of pairwise (2-SG) and higher-order (3-SG) payoffs. By integrating these two sources into a composite decision signal, α shapes the agents' expectations of strategy values. Under low dilemma intensities (small r), cooperation is the dominant strategy in each individual layer. At intermediate α, exploration enables agents to identify that mutual cooperation across both layers (CC) yields combined payoffs that exceed those from any other joint action. This interlayer payoff coupling amplifies the incentive to cooperate, pushing f_C to a maximum. Conversely, under high dilemma intensities (large r), defection yields a higher immediate payoff in each single-layer game. At intermediate α, agents receive conflicting signals from the two layers: one layer favors defection, while the other, due to the collective dilemma, favors cooperation. This conflict in payoffs hinders the convergence of Q-value updates, thereby suppressing the stability of cooperative clusters at the macroscopic level. Thus, the coupling strength α modulates cooperative behavior by adjusting the weight assigned to payoffs from each layer, thereby governing the relative dominance of the interlayer payoff coupling and the signal conflict mechanism within the evolutionary game dynamics. These results reveal that the effect of interlayer coupling strength on cooperation is not fixed but shifts with dilemma intensity. Under weak dilemmas, moderate interlayer payoff coupling promotes cooperation; under strong dilemmas, the same coupling intensity instead suppresses it most strongly.

The joint effects of interlayer coupling strength α and dilemma intensity r on cooperation frequency are further illustrated in Fig. 3. A clear gradient in cooperation frequency is observed across the parameter space. As r increases, the system transitions gradually from robust cooperation to widespread defection. This pattern confirms the findings from Fig. 2(a). At low dilemma intensity (small r), moderate coupling ($\alpha \approx 0.1$) most effectively promotes cooperation, visible as a warm-colored high-value band extending along the α-axis. At high dilemma intensity (large r), intermediate coupling ($\alpha \approx 0.4$) most strongly suppresses cooperation, forming a cool-colored trough. Notably, near $r = 0.5$, a broad green transition band with a gentle gradient appears, indicating that cooperation frequency is insensitive to α in this region. Here, the system evolves smoothly, with no sharp phase transition boundaries.

3.2. Time evolution of microscopic strategy switching

To elucidate the dynamics underlying the macroscopic patterns, we further analyze the time evolution of cooperation frequency and agents' strategy combinations. Fig. 4 shows the evolution of f_C over Monte Carlo steps T for different dilemma intensities r. From the dynamic perspective of reinforcement learning, this evolutionary process essentially reflects the dynamic balancing process of agents weighing the temptation of short-term defection against long-term accumulated payoffs.

The evolution under a low dilemma intensity ($r = 0.1$) is shown in Fig. 4(a). When the interlayer coupling is very weak, the comprehensive payoff of the agent is primarily determined by the payoff from 2-SG layer. At low dilemma intensity, cooperation can yield high payoffs. Agents rapidly reinforce the Q-value of the cooperative strategy through Q-learning, forming a positive feedback loop; therefore, f_C exhibits a monotonically increasing trend and stabilizes at a relatively high level. Conversely, under strong coupling, the agent's comprehensive payoff is mainly determined by the 3-SG layer, where defectors often obtain higher immediate rewards than cooperators. Q-learning continuously strengthens the Q-values of actions that yield high immediate payoffs. In the early evolutionary stage, defection arising from random exploration in the 3-SG layer often gains higher payoffs due to free-riding. Through Q-learning updates, such payoffs increase the Q-value of the corresponding defection action in the shared Q-table.

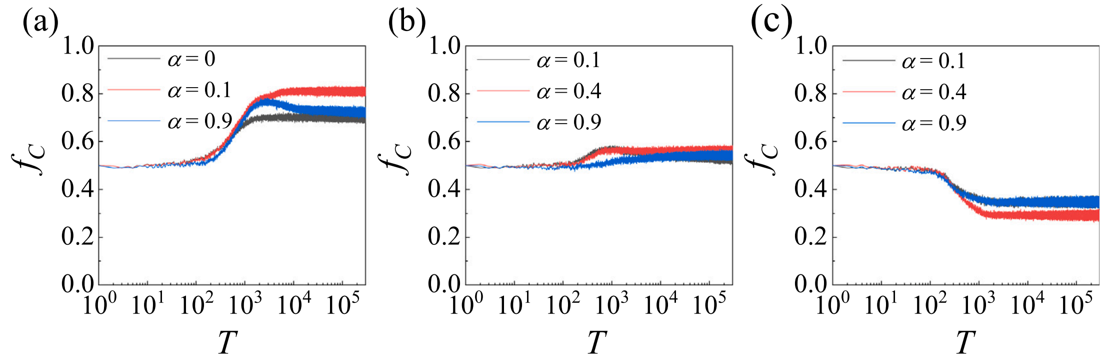

Fig. 4. Time evolution of the cooperation frequency f_C over time T. Panels (a), (b), and (c) correspond to $r = 0.1$, $r = 0.5$, and $r = 0.9$.

