Game-theoretical approach for opinion dynamics on social networks

Zhifang Li and Xiaojie Chen*
School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 611731, China

Han-Xin Yang
Department of Physics, Fuzhou University, Fuzhou 350108, China

Attila Szolnoki
Institute of Technical Physics and Materials Science, Centre for Energy Research, P.O. Box 49, H-1525 Budapest, Hungary

Opinion dynamics on social networks have been received considerable attentions in recent years. Nevertheless, just a few works have analyzed theoretically the condition in which a certain opinion can spread in the whole structured population. In this paper, we propose an evolutionary game approach for a binary opinion model to explore the conditions for an opinion's spreading. Inspired by real-life observations, we assume that an agent's choice to select an opinion is not random, but is based on a score rooted from public knowledge and the interactions with neighbors. By means of coalescing random walks, we obtain a condition in which opinion A can be favored to spread in the weak selection limit. In particular, when individuals adjust their opinions based solely on the public information, the vitality of opinion A depends solely on the difference of basic scores of A and B. Furthermore, when there are no negative feedback interactions between connected individuals, we find that opinion A can be favored if the ratio of the obtained positive feedback scores of competing opinions exceeds a critical value. To complete our study, we perform computer simulations on fully-connected, small-world, and scale-free networks, respectively, which support and confirm our theoretical findings.

The spreading of opinions on social networks can be detected in several ways in modern times and to understand related collective behaviors is the focus of research interest in last decades. Several examples can be raised ranging from political elections to fashion and marketing which influence our daily life. Albeit the microscopic process seems to be simple, it is still challenging to find analytical solutions. Our present theoretical approach utilizes the fact that to change an opinion, which can be considered as a decision-making process, depends not only public knowledge, but also on local interactions. In particular, we introduce an evolutionary game-theoretical approach with which an opinion update depends directly on a score. This score is calculated both from globally reachable public knowledge and also from restricted interactions with neighbors. Our principal goal is to provide theoretical conditions which ensure the spreading of opinion A in a binary opinion model. By means of coalescing random walks, we obtain a theoretical condition and study different cases to identify decisive factors. When only public information is available, the final evolutionary outcome depends exclusively on the difference of basic scores of binary opinions. However, when only local interactions without negative feedbacks with neighbors are considered, an opinion's final diffusion depends sensitively on the ratio of the obtained positive feedback scores of competing opinions and this ratio should exceed a critical value. For a more comprehensive study, we have checked interaction graphs with different topologies and confirmed our theoretical predictions by computer simulations.

* xiaojiechen@uestc.edu.cn (corresponding author)

I. INTRODUCTION

Because of its paramount importance in several social phenomena, the study of opinion dynamics in structured populations has become an intensively studied research area in the last decades [1–7]. In order to study the evolution and diffusion of opinions among interacting agents, a huge variety of mathematical models have been proposed [8–17]. In general, these models can be divided into two main categories: the first branch assumes discrete opinion model in which individuals take the discrete opinion values, like voter model [12], majority rule model [13], or Sznajd model [14]. The other approach uses continuous opinion models in which continuously distributed opinion values are considered, including Defffuant model [15], HK model (Hegselmann-Krause model) [16, 17], and so on.

It is worth mentioning that the above mentioned models have provided theoretical paradigms for studying opinion dynamics on social networks. In a realistic world, when an individual chooses or changes a certain opinion, then it could be the consequence of interactions with neighbors. From this viewpoint, how to choose an opinion can be regarded as a decision-making process with strategic interactions [18, 19]. In general this is a missing feature from previous theories, and thus it is meaningful to characterize the microscopic process of opinion choice in the mentioned way [20–26]. Evolutionary game theory, as a powerful mathematical tool, can be used to achieve this goal. Recently, some related efforts have been made along this research path. For example, Di Mare and Latora studied how to use game theory in the modeling of opinion formations in a way to simulate the basic interaction mechanisms between two individuals and particularly have shown how opinion formation can be obtained by just changing the rules of the game [20]. Subsequently, Yang proposed a consensus model of binary opinion in the framework of evolutionary games and studied how the necessary time to reach a consensus state can be reduced [23]. It was found that there exists an optimal cost-benefit ratio in the game leading to the shortest consensus time. Furthermore, Zhou *et al.* introduced conformity-driven teaching ability into the evolutionary process of opinion dynamics [24]. By means of computer simulations, they found that when the teaching ability strongly depends on the conformity, the consensus time can be shortened significantly. On the other hand, Lorenzo *et al.* proposed a novel model that captures the coevolution of actions and opinions on social networks and considers the interplay between the dynamics of actions and opinions. To be more specific, each agent updates his/her opinion, depending on the opinions shared by others, the actions observed on the network, and possibly an external influence source [25]. However, the mentioned work does not take into account the microscopic opinion updating with strategic interactions. It is worth pointing out that most of previously related works are based on computer simulations and thus far a few studies have tried analytical calculations to identify the theoretical conditions of successful opinion spreading in a structured population.

In this work, we thus would like to analyze the evolutionary process of opinion spreading on social networks theoretically where we integrate the evolutionary game approach. Accordingly, we consider binary opinion and assume that each individual can choose one of the two opinions A and B. When an individual chooses an opinion, he/she can obtain a basic score based on the available public information. Practically, in many situations the opinion choice of individuals are also influenced by the decisions of their neighbors or friends. We therefore also assume that individuals can interact with each others via pairwise interactions [27]. In particular, when an individual interacts with a neighbor who shares the same opinion, our agent receives a positive score due to a positive supporting feedback effect. Otherwise, our agent gets a negative score because of the conflicting decisions which can be implemented as a negative feedback. This is a psychologically reasonable assumption because the positive score received from an interaction with akin neighbor expresses a sense of belonging to the same group, and the negative score represents a stress of nonconformity [28–31]. This argument establishes a certain interaction of neighboring agents that can be handled in the framework of a game-theoretical model [32–35]. Having considered the mentioned interactions, the final score can be calculated which determines how an agent updates personal opinion. In general, we can say that during the evolutionary process individuals will imitate neighboring opinions which provide higher individual score for them [36].