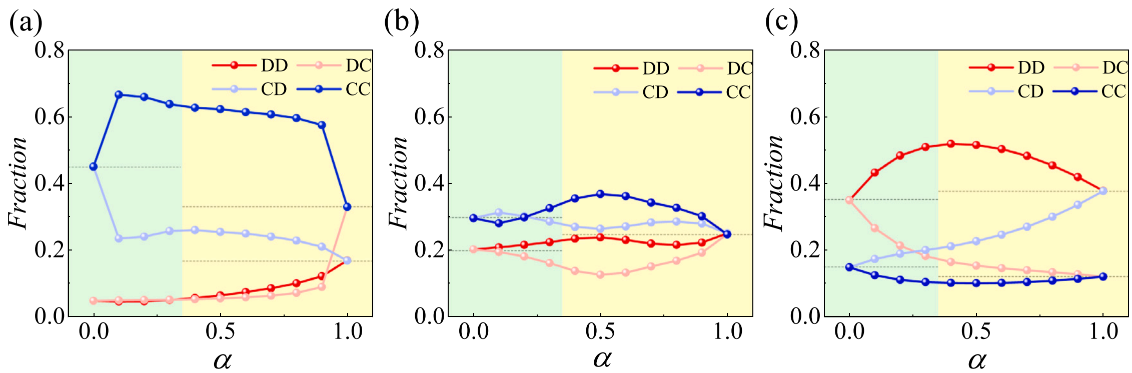

Fig. 5. Evolution of agent strategy combinations (DD: defection in both layers; DC: defection in 2-SG and cooperation in 3-SG; CD: cooperation in 2-SG and defection in 3-SG; CC: cooperation in both layers) as a function of interlayer coupling intensity. Panels (a), (b), and (c) correspond to $r = 0.1$, $r = 0.5$, and $r = 0.9$.

Because both layers share the same decision system, the strategy preference formed in the 3-SG layer may spill over to the 2-SG layer, influencing action selection there. This payoff-driven evaluation bias can be amplified through positive feedback under strong coupling, causing defection to spread within the system and potentially undermining the initial cooperation. As a result, after a brief initial rise, f_C peaks and then declines, eventually converging to a steady state around 0.7.

The evolution under a medium dilemma intensity ($r = 0.5$) presents distinct dynamic characteristics. As depicted in Fig. 4(b), regardless of the value of α, the cooperation frequency f_C eventually stabilizes at a moderate level, but the evolutionary speeds vary significantly across different interlayer couplings. Under weak coupling, the simple reciprocal structure of the 2-SG layer allows agents to quickly recognize the value of cooperation through Q-learning, causing f_C to rapidly reach a steady state. In contrast, under strong coupling, agents must integrate payoff information from both layers. This dual-layer feedback complicates their decision-making and slows down the convergence of Q-learning, resulting in a gradual increase in f_C. This further confirms that under medium dilemma conditions, the dilemma intensity r determines the final steady-state level of cooperation, while the interlayer coupling strength α governs the evolutionary path and the convergence speed of the system.

When the dilemma intensity is high ($r = 0.9$), the evolutionary behavior of the system changes. As shown in Fig. 4(c), for all values of α, the cooperation frequency f_C exhibits a rapid and monotonically decreasing trend. This indicates that when the temptation to defect is extremely high, the dilemma intensity becomes the dominant selection pressure, offsetting any potential positive effects from interlayer payoff coupling. Under such high dilemma intensity, the defection strategy consistently maintains a significant expected advantage during Q-table updates. However, the steady-state cooperation level varies distinctly with the coupling strength: $\alpha = 0.1$, $\alpha = 0.4$, and $\alpha = 0.9$ lead to final values of approximately 0.28, 0.22, and 0.30, respectively. Notably, $\alpha = 0.4$ yields the lowest cooperation frequency, which is consistent with the non-monotonic effect observed in Fig. 2(a), where intermediate coupling most strongly suppresses cooperation under high dilemma intensity.

The non-monotonic behavior of macroscopic cooperation originates from the evolution of microscopic strategy combinations. To examine how interlayer coupling strength α influences agents' joint strategies in the two-layer network, we analyze the fractions of the four strategy combinations (DD, DC, CD, CC) under different dilemma intensities, as shown in Fig. 5. The results reveal that the system's cooperation frequency f_C closely tracks the evolutionary patterns of these joint strategies, illustrating how α nonlinearly reshapes agents' value expectations. Specifically, under low dilemma intensity ($r = 0.1$), the peak in f_C coincides with the maximum fraction of the all-cooperation combination CC at $\alpha = 0.1$, indicating that moderate coupling induces an interlayer payoff synergistic