Based on the above description, we propose an evolutionary game-theoretical model of binary opinion on social networks and accordingly study the evolution of binary opinion with strategic interactions. Indeed, one crucial quantity for studying evolutionary dynamics of binary opinions on networks is the fixation probability ρ_A of opinion A, which means the probability that individuals with opinion A take over the whole structured population given that initially an individual at a vertex is chosen randomly to have opinion A in a population of individuals with opinion B. Furthermore, if $\rho_A > 1/N$, then opinion A is favored, which indicates that opinion A can spread on the whole network. Hence, using evolutionary game theory we derive the formula of fixation probability of opinion A by calculating coalescence times [37–40], and obtain the condition in which opinion A can be favored to spread in the weak selection limit. We find that whether opinion A can diffuse or not depends on the score parameters we considered and the weight values in the network structure. Particularly, when individuals only adjust their opinions based on the basic score derived from the information of public knowledge, we find that the success spreading of opinion A depends solely on the difference of basic scores of A and B. Besides, we consider a special case of our game model in which the negative feedback effects of strategic interactions of opinions are ignored and the two basic scores about opinion A and B are the same for the evolution of binary opinion. In this particular case, we find that opinion A can be favored in the weak selection limit if the ratio of positive scores of competing opinions exceeds a critical value. We finally carry out computer simulations on fully-connected, small-world, and scale-free networks to confirm the robustness of our theoretical predictions.

The rest of this paper is structured as follows. We first introduce the basic definition and construct our model in Section 2. Then, we obtain the condition in which opinion A can be favored to spread under weak selection and carry out simulations to validate our theoretical results in Section 3. Finally, we summarize our conclusions in Section 4.

II. MODEL

We consider a structured population of individuals with size N. The interaction graph is represented by a weighted graph G with edge weight $\omega_{ij} \geq 0$, where individuals are located on the nodes and interactions with neighbors are linked by edges as shown in Fig. 1(a). Here, graph G is undirected and self-loops are not allowed. The weight of vertex i is indicated by $\omega_i = \sum_{j \in G} \omega_{ij}$ and the total weight value of all vertices is expressed as $W = \sum_{i \in G} \omega_i = \sum_{i,j \in G} \omega_{ij}$.

During the evolutionary process, each individual will choose between the optional opinions A and B, which are marked as 1 and 0, respectively. In this way, a particular state of all individuals can be represented by a vector $\mathbf{s} = (s_i)_{i \in G}$, where $s_i \in \{0, 1\}$ denotes the opinion that individual i chooses. When an individual chooses an opinion, we assume that he/she can obtain a basic score about the chosen opinion based on the information of public knowledge, which is set as δ_A for opinion A and δ_B for opinion B. Here, the basic score is derived from the information of public knowledge. In addition, each individual with the chosen opinion can interact with his/her neighbors via pairwise interactions [41]. Such interactions can induce some feedback scores for all individuals. To be specific, we assume that when an individual with opinion $A(B)$ interacts with a neighbor who shares the same opinion $A(B)$, he/she can receive a positive score a (d) due to the positive feedback effect. Otherwise, when an

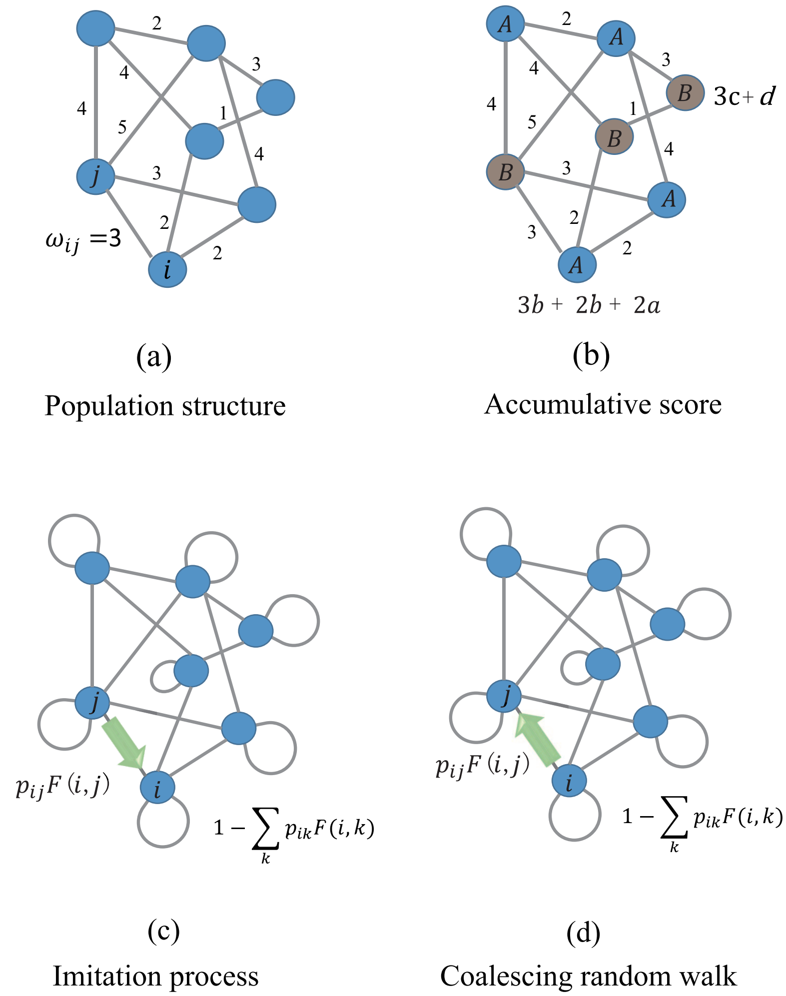

Fig. 1. Illustration of interaction-based opinion dynamics on weighted graphs. In panel (a), population structure is represented by a weighted and connected graph G with edge weights ω_{ij}. In panel (b), individual i interacts with nearest neighbors and gets an edge-weighted accumulated score in state $\mathbf{s}$, denoted by $P_i(\mathbf{s})$, which means the sum of scores obtained from each neighbor multiplied by the corresponding edge weight. In panel (c), a new graph G' is generated based on the graph G. The green arrow on G' indicates that individual i imitates the opinion of individual j. The imitation probability is $p_{ij}F(i,j)$ for $j \neq i$, otherwise individual i maintains the original opinion with probability $1-\sum_k p_{ik}F(i,k)$, indicated by the self-loop in the figure. In panel (d), individual i performs coalescing random walks on G'. The arrow in panel (d) indicates that a step from i to j is taken with probability $p_{ij}F(i,j)$. Here, coalescing random walk is a process of looking for ancestors backward, which is dual to the neutral case in our model.

individual with opinion $A(B)$ interacts with a neighbor having the opposite opinion $B(A)$, he/she gets a negative score b (c). The score matrix for the pairwise interactions [32–35] between two individuals can be thus described as

$$\begin{array}{c} \\ A \\ B \end{array} \begin{array}{c} \begin{array}{cc} A & B \end{array} \\ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \end{array}, \tag{1}$$

where the a and d matrix elements are positive representing positive feedback score values, while elements b and c are negative because of adverse response.

After interacting with all nearest neighbors, each individual i receives a total score $f_i(\mathbf{s})$, which means the sum of the basic score and the accumulative scores from the interactions with his/her neighbors, given as

$$f_i(\mathbf{s}) = s_i\delta_A + (1-s_i)\delta_B + P_i(\mathbf{s}), \tag{2}$$

where $P_i(\mathbf{s})$ represents the accumulated score value of individual i in state $\mathbf{s}$. Here, $P_i(\mathbf{s})$ means that the score value obtained from neighbors are multiplied by the corresponding edge weights and then summed, as illustrated in Fig. 1(b). Accordingly, $P_i(\mathbf{s})$ can be written as

$$\begin{aligned} P_i(\mathbf{s}) &= W\pi_i\left(as_is_i^{(1)} + bs_i\left(1-s_i^{(1)}\right)\right) \\ &\quad + W\pi_i\left(c\left(1-s_i\right)s_i^{(1)} + d(1-s_i)\left(1-s_i^{(1)}\right)\right) \\ &= W\pi_i\left(\left(a-b-c+d\right)s_is_i^{(1)} + (b-d)s_i\right) \end{aligned}$$

$$+ W\pi_i \left((c-d)\, s_i^{(1)} + d \right), \tag{3}$$

where $s_i^{(1)} = \Sigma_{i=1}^{N} p_{ij} s_j$ describes the expected type of nearest neighbors of individual i, $p_{ij} = \omega_{ij}/\omega_i$ represents the probability of the step from i to j for a random walk on G, and $\pi_i = \omega_i/W$ can be understood as the stationary probability of i in the stationary distribution of random walks on G [40].

Furthermore, we have

$$f_i(\mathbf{s}) = W\pi_i \left((a-b-c+d)\, s_i s_i^{(1)} + (b-d) s_i \right) + W\pi_i \left((c-d)\, s_i^{(1)} + d \right) + s_i\delta_A + (1-s_i)\delta_B. \tag{4}$$

Subsequently, each individual can update his/her opinion choice based on the total score information. By following previous studies about binary opinions [23, 24], the microscopic updating procedure of opinions on networks is governed by the "pairwise comparison" updating in our study [41–44]. To be specific, at each time step an individual i is randomly selected from the population to imitate the opinion of a randomly chosen neighbor j with probability $F(s_i, s_j)$, given as

$$F(s_i, s_j) = \frac{1}{1 + \exp[-\beta(f_j(\mathbf{s}) - f_i(\mathbf{s}))]}, \tag{5}$$

where β denotes the intensity of selection. For $\beta \to 0$, the selection is weak [45, 46], which means that the score value is only a small perturbation to the neutral drift that is a baseline at $\beta = 0$ [47]. In contrast, for $\beta \to +\infty$ limit the selection is strong in the sense that individual i will deterministically imitate the opinion of his/her neighbors with higher score values [48–51]. In this paper, we consider weak selection which is reasonable, because current score is only one of the factors considered by individual i who makes a decision about the opinion change.

III. RESULTS

Based on the above description, we can consider the evolutionary process as a continuous-time *Markov chain* $(\mathbf{S}(t))_{t\geq 0}$ and call it as *evolutionary Markov chain*, in which state transitions occur via imitation events [40, 52]. We indicate the imitation event that individual i randomly imitates the opinion of neighboring j by $j \to i$. Accordingly, the $j \to i$ event occurs with probability $R[j \to i]$, given as

$$\mathrm{R}[j \to i] = p_{ij} F(s_i, s_j). \tag{6}$$

Hence the probability that individual i keeps the original opinion is given as

$$\mathrm{R}[i \to i] = p_{ij}(1 - F(s_i, s_j)). \tag{7}$$

The *evolutionary Markov chain* in our model will be absorbed in one of the two states: all-A and all-B, which represents to the fixation of A and B, respectively. Notably, the system spends only a short intermediate time in mixed states [52].

Here, we focus on the fixation probability ρ_A of opinion A under weak selection [40], which represents the probability that the population state eventually reaches the absorbing state of all-A from an initial state $\mathbf{s}_0$ in which there are only one individual with opinion A and $N-1$ individuals with opinion B. We can obtain the mathematical expression of fixation probability ρ_A (for detailed derivations see Appendix), given as

$$\rho_A = \frac{1}{N} + \beta \langle D'(\mathbf{s}) \rangle_{\mathbf{u}}^{\circ} + o(\beta^2), \tag{8}$$

where $D'(\mathbf{s})$ is the first-order term of the Taylor-expansion of $D(\mathbf{s})$ at $\beta = 0$ and $D(\mathbf{s})$ represents the instantaneous rate of change in the degree-weighted frequency of opinion A from state $\mathbf{s}$. Here $\langle D'(\mathbf{s}) \rangle_{\mathbf{u}}^{\circ}$ is expressed as

$$\begin{aligned}
&\langle D'(\mathbf{s}) \rangle_{\mathbf{u}}^{\circ} \\
&= \frac{W}{2}(b-d) \left(\sum_{i,j} \pi_i^2 p_{ij} \left\langle s_i^2 - s_i s_j \right\rangle_{\mathbf{u}}^{\circ} \right) \\
&+ \frac{W}{2}(a-b-c+d) \left(\sum_{i,j,k} \pi_i^2 p_{ij} p_{ik} \left\langle s_i^2 s_k - s_i s_j s_k \right\rangle_{\mathbf{u}}^{\circ} \right) \\
&+ \frac{W}{2}(c-d) \left(\sum_{i,j,k} \pi_i^2 p_{ij} p_{ik} \left\langle s_i s_k - s_j s_k \right\rangle_{\mathbf{u}}^{\circ} \right) \\
&+ \frac{Wd}{2} \left(\sum_{i,j} \pi_i^2 p_{ij} \left\langle s_i - s_j \right\rangle_{\mathbf{u}}^{\circ} \right) \\
&+ \frac{(\delta_A - \delta_B)}{2} \left(\sum_{i,j} \pi_i p_{ij} \left\langle s_i^2 - s_i s_j \right\rangle_{\mathbf{u}}^{\circ} \right),
\end{aligned} \tag{9}$$

where

$$\langle s_i \rangle_{\mathbf{u}}^{\circ} = \int_0^{\infty} E_{\mathbf{u}}^{\circ}[S_i(t)]\mathrm{d}t, \tag{10}$$

$$\langle s_i s_j \rangle_{\mathbf{u}}^{\circ} = \int_0^{\infty} E_{\mathbf{u}}^{\circ}[S_i(t) S_j(t)]\mathrm{d}t, \tag{11}$$

and

$$\langle s_i s_j s_k \rangle_{\mathbf{u}}^{\circ} = \int_0^{\infty} E_{\mathbf{u}}^{\circ}[S_i(t) S_j(t) S_k(t)]\mathrm{d}t. \tag{12}$$

For calculating the quantities in Eq. (9), we use the concept of coalescence times which will be introduced in the next subsection.