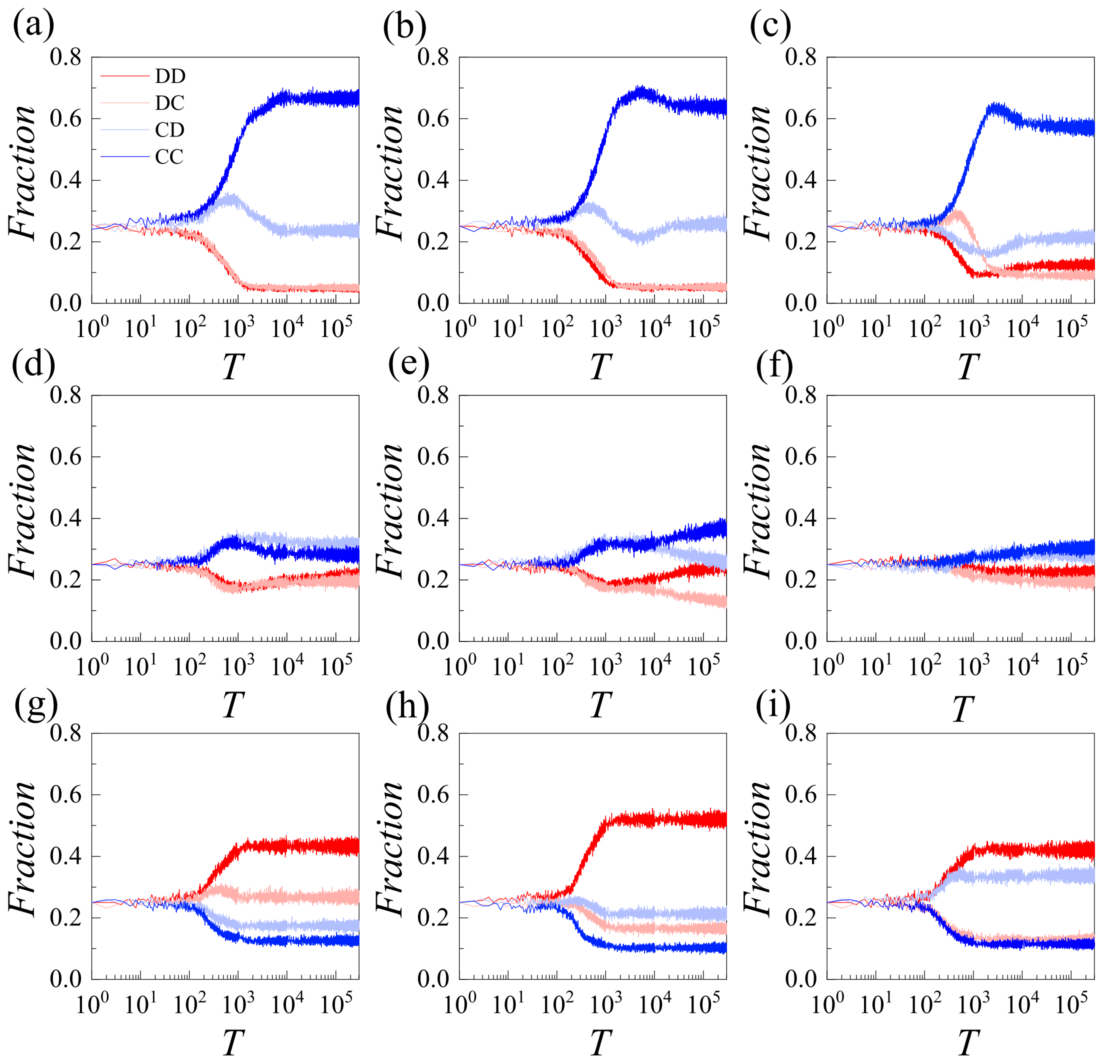

Fig. 6. Time evolution of agent strategy combinations (DD: defection in both layers; DC: defection in 2-SG and cooperation in 3-SG; CD: cooperation in 2-SG and defection in 3-SG; CC: cooperation in both layers) under different dilemma intensities and interlayer coupling strengths. The first row (a–c) corresponds to $r = 0.1$ with α values of 0.0, 0.1, and 0.9, respectively. The second row (d–f) corresponds to $r = 0.5$ with α values of 0.1, 0.4, and 0.9, respectively. The third row (g–i) corresponds to $r = 0.9$ with α values of 0.1, 0.4, and 0.9.

effect. Conversely, under high dilemma intensity ($r = 0.9$), the DD strategy reaches its peak fraction at $\alpha = 0.4$, directly driving the decline in macroscopic cooperation. This confirms that intermediate coupling amplifies the spread of defection under high dilemma intensity.

Closer inspection of the evolutionary paths of strategy fractions reveals a clear threshold in agents' switching behavior: when the coupling strength is weak ($\alpha < 0.35$), the comprehensive payoff is primarily determined by the 2-SG layer. To maximize the cumulative reward, the Q-learning algorithm tends to keep the strategy in the high-weight layer unchanged while exploring adjustments in the low-weight layer, since altering the high-weight strategy would cause larger payoff fluctuations. This results in the system exhibiting transitions between DD and DC, and between CC and CD. Conversely, when the coupling strength is strong ($\alpha \geq 0.35$), the comprehensive payoff is dominated by the 3-SG layer, and the algorithm exhibits the opposite tendency, leading to transitions between CD and DD, and between DC and CC. Thus, the interlayer coupling strength α not only modulates the overall cooperation frequency but also governs the priority of agent decision-making, which naturally emerges from the Q-learning process under different payoff weight distributions.