A. Coalescing random walks

In order to calculate $\langle s_i \rangle_{\mathbf{u}}^{\circ}$, $\langle s_i s_j \rangle_{\mathbf{u}}^{\circ}$, and $\langle s_i s_j s_k \rangle_{\mathbf{u}}^{\circ}$, we consider coalescing random walks on graphs [37, 38, 40, 53]. Coalescing random walks on graph G are defined as a collection of random walks, which is a process corresponding to tracing backwards ancestors. Specifically, if individual i imitates the opinion of neighboring j at a certain step during the evolutionary process, then individual j is called the "ancestor" of individual i at that corresponding step.

During the evolutionary process in our model, individual i imitates the opinion of individual j under neutral drift ($\beta = 0$) with probability $\tilde{p}_{ij} = p_{ij} \frac{1}{1+exp[-\beta(f_j(\mathbf{s})-f_i(\mathbf{s}))]} = \frac{1}{2}p_{ij}$ when $j \neq i$, hence individual i keeps the original opinion with probability $\tilde{p}_{ii} = 1 - \sum_j \tilde{p}_{ij} = 1 - \sum_j \frac{1}{2}p_{ij} = \frac{1}{2}$. Therefore a new graph G' can be generated based on graph G, on which self-loops are introduced due to the possible opinion keeping by some individuals. We can then note that for a random walk on G', a step from i to j is taken with probability $\tilde{p}_{ij} = \frac{1}{2}p_{ij}$ for $j \neq i$, and a step from i to i is taken with probability $\tilde{p}_{ii} = \frac{1}{2}$. Hence, the case under neutral drift in our model is dual to the continuous-time coalescing random walks on G' as depicted in Fig. 1(c) and (d). Furthermore, we need to consider one-, two-, and three-dimensional coalescing random walks on G' to calculate the values of $\langle s_i \rangle^\circ_{\mathbf{u}}$, $\langle s_i s_j \rangle^\circ_{\mathbf{u}}$, and $\langle s_i s_j s_k \rangle^\circ_{\mathbf{u}}$, respectively.

B. One-dimensional coalescing random walks

For one-dimensional coalescing random walks on G', we assume that there is a walker $X(t)$ where $t \geq 0$, which occupies an arbitrary vertex i at the initial time, i.e., $X(0) = i$. At each time step, $X(t)$ takes a random step on G'. Each step is taken with the probability corresponding to the rate at which the imitation event occurs in the continuous-time evolutionary process.

As stated above, the case under neutral drift in our model is dual to the continuous-time coalescing random walks on G'. We thus can use the following equation to illustrate the duality relationship: for any initial state $\mathbf{s}_0$ and any two types of opinions $m, n \in \{1, 0\}$,

$$P^\circ_{\mathbf{s}_0}[S_i(t) = m] = P^{CRW}_i[(\mathbf{s}_0)_{X(t)} = m], \tag{13}$$

where $P^\circ_{\mathbf{s}_0}[]$ denotes the probability in the neutral evolutionary process started from state s_0, and $P^{CRW}_i[]$ represents the probability value in one-dimensional coalescing random walks started from i. Eq. (13) means that in the neutral case of evolutionary process, the probability of individual i with type m is the same to the probability that the walker $X(t)$ steps to the ancestor of individual i at time t in the coalescing random walks started from i. In other words, the type of individual i at time t is identical to the type of his/her ancestor at the initial state.

For a special initial state $\mathbf{s}_0$ satisfying $(\mathbf{s}_0)_l = 1$ and $(\mathbf{s}_0)_k = 0$ for all $k \neq l$, according to Eq. (13) we have

$$\begin{aligned} E^\circ_{\mathbf{s}_0}[S_i(t)] &= E^\circ_{\mathbf{s}_0}[S_i(t) = 1] = P^\circ_{\mathbf{s}_0}[S_i(t) = 1] \\ &= P^{CRW}_i[(\mathbf{s}_0)_{X(t)} = 1] = P^{CRW}_i[X(t) = l], \end{aligned} \tag{14}$$

where $E^\circ_{\mathbf{s}_0}[]$ denotes the expectation in the neutral evolutionary process started from state s_0.

We now define a probability distribution $\mathbf{u}$ over all states $\mathbf{s}$, assigning probability $\frac{1}{N}$ to the special states in which there is only one vertex with opinion A, and probability zero to other states. When the initial state is sampled from the probability distribution $\mathbf{u}$, according to Eq. (14) we have

$$E^\circ_{\mathbf{u}}[S_i(t)] = \frac{1}{N}\sum_l P^{CRW}_i[X(t) = l] = \frac{1}{N}. \tag{15}$$

Furthermore, we obtain $\langle s_i \rangle^\circ_{\mathbf{u}}$ as

$$\langle s_i \rangle^\circ_{\mathbf{u}} = \int_0^\infty E^\circ_{\mathbf{u}}(S_i(t))\mathrm{d}t = \int_0^\infty \frac{1}{N}\mathrm{d}t. \tag{16}$$

C. Two-dimensional coalescing random walks

For two-dimensional coalescing random walks on G', we consider continuous- and discrete-time versions, respectively [40]. In the continuous-time coalescing random walks, we assume that there are two walkers $(X(t), Y(t))_{t\geq 0}$, which occupy arbitrarily chosen i and j vertices at the initial time, i.e., $X(0) = i$ and $Y(0) = j$. At each time step, they walk independently until they meet for the first time (coalescence). The first meeting time in two-dimensional coalescing random walks on G' is denoted by $T^{(2)}_{coal}$. After this time, they walk together, that is, $X(t) = Y(t)$ for all $t > T^{(2)}_{coal}$. We define $P^{CRW}_{(i,j)}[]$ and $E^{CRW}_{(i,j)}[]$ to respectively represent the probabilities and expectations in continuous-time coalescing random walks started from i and j.

In the discrete-time coalescing random walks, we also assume that there are two walkers $(X(t), Y(t))^\infty_{t=0}$, which occupy randomly chosen i and j vertices at the initial time, i.e., $X(0) = i$ and $Y(0) = j$. Unlike continuous-time coalescing random walks, at each time step if $X(t) \neq Y(t)$, then one of $X(t)$ and $Y(t)$ will be randomly chosen to make a random step until their time of coalescing. If $X(t) = Y(t)$, they will take the same step at the next time step. We define $\tilde{P}^{CRW}_{(i,j)}[]$ and $\tilde{E}^{CRW}_{(i,j)}[]$ to respectively represent the probabilities and expectations in discrete-time and two-dimensional coalescing random walks on G' started from i and j.

Obviously, it can be seen that in the continuous-time version two steps are taken per unit time, while in the discrete-time version one step is taken every time step. In the discrete-time coalescing random walk started from i and j, the expected coalescence time is defined by $\tau_{ij} = \tilde{E}^{CRW}_{(i,j)}[T^{(2)}_{coal}]$. Due to the different numbers of steps per unit time between the two versions, we obtain the expected coalescence time in continuous-time coalescing random walks $E^{CRW}_{(i,j)}[T^{(2)}_{coal}] = \tau_{ij}/2$.

Due to the fact that the neutral drift case of our model is dual to the continuous-time coalescing random walks on G', for any initial state $\mathbf{s}_0$ and any two types of opinions $m, n \in \{1, 0\}$ we have

$$\begin{aligned} &P^\circ_{\mathbf{s}_0}[S_i(t) = m, S_j(t) = n] \\ &= P^{CRW}_{(i,j)}[(\mathbf{s}_0)_{X(t)} = m, (\mathbf{s}_0)_{Y(t)} = n], \end{aligned} \tag{17}$$

which indicates that the types of individual i and j at time t in the evolutionary process under neutral drift are the same to their corresponding ancestors at the initial state.

Furthermore, we consider a special initial state $\mathbf{s}_0$ satisfying $(\mathbf{s}_0)_l = 1$ and $(\mathbf{s}_0)_k = 0$ for all $k \neq l$. According to Eq. (17), we have

$$\begin{aligned} E^{\circ}_{\mathbf{s}_0}[S_i(t)S_j(t)] &= P^{\circ}_{\mathbf{s}_0}[S_i(t)=1, S_j(t)=1] \\ &= P^{CRW}_{(i,j)}[X(t)=Y(t)=l] \\ &= P^{CRW}_{(i,j)}[T^{(2)}_{coal} < t, X(t)=l]. \end{aligned} \tag{18}$$

When the initial state is sampled from the probability distribution $\mathbf{u}$, we have

$$\begin{aligned} E^{\circ}_{\mathbf{u}}[S_i(t)S_j(t)] &= \frac{1}{N}\sum_l P^{CRW}_{(i,j)}[T^{(2)}_{coal} < t, X(t)=l] \\ &= \frac{1}{N} P^{CRW}_{(i,j)}[T^{(2)}_{coal} < t], \end{aligned} \tag{19}$$

where the second equality can be derived by the law of total probability.

Based on the calculation of $E^{\circ}_{\mathbf{u}}[S_i(t)S_j(t)]$, we can have

$$\begin{aligned} \left\langle \frac{1}{N} - s_i s_j \right\rangle^{\circ}_{\mathbf{u}} &= \int_0^{\infty} \left(\frac{1}{N} - E^{\circ}_{\mathbf{u}}[S_i(t)S_j(t)]\right) \mathrm{d}t \\ &= \frac{1}{N}\int_0^{\infty}\left(1 - P^{CRW}_{(i,j)}[T^{(2)}_{coal} < t]\right)\mathrm{d}t \\ &= \frac{1}{N} E^{CRW}_{(i,j)}[T^{(2)}_{coal}] \\ &= \frac{\tau_{ij}}{2N}, \end{aligned} \tag{20}$$

where $\tau_{ij}/2$ denotes the expected coalescence time of continuous-time coalescing random walks.

Accordingly, we obtain the expression of $\langle s_i s_j \rangle^{\circ}_{\mathbf{u}}$ as

$$\langle s_i s_j \rangle^{\circ}_{\mathbf{u}} = \int_0^{\infty} \frac{1}{N}\mathrm{d}t - \frac{\tau_{ij}}{2N}. \tag{21}$$

The expected coalescence time τ_{ij} of two-dimensional and discrete-time coalescing random walks started from i and j satisfies the recurrence relation [40]

$$\tau_{ij} = \begin{cases} 0 & i = j \\ 1 + \frac{1}{2}\sum_{x\in G}\left(\tilde{p}_{ix}\tau_{jx} + \tilde{p}_{jx}\tau_{ix}\right) & i \neq j, \end{cases} \tag{22}$$

where Eq. (22) is a system of $\binom{N}{2}$ linear equations.

D. Three-dimensional coalescing random walks

Similar to the two-dimensional case, for three-dimensional coalescing random walks on G' we also consider continuous- and discrete-time versions. In the continuous-time coalescing random walks, we assume that there are three walkers $(X(t), Y(t), Z(t))_{t\geq 0}$, which occupy randomly chosen i, j, and k vertices at the initial time, i.e., $X(0) = i$, $Y(0) = j$, and $Z(0) = k$. At each time step, they walk independently until they meet for the first time. The first meeting time is represented by $T^{(3)}_{coal}$. After this time, they will walk together, i.e., $X(t) = Y(t) = Z(t)$ for $t > T^{(3)}_{coal}$. Particularly, if any two of them meet before $T^{(3)}_{coal}$, then the two walkers who have met will walk together afterwards. We define $P^{CRW}_{(i,j,k)}[]$ and $E^{CRW}_{(i,j,k)}[]$ to respectively denote the probabilities and expectations in the continuous-time and three-dimensional coalescing random walks on G' started from i, j, and k.