Fig. 6 presents the evolutionary trajectories of the four strategy combinations (DD, DC, CD, CC) over time steps T, under different dilemma intensities r and interlayer coupling strengths α. These dynamics illustrate how the social dilemma and interlayer interactions jointly shape agents' learning process. From the perspective of evolutionary dynamics, the dilemma intensity r

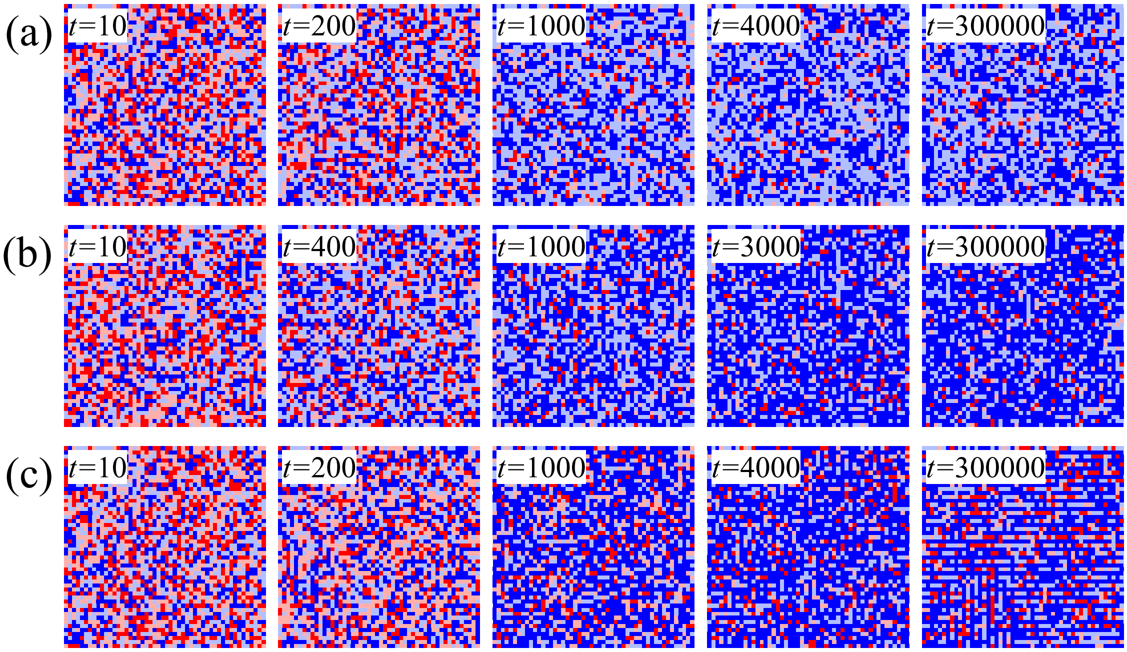

Fig. 7. Snapshots at different time steps for $r = 0.1$, where different colors distinguish the strategy combinations of agents. Four colors are used to denote the strategy combinations (dark red: DD, defection in both layers; light red: DC, defection in 2-SG and cooperation in 3-SG; light blue: CD, cooperation in 2-SG and defection in 3-SG; dark blue: CC, cooperation in both layers). Panels (a), (b), and (c) correspond to $\alpha = 0$, $\alpha = 0.1$, and $\alpha = 0.9$.

constitutes the background evolutionary pressure of the system. Under low dilemma intensity ($r = 0.1$), strategy CC rapidly establishes dominance under weak coupling through the positive feedback mechanism of Q-learning. Conversely, under high dilemma intensity ($r = 0.9$), the strong temptation to defect drives a rapid and monotonic rise of strategy DD, indicating that dilemma intensity governs the evolutionary outcome under extreme survival pressure. Meanwhile, the coupling strength α plays a dual role in this process: it modulates the payoff and affects the convergence time of evolution, imparting significant nonlinear characteristics to the strategy dynamics.

Specifically, under low dilemma intensity, strong coupling ($\alpha = 0.9$) introduces dynamic instability: the fraction of strategy CC initially rises but subsequently declines. This indicates that excessive interlayer information flow allows defection signals from the 3-SG layer to erode the cooperative equilibrium established in the 2-SG layer. More notably, under intermediate dilemma intensity ($r = 0.5$), the system exhibits a significantly prolonged convergence time and increased complexity in its evolutionary path. This arises from a delay in value updates as agents integrate conflicting payoffs from the two layers. When the immediate rewards from each layer point toward opposing strategies, the resulting fluctuations in the interlayer payoff coupling prevent the Q-table from forming stable strategy valuations. Agents thus undergo extended exploration, substantially slowing convergence to a steady state. In summary, Fig. 6 not only verifies the dynamic shift of strategy dominance over time but also reveals how coupling strength shapes the system's evolutionary path by modulating payoff structure, thereby establishing a coherent link between macroscopic evolution and microscopic learning.

3.3. Spatial dynamics of cooperative behavior

To further unveil the evolutionary logic of macroscopic cooperative behavior within the topological space, we analyze the dynamic evolution snapshots of strategy combinations on the square lattice over time (Figs. 7 and 8). This analysis explores how agents resist the invasion of defectors through spatial clustering effects under varying environmental pressures.