For discrete-time coalescing random walks, we also assume that there are three walkers $(X(t), Y(t), Z(t))_{t=0}^{\infty}$, which occupy arbitrary three i, j, and k vertices at the initial time, i.e., $X(0) = i$, $Y(0) = j$, and $Z(0) = k$. Unlike continuous-time coalescing random walks, at each time step if $X(t)$, $Y(t)$, and $Z(t)$ occupy different vertices, one of $X(t)$, $Y(t)$, and $Z(t)$ will be randomly chosen to make a random step until their time of coalescing. Particularly, if any two of them meet before coalescing, then these walkers stay together. Evidently, if $X(t) = Y(t) = Z(t)$, they all stay together in the following time. As previously, here we define $\tilde{P}^{CRW}_{(i,j,k)}[]$ and $\tilde{E}^{CRW}_{(i,j,k)}[]$ to respectively represent the probabilities and expectations in the discrete-time and three-dimensional coalescing random walks on G' started from i, j, and k.

Obviously, it can be seen that in a continuous-time three-dimensional coalescing random walk three steps are taken per unit time, while in a discrete-time coalescing random walk one step is taken every time. In the discrete-time version, the expected coalescence time of three-dimensional coalescing random walks started from i, j, and k is defined by $\tau_{ijk} = \tilde{E}^{CRW}_{(i,j,k)}[T^{(3)}_{coal}]$. Since the different numbers of steps per unit time between two versions, we obtain the expected coalescence time in continuous-time coalescing random walks $E^{CRW}_{(i,j,k)}[T^{(3)}_{coal}] = \tau_{ijk}/3$.

Similarly, the three-dimensional coalescing random walks in the continuous-time version have the same relationship with the neutral drift case in our model as found in two-dimensional coalescing random walks. For any initial state $\mathbf{s}_0$ and any two types $m, n \in \{1, 0\}$, we thus have

$$\begin{aligned} &P^{\circ}_{\mathbf{s}_0}[S_i(t)=m, S_j(t)=n, S_k(t)=m] \\ &= P^{CRW}_{(i,j,k)}[(\mathbf{s}_0)_{X(t)}=m, (\mathbf{s}_0)_{Y(t)}=n, (\mathbf{s}_0)_{Z(t)}=m], \end{aligned} \tag{23}$$

which indicates that the types of individual i, j, and k at time t in the evolutionary process under neutral drift are the same to their corresponding ancestors at the initial state.

For a special initial state $\mathbf{s}_0$ satisfying $(\mathbf{s}_0)_l = 1$ and $(\mathbf{s}_0)_k = 0$ for all $k \neq l$, we have

$$\begin{aligned} E^{\circ}_{s_0}[S_i(t)S_j(t)S_k(t)] &= P^{CRW}_{(i,j,k)}[X(t)=Y(t)=Z(t)=l] \\ &= P^{CRW}_{(i,j,k)}[T^{(3)}_{coal} < t, X(t)=l]. \end{aligned} \tag{24}$$

If the initial state is chosen from $\mathbf{u}$, according to Eq. (24) we obtain

$$E_{\mathbf{u}}^{\circ}[S_i(t)S_j(t)S_k(t)] = \frac{1}{N}\sum_{l} P_{(i,j,k)}^{CRW}[T_{coal}^{(3)} < t, X(t) = l] \\ = \frac{1}{N} P_{(i,j,k)}^{CRW}[T_{coal}^{(3)} < t]. \tag{25}$$

Based on the calculation of $E_{\mathbf{u}}^{\circ}[S_i(t)S_j(t)S_k(t)]$, we can have

$$\left\langle \frac{1}{N} - s_i s_j s_k \right\rangle_{\mathbf{u}}^{\circ} = \int_0^{\infty} \left(\frac{1}{N} - E_{\mathbf{u}}^{\circ}[S_i(t)S_j(t)S_k(t)]\right) \mathrm{d}t \\ = \frac{1}{N}\int_0^{\infty} \left(1 - P_{(i,j,k)}^{CRW}[T_{coal}^{(3)} < t]\right) \mathrm{d}t \\ = \frac{1}{N} E_{(i,j,k)}^{CRW}[T_{coal}^{(3)}] \\ = \frac{\tau_{ijk}}{3N}. \tag{26}$$

Accordingly we obtain the mathematical expression of $\langle s_i s_j s_k \rangle_{\mathbf{u}}^{\circ}$ as

$$\langle s_i s_j s_k \rangle_{\mathbf{u}}^{\circ} = \int_0^{\infty} \frac{1}{N} \mathrm{d}t - \frac{\tau_{ijk}}{3N}. \tag{27}$$

The expected coalescence time τ_{ijk} of discrete-time and three-dimensional coalescing random walks started from $i,\ j,\ k$ satisfies the recurrence relation

$$\tau_{ijk} = \begin{cases} 0 & i = j \neq k \\ 1 + \frac{1}{3}\sum_{x\in G}\left(2\tilde{p}_{ix}\tau_{xxk} + \tilde{p}_{kx}\tau_{iix}\right) & i \neq j = k \\ 1 + \frac{1}{3}\sum_{x\in G}\left(\tilde{p}_{ix}\tau_{xjj} + 2\tilde{p}_{jx}\tau_{ixx}\right) & i \neq j = k \\ 1 + \frac{1}{3}\sum_{x\in G}\left(\tilde{p}_{jx}\tau_{ixi} + 2\tilde{p}_{kx}\tau_{xjx}\right) & i = k \neq j \\ 1 + \frac{1}{3}\sum_{x\in G}\left(\tilde{p}_{ix}\tau_{xjk} + \tilde{p}_{jx}\tau_{ixk} + \tilde{p}_{kx}\tau_{ijx}\right) & i \neq j \neq k, \end{cases} \tag{28}$$

where Eq. (28) is a system of $\binom{N}{3}$ linear equations.

E. Condition for $\rho_A > \frac{1}{N}$

Applying Eqs. (9), (16), (21), and (27), we obtain the fixation probability of opinion A under weak selection as

$$\rho_A = \frac{1}{N} + \beta\langle D'(\mathbf{s})\rangle_{\mathbf{u}}^{\circ} + o(\beta^2), \tag{29}$$

where

$$\langle D'(\mathbf{s})\rangle_{\mathbf{u}}^{\circ} \\ = \frac{W}{2}(a - b - c + d)\sum_{i,j,k}\pi_i^2 p_{ij}p_{ik}\frac{\tau_{ijk} - \tau_{iik}}{3N} \\ + \frac{W}{2}(b - d)\sum_{i,j}\pi_i^2 p_{ij}\frac{\tau_{ij}}{2N} \\ + \frac{W}{2}(c - d)\sum_{i,j,k}\pi_i^2 p_{ij}p_{ik}\frac{\tau_{jk} - \tau_{ik}}{2N} \\ + \frac{(\delta_A - \delta_B)}{2}\sum_{i,j}\pi_i p_{ij}\frac{\tau_{ij}}{2N}. \tag{30}$$

Hence opinion A is favored to spread under weak selection, if and only if $\rho_A > \frac{1}{N}$, i.e.,

$$\langle D'(\mathbf{s})\rangle_{\mathbf{u}}^{\circ} > 0. \tag{31}$$

F. Theoretical results for representative cases

In this subsection, we consider three special but representative cases. For each case, we aim to derive the condition in which opinion A is favored in the weak selection limit.