When $r = 0.1$, as illustrated in Fig. 7, cooperation exhibits a robust capacity for spatial expansion. Initially, the four joint strategies are randomly distributed across the lattice. As evolution proceeds, strategy CC rapidly forms local cooperative clusters through spatial reciprocity and expands outward by leveraging its payoff advantages acquired during Q-learning. The efficiency of this expansion depends nonlinearly on the interlayer coupling strength α. Under moderate coupling ($\alpha = 0.1$), positive interlayer payoff coupling enhances the ability of cooperative clusters to assimilate neighboring defectors, enabling the system to converge to a nearly all-cooperative steady state within a short period. In contrast, under strong coupling ($\alpha = 0.9$), the spatial dynamics exhibit increased complexity and strategy differentiation. Here, defection signals from the 3-SG layer persistently interfere with the cooperative stability of the 2-SG layer through strong coupling. This drives the system from a uniform cooperative equilibrium to a dynamic steady state characterized by spatial competition and stripe-like coexistence between cooperative and heterogeneous strategies. Such

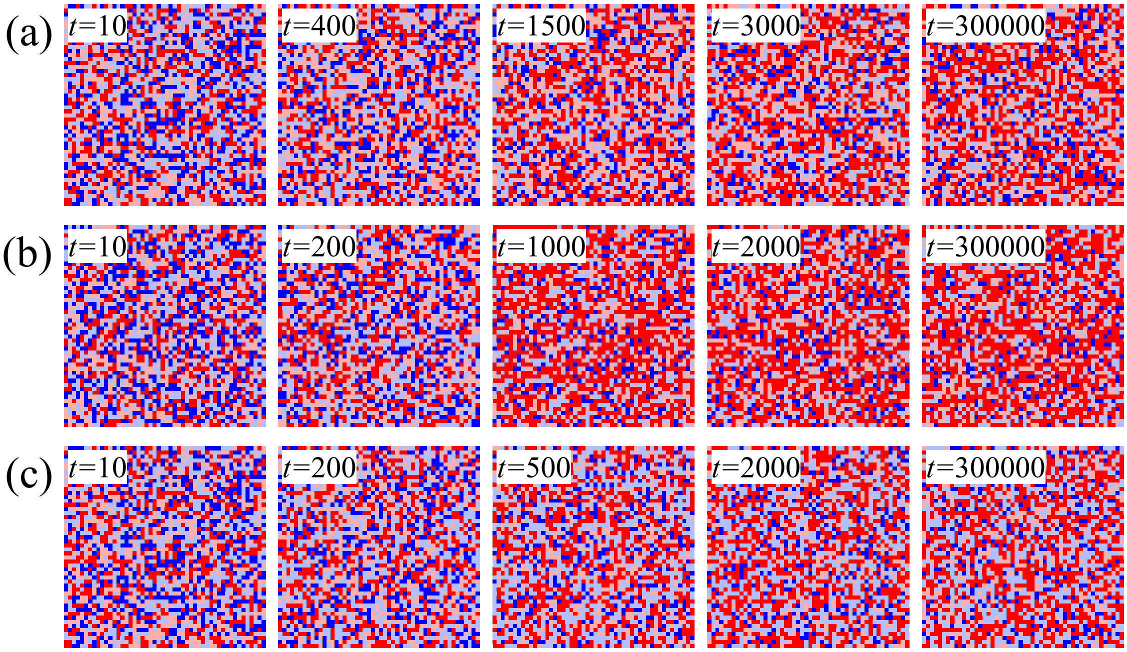

Fig. 8. Snapshots at different time steps for $r = 0.9$, where different colors distinguish the strategy combinations of agents. Four colors are used to denote the strategy combinations (dark red: DD, defection in both layers; light red: DC, defection in 2-SG and cooperation in 3-SG; light blue: CD, cooperation in 2-SG and defection in 3-SG; dark blue: CC, cooperation in both layers). Panels (a), (b), and (c) correspond to $\alpha = 0.1$, $\alpha = 0.4$, and $\alpha = 0.9$.

stripe patterns reflect spatial self-organization induced by interlayer game conflicts under strong coupling, representing a balance between local competition and evolutionary dynamics driven by payoff incentives.

Under high dilemma intensity ($r = 0.9$), as shown in Fig. 8, defection becomes the dominant strategy. In the initial stage, even if scattered cooperative clusters exist, the DD strategy (dark red regions) rapidly invades and dismantles them by exploiting its high immediate payoffs. Spatially, this invasion manifests as the expansion of DD from its initial random seeds, continuously reinforcing its dominance through its payoff advantage within the Q-learning algorithm. Under weak coupling ($\alpha = 0.1$), DD expands rapidly and ultimately leads to a steady state co-dominated by DD (full defection) and the mixed strategy DC. The coupling strength α modulates both the invasion efficiency of defection and the resulting spatial configuration. When coupling is increased to a moderate level ($\alpha = 0.4$), the system reaches its most pronounced defection equilibrium. The snapshots show that in this regime, DD dominates the entire network within a very short time, nearly eliminating all other strategies. This indicates that under extreme dilemma, moderate interlayer coupling unexpectedly accelerates the spread and consolidation of defection signals. However, under strong coupling ($\alpha = 0.9$), the spatial dynamics exhibit pronounced strategy differentiation. The system does not converge to a single homogeneous strategy; instead, it evolves into a pattern of spatial competition, where DD (dark red) and CD (light blue) form stable clusters of comparable size.

In summary, the spatial evolution of cooperation strongly depends on the joint effects of dilemma intensity r and interlayer coupling α. Under low dilemma intensities, cooperative clusters expand by converting neighboring defectors, with this efficiency peaking at moderate coupling. Conversely, under high dilemma intensities, defection possesses an inherent invasion advantage.

3.4. Topological robustness analysis

To further verify the robustness of our model across different network topologies, we replace the regular square lattice with a heterogeneous scale-free simplicial complex network in this section. With its power-law degree distribution, this network captures the heterogeneity of agent connectivity observed in real-world systems, allowing a more accurate characterization of structural heterogeneity in higher-order interactions.

Numerical results show that the dependence of the steady-state cooperation frequency f_C on the interlayer coupling strength α remains robust under this heterogeneous topology. As shown in Fig. 9(a), for different dilemma intensities r, the effect of α on f_C persists: in the low dilemma region (small r), f_C first increases and then decreases with α; conversely, in the high dilemma region (large r), it first decreases and then increases. This indicates that the modulating effect of α is robust across different network topologies, further confirming that the interlayer payoff coupling mechanism fundamentally drives the evolution of cooperation. As shown in Fig. 9(b), the average population payoff on the two-layer simplicial complex network exhibits a similar dependence on α, confirming the robustness of the welfare dynamics to topological changes. The overall trend remains consistent with the

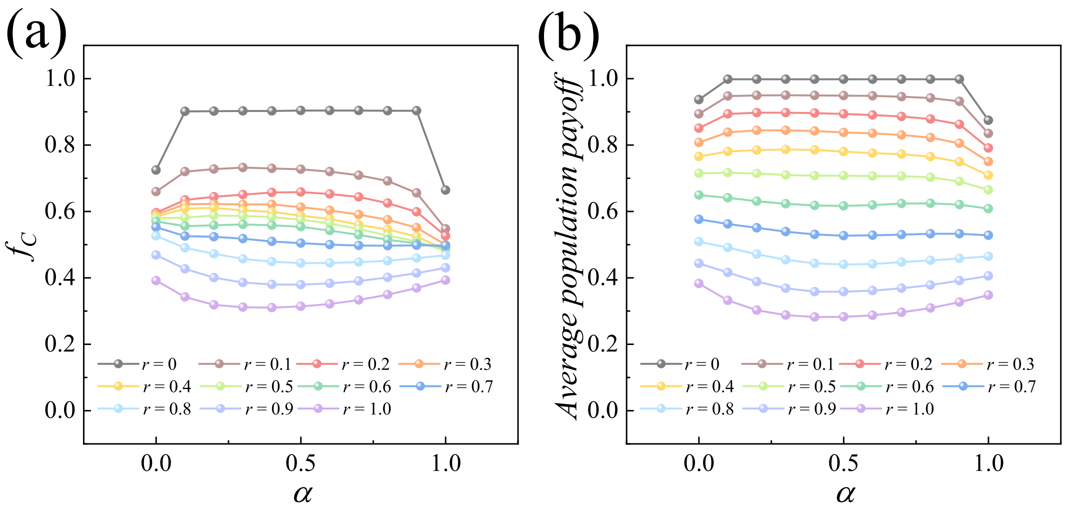

Fig. 9. The frequency of cooperators f_C (a) and the average population payoff (b) as functions of the coupling strength α under different dilemma intensities r on the two-layer simplicial complex network. The average population payoff is computed as $\frac{1}{N}\sum_{x=1}^{N}\frac{\bar{\pi}_x^{/}+\bar{\pi}_x^{\triangle}}{2}$ during the steady state.

cooperation frequency, and the local discrepancy between welfare and cooperation level is also observed in this heterogeneous topology, consistent with the findings on regular lattices. Furthermore, we observe that under identical parameter configurations, the steady-state cooperation frequency on the two-layer simplicial complex network is consistently higher than that on the regular lattice. This phenomenon corroborates the classic conclusion in complex network theory that "heterogeneity promotes cooperation": the heterogeneous degree distribution enables high-degree nodes (hubs) to more easily form and stabilize cooperative clusters. As primary drivers of cooperation, these hubs spread cooperative behavior to their numerous neighbors, thereby enhancing both the robustness and frequency of cooperation at the global scale.

4. Conclusion

In summary, this study explores the emergence of cooperation in multilayer higher-order systems by integrating a novel self-regarding Q-learning algorithm. Our results reveal that the interlayer coupling strength α exerts a non-monotonic modulatory effect on cooperation, a process significantly governed by the dilemma intensity r. Specifically, a micro-level transition in decision-making logic, which shifts from pairwise to higher-order interaction priority, serves as the fundamental driver for the observed macroscopic phase transitions. These findings are robust across both regular lattices and heterogeneous simplicial complexes, confirming that the interplay between interlayer payoff coupling and network topology fundamentally shapes evolutionary dynamics. By highlighting how agents adaptively weigh payoffs across different interaction scales, this research provides a nuanced understanding of strategy evolution in complex social dilemmas.

While the proposed framework offers novel insights into payoff integration mechanisms, several important directions merit further investigation. First, the current model assumes a static network topology; however, real-world social systems are inherently co-evolutionary. Future studies could incorporate adaptive learning parameters or dynamic interaction structures, wherein agents rewire their connections based on payoff feedback. Second, this study focuses on the competition between two-player and three-player snowdrift games. Extending the analysis to general n-player hypergraphs or alternative game paradigms would help establish more universal principles governing higher-order cooperation. Finally, examining the influence of environmental factors on strategy evolution would provide a more comprehensive understanding of cooperative dynamics in complex systems.