Case I: We consider that individuals adjust their opinions only according to the basic evaluation score derived from the information of public knowledge. In other words, the positive and negative feedback scores from pairwise interactions are all zero, i.e., $a = b = c = d = 0$. In this case, we find that opinion A is favored under weak selection if and only if

$$\langle D'(\mathbf{s})\rangle_{\mathbf{u}}^{\circ} = \frac{(\delta_A - \delta_B)}{2}\sum_{i,j}\pi_i p_{ij}\frac{\tau_{ij}}{2N} > 0. \tag{32}$$

It can be seen that since $\sum_{i,j}\pi_i p_{ij}\frac{\tau_{ij}}{2N} > 0$, thus whether opinion A can be favored to spread or not depends entirely on the difference of basic scores of A and B. To be specific, opinion A is favored if and only if $\delta_A > \delta_B$.

Case II: We consider that $\delta_A = \delta_B$ and $b = c = 0$. In this case, individuals adjust their opinion mainly based on the positive feedback scores. Accordingly, we can obtain the mathematical condition in which opinion A is favored, given as

$$\frac{a}{d} > (\frac{a}{d})^* \\ = \frac{3\sum_{i,j}\pi_i^2 p_{ij}\tau_{ij} + 3\sum_{i,j,k}\pi_i^2 p_{ij}p_{ik}(\tau_{jk} - \tau_{ik})}{2\sum_{i,j,k}\pi_i^2 p_{ij}p_{ik}(\tau_{ijk} - \tau_{iik})} - 1, \tag{33}$$

which implies that when the ratio of the obtained positive scores of competing opinions exceeds the critical value $(\frac{a}{d})^*$, opinion A is favored.

Case III: We consider the donation game for our model and then the payoff matrix of donation game is given as

$$\begin{array}{c} \\ A \\ B \end{array}\!\!\begin{array}{c} \begin{array}{cc} A & B \end{array} \\ \begin{pmatrix} b - c & -c \\ b & 0 \end{pmatrix} \end{array}, \tag{34}$$

Furthermore, when $\delta_A = \delta_B$, we find that opinion A is favored under weak selection if and only if

$$\frac{b}{c} > \frac{\sum_{i,j} \pi_i^2 p_{ij} \tau_{ij}}{\sum_{i,j,k} \pi_i^2 p_{ij} p_{ik} (\tau_{jk} - \tau_{ik})}. \tag{35}$$

In this special case, we find that our prediction is consistent with the finding in Ref. [40], which verifies the effectiveness of our theoretical analysis.

G. Simulation results

In this subsection, to verify our theoretical results derived from Eq. (33), we carry out Monte Carlo simulations on three representative network structures including fully-connected, NW small-world [54], and BA scale-free networks [55]. All simulation results are obtained by performing the pairwise comparison updating rule. To reach the expected accuracy, we average the results over 10^4 independent runs for a specific structure. In each simulation run, all individuals initially choose opinion B, except a randomly chosen individual who chooses opinion A. The fixation probability is approximated as the fraction of Monte Carlo simulation runs which eventuate in all-A state. In Fig. 2, we show the fixation probability multiplied by the system size N as a function of the ratio a/d for the three representative graphs. We find that the fixation probability of A is close to $1/N$ to the greatest extent when a/d reaches the theoretical critical value $(a/d)^*$ indicated by the green dashed vertical line. When the ratio exceeds $(a/d)^*$, the fixation probability is larger than $1/N$, which means that A is favored to spread under weak selection. From Fig. 2, we can see that all simulation results are in good agreement with our numerical results obtained from Eq. (33).

IV. CONCLUSION

Our basic motivation in this work is to provide a theoretical framework for an opinion dynamics model where the microscopic rule of opinion update is based on a score value which is rooted from public knowledge about competing opinions and also depends on the interactions with neighbors. According to the obtained total score, each agent in the population can adjust individual opinion during the evolutionary process. Importantly, we have studied the opinion dynamics on any network structures in the framework of our model. By means of theoretical analysis, we have derived a mathematical expression of fixation probability of opinion A and further obtained the condition in which opinion A is favored under weak selection. We find that whether opinion A is favored or not depends sensitively on the score parameters and weight parameters of the network structure. In particular, when individuals adjust their opinions based only on the basic score, the diffusion of opinion A depends on the difference of basic scores of opinion A and B. In addition, we consider a special case in which the negative feedback effects of strategic interactions of opinions are ignored and find that there exists a critical value of the ratio a/d feedback parameters above which opinion A is favored under weak selection. Interestingly, the value of this critical ratio is related to the geometry of interaction graph and predicts the smallest value for strongly heterogeneous scale-free networks. To complete our study, we also carry out computer simulations where we use the above specified interaction graphs. These simulation data are in good agreement with our theoretical predictions and confirm the robustness of our findings. We hope that our work will contribute to a deeper understanding of different spreading process on complex networks [56–60].

In this study, we do not consider the issue of fixation time. Indeed, beside the fixation probability, the fixation time depicting the average time until the fixation occurs is another important quantity for studying evolutionary dynamics of binary opinions on social networks. In the framework of evolutionary games on graphs, some previous work have studied the fixation time of competing strategies on graphs [61–66]. For a future study it could be a promising extension to investigate the fixation time based on the theoretical approaches from evolutionary games on graphs.

APPENDIX

In the following, we mainly carry out the derivation of the fixation probability of opinion A. In order to describe the distribution of opinion A in the population in state $\mathbf{s}$, we define the degree-weighted frequency of opinion A as

$$\hat{s} = \sum_{i \in G} \pi_i s_i. \tag{A1}$$

The degree-weighted frequency of A at time T can be denoted by a random variable $\hat{S}(T)$, given as

$$\hat{S}(T) = \sum_{i \in G} \pi_i S_i(T). \tag{A2}$$

The weighting π_i in Eq. (A2) can be regraded as the reproductive value of vertex i in evolutionary game theory, which quantifies the fixation probability of opinion A under neutral drift [67–70].