From a methodological perspective, the proposed two-layer payoff coupling mechanism and Q-learning evolutionary framework can be transferred to broader research scenarios beyond snowdrift games. For the evolution of fairness, the interlayer coupling can characterize the interaction between individual equity preferences and population-level distribution rules, while the Q-learning framework captures the adaptive updating of fairness strategies [38]. For trust evolution, the dual-layer structure is suitable for modeling the coevolution of pairwise direct trust and group-level indirect reputation [39]. The framework also has application potential in engineering and governance domains, including hierarchical regulation of AI safe development [40], technological coordination among firms with commitment constraints [50], and cooperative incentive design for open data management [51]. Exploring the applicability of the two-layer coupling framework in these domains will be a meaningful direction for future research.

CRediT authorship contribution statement

Rongrong Lin: Writing – original draft, Software, Methodology, Investigation, Conceptualization. **Sihang Li:** Writing – review & editing, Formal analysis, Conceptualization. **Haojie Xu:** Methodology, Formal analysis, Conceptualization. **Matjaž Perc:** Writing –

review & editing, Conceptualization. **Luoluo Jiang:** Writing – review & editing, Investigation, Conceptualization. **Changwei Huang:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization.

Declaration of competing interest

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

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant No. 12565005, in part by the Guangxi Science and Technology Base and Talent Special Project under Grant No. AD25069071, in part by the Open Project of Key Laboratory of Mathematics and Information Networks (Beijing University of Posts and Telecommunications), Ministry of Education, China, under Grant No. KF202402, and in part by the Natural Science and Technology Innovation Development Foundation of Guangxi University under Grant No. 2024BZRC011. Matjaž Perc was supported by the Slovenian Research and Innovation Agency (Grant No. P1-0403).

Data availability

Data will be made available on request.

References

[1] Smith JM, Price GR. The logic of animal conflict. Nature 1973;246(5427):15–8.
[2] Axelrod R, Hamilton WD. The evolution of cooperation. Science 1981;211(4489):1390–6.
[3] Christakis NA, Fowler JH. Friendship and natural selection. Proc Natl Acad Sci 2014;111(Suppl. 3):10796–801.
[4] Hardin G. The tragedy of the commons: The population problem has no technical solution; it requires a fundamental extension in morality. Science 1968;162(3859):1243–8.
[5] Taylor PD, Jonker LB. Evolutionary stable strategies and game dynamics. Math Biosci 1978;40(1–2):145–56.
[6] Darwin C. On the origin of species: A facsimile of the first edition. Harvard University Press; 1964.
[7] Smith JM. Evolution and the theory of games. Cambridge University Press; 1982.
[8] Cong R, Li K. Disguise as a medium for the evolution of cooperation. Appl Math Comput 2026;526:130084.
[9] Broom M, Rychtář J. A general framework for analysing multiplayer games in networks using territorial interactions as a case study. J Theoret Biol 2012;302:70–80.
[10] Battiston F, Cencetti G, Iacopini I, Latora V, Lucas M, Patania A, et al. Networks beyond pairwise interactions: Structure and dynamics. Phys Rep 2020;874:1–92.
[11] Lim IS, Masuda N. To trust or not to trust: Evolutionary dynamics of an asymmetric N-player trust game. IEEE Trans Evol Comput 2024;28(1):117–31.
[12] Zou K, Huang C. Incorporating reputation into reinforcement learning can promote cooperation on hypergraphs. Chaos Solitons Fractals 2024;186:115203.
[13] Pan J, Zhang L, Han W, Huang C. Heterogeneous investment promotes cooperation in spatial public goods game on hypergraphs. Phys A 2023;609:128400.
[14] Kovalenko K, Sendiña-Nadal I, Khalil N, Dainiak A, Musatov D, Raigorodskii AM, et al. Growing scale-free simplices. Commun Phys 2021;4(1):43.
[15] Guo S, Wang J, Zhao D, Xia C. Role of second-order reputation evaluation in the multi-player snowdrift game on scale-free simplicial complexes. Chaos Solitons Fractals 2023;172:113539.
[16] Alvarez-Rodriguez U, Battiston F, de Arruda GF, Moreno Y, Perc M, Latora V. Evolutionary dynamics of higher-order interactions in social networks. Nat Hum Behav 2021;5(5):586–95.
[17] Perc M. Diffusion dynamics and information spreading in multilayer networks: An overview. Eur Phys J Spec Top 2019;228(11):2351–5.
[18] Govaert A, Ramazi P, Cao M. Rationality, imitation, and rational imitation in spatial public goods games. IEEE Trans Control Netw Syst 2021;8(3):1324–35.
[19] Wang Z, Wang L, Szolnoki A, Perc M. Evolutionary games on multilayer networks: A colloquium. Eur Phys J B 2015;88(5):124.
[20] Battiston F, Perc M, Latora V. Determinants of public cooperation in multiplex networks. New J Phys 2017;19(7):073017.
[21] Sadekar O, Civilini A, Latora V, Battiston F. Drivers of cooperation in social dilemmas on higher-order networks. J R Soc Interface 2024;21(218):20240134.
[22] Chen J, Zou K, Zhang L, Huang C. Cooperation dynamics on higher-order networks: A brief review. Europhys Lett 2025;151(4):41001.
[23] Su Q, McAvoy A, Mori Y, Plotkin JB. Evolution of prosocial behaviours in multilayer populations. Nat Hum Behav 2022;6(3):338–48.
[24] Battiston F, Amico E, Barrat A, Bianconi G, Ferraz de Arruda G, Franceschiello B, et al. The physics of higher-order interactions in complex systems. Nat Phys 2021;17(10):1093–8.
[25] Wang F, Yu N, Li K. A novel reliability assessment method for binary-state networks. Chaos Solitons Fractals 2026;205:117868.
[26] You T, Yang H, Wang J, Zhang P, Chen J, Zhang Y. Cooperative behavior under the influence of multiple experienced guiders in prisoner's dilemma game. Appl Math Comput 2023;458:128234.
[27] You T, Zhang H, Zhang Y, Li Q, Zhang P, Yang M. The influence of experienced guider on cooperative behavior in the prisoner's dilemma game. Appl Math Comput 2022;426:127093.
[28] Zhu W, Wang X, Wang C, Xing W, Liu L, Zheng H, et al. Evolution of cooperation and competition in multilayer networks. Nonlinear Dynam 2025;113:31619–35.
[29] Civilini A, Sadekar O, Battiston F, Gómez-Gardeñes J, Latora V. Explosive cooperation in social dilemmas on higher-order networks. Phys Rev Lett 2024;132(16):167401.
[30] Kumar A, Chowdhary S, Capraro V, Perc M. Evolution of honesty in higher-order social networks. Phys Rev E 2021;104(5):054308.
[31] Wang C, Perc M, Szolnoki A. Evolutionary dynamics of any multiplayer game on regular graphs. Nat Commun 2024;15(1):5349.
[32] Ohtsuki H, Nowak MA. The replicator equation on graphs. J Theoret Biol 2006;243(1):86–97.
[33] Xu Y, Zhao D, Benko TP, Xia C, Perc M. Reinforcement learning can be a double-edged sword for cooperation on higher-order networks. IEEE Trans Syst Man, Cybern: Syst 2026;56(1):231–44.
[34] Shi Y, Rong Z. Analysis of Q-learning like algorithms through evolutionary game dynamics. IEEE Trans Circuits Syst II: Express Briefs 2022;69(5):2463–7.

[35] Wang L, Fan L, Zhang L, Zou R, Wang Z. Synergistic effects of adaptive reward and reinforcement learning rules on cooperation. New J Phys 2023;25(7):073008.
[36] Xu Y, Wang J, Chen J, Zhao D, Özer M, Xia C, Perc M. Reinforcement learning and collective cooperation on higher-order networks. Knowl-Based Syst 2024;301:112326.
[37] Zou K, Huang C. Cooperation dynamics on hypergraphs with punishment and Q-learning. Expert Syst Appl 2025;263:128989.
[38] Cimpeanu T, Di Stefano A, Perret C, Han TA. Social diversity reduces the complexity and cost of fostering fairness. Chaos Solitons Fractals 2023;167:113051.
[39] Kumar A, Capraro V, Perc M. The evolution of trust and trustworthiness. J R Soc Interface 2020;17(169):20200491.
[40] Cimpeanu T, Santos FC, Pereira LM, Lenaerts T, Han TA. Artificial intelligence development races in heterogeneous settings. Sci Rep 2022;12(1):1723.
[41] Jiang LL, Li YM, Li WJ, Szolnoki A. Combined effect of incentives and coupling in multigames in two-layer networks. J Phys: Complex 2024;5(2):025010.
[42] You T, Yang L, Wang J, Zhang P, Chen J, Zhang Y. The guidance of neutral human populations maintains cooperation in the prisoner's dilemma game. Appl Math Comput 2025;486:129071.
[43] Huang C, Han W, Li H, Cheng H, Dai Q, Yang J. Public cooperation in two-layer networks with asymmetric interaction and learning environments. Appl Math Comput 2019;340:305–13.
[44] Wang J, Guo S, Xia C, Perc M. Utility coupling promotes cooperation in multiplayer snowdrift games on interdependent simplicial networks. Eur Phys J Spec Top 2024;233(4):831–42.
[45] Xu Y, Feng M, Zhu Y, Xia C. Multi-player snowdrift game on scale-free simplicial complexes. Phys A 2022;604:127698.
[46] Kahneman D, Tversky A. Prospect theory: An analysis of decision under risk. Econometrica 1979;47(2):263–91.
[47] Tversky A, Kahneman D. Advances in prospect theory: Cumulative representation of uncertainty. J Risk Uncertain 1992;5(4):297–323.
[48] Han TA, Song Z, Cimpeanu T, Duong MH, Krellner M, Capraro V, Perc M. Cooperation versus social welfare. Phys Life Rev 2026;56:33–60.
[49] Song Z, Han TA. Emergence of cooperation and commitment in optional prisoner's dilemma. Appl Math Model 2026;155:116603.
[50] Ogbo NB, Cimpeanu T, Di Stefano A, Han TA. Shake on it: The role of commitments and the evolution of coordination in networks of technology firms. Artif Life Conf Proc 2022;34(1):41.
[51] Benko TP, Pi B, Li Q, Feng M, Perc M, Vošner HB. Evolutionary games for cooperation in open data management. Appl Math Comput 2025;496:129364.