We consider the *evolutionary Markov chain* started from arbitrary initial state $\mathbf{S}(0) = \mathbf{s}_0 \in \{0,1\}^G$. According to the Fundamental Theorem of Calculus [40], the expectation of degree-weighted frequency $E_{\mathbf{s}_0}[\hat{S}(T)]$ satisfies

$$E_{\mathbf{s}_0}[\hat{S}(T)] = \hat{s}_0 + \int_0^T \frac{\mathrm{d}}{\mathrm{d}t} E_{\mathbf{s}_0}[\hat{S}(t)] \mathrm{d}t. \tag{A3}$$

As stated above, the *evolutionary Markov chain* will be absorbed in the fixation states, all-A or all-B, for any given initial state. Thus, in the limit $T \to \infty$ the expectation of degree-weighted frequency $E_{s_0}[\hat{S}(T)]$ is equivalent to the fixation probability of type A, that is,

$$\rho_{\mathbf{s}_0} = \hat{s}_0 + \int_0^\infty \frac{\mathrm{d}}{\mathrm{d}t} E_{\mathbf{s}_0}[\hat{S}(t)] \mathrm{d}t. \tag{A4}$$

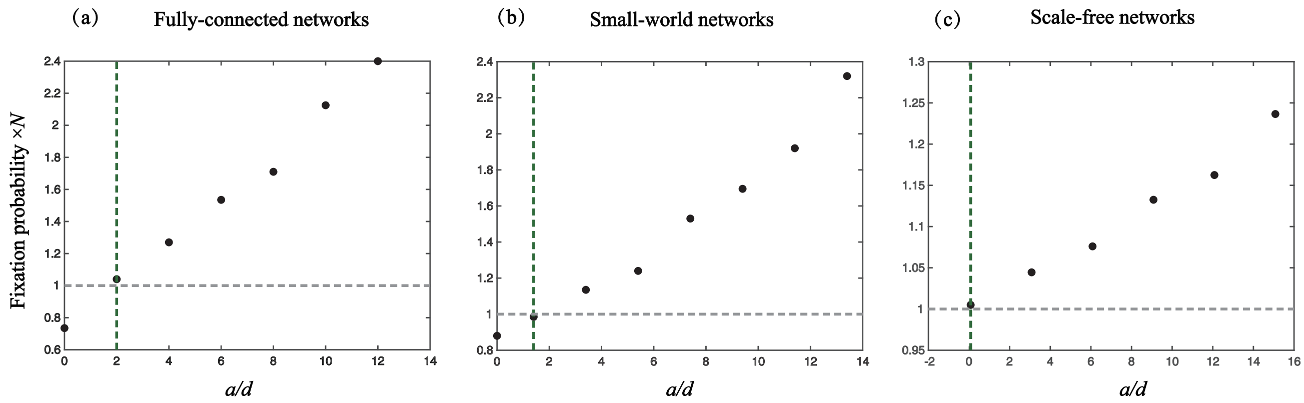

Fig. 2. The value of fixation probability ρ_A multiplied by the system size N as a function of the ratio a/d for three representative graphs, including complete graph in panel (a), small-world networks with the initial number of neighbors $k = 8$ and edge creation probability $p = 0.4$ in panel (b), and scale-free networks with initial number of nodes $m_0 = 3$ and linking number $m = 3$ in panel (c). The network size is all set to $N = 50$ and all edge weight values are set to 1. Simulation results are represented by black solid circles and all green vertical dashed lines indicate the critical value $(a/d)^*$ obtained from Eq. (33). When the a/d value is larger than $(a/d)^*$, opinion A is favored. Parameters are: $\delta_A = \delta_B$ and $b = c = 0$.

In the following part, we will focus on the calculation of $\rho_{\mathbf{s}_0}$. The crucial part is the integrand in Eq. (A4). Therefore, a state function $D(\mathbf{s})$ is defined in this section, which describes the expected instantaneous rate of change about degree-weighted frequency of opinion A in state $\mathbf{s}$, exactly corresponding to the differential in the integrand. $D(\mathbf{s})$ satisfies

$$E[\hat{S}(t+\varepsilon) - \hat{S}(t) \mid S(t) = \mathbf{s}] = D(\mathbf{s})\varepsilon + o(\epsilon) \quad (\varepsilon \to 0^+). \tag{A5}$$

Substituting Eq. (A5) into Eq. (A4), we obtain

$$\rho_{\mathbf{s}_0} = \hat{s}_0 + \int_0^\infty E_{\mathbf{s}_0}[D(\mathbf{S}(t))]\mathrm{d}t. \tag{A6}$$

Note that it is challenging to compute $\rho_{\mathbf{s}_0}$ exactly for arbitrary s_0. Therefore we here concentrate on the effects of weak selection on the fixation probability, meaning that $\beta \to 0$. In order to derive the fixation probability $\rho_{\mathbf{s}_0}$ under weak selection, i.e., the first order in β as $\beta \to 0^+$, we write the Taylor series expansion of $D(\mathbf{s})$ in β when $\beta \to 0$ as

$$\begin{aligned} D(\mathbf{s}) &= D^\circ(\mathbf{s}) + \beta D'(\mathbf{s}) + o(\beta^2) \\ &= \beta D'(\mathbf{s}) + o(\beta^2), \end{aligned} \tag{A7}$$

where $D^\circ(\mathbf{s})$ denotes the value of $D(\mathbf{s})$ under neutral drift ($\beta = 0$). Here, the superscript $^\circ$ is used to denote the case of neutral drift. We will show that $D^\circ(\mathbf{s}) = 0$ for all $\mathbf{s} \in \{0,1\}^G$ under the pairwise comparison updating in the following part.

The expansion of integrand in Eq. (A6) can be written as

$$\begin{aligned} E_{\mathbf{s}_0}[D(\mathbf{S}(t))] &= \sum_{\mathbf{s}} P_{\mathbf{s}_0}[\mathbf{S}(t) = \mathbf{s}]D(\mathbf{s}) \\ &= \beta \sum_{\mathbf{s}} P^\circ_{\mathbf{s}_0}[\mathbf{S}(t) = \mathbf{s}]D'(\mathbf{s}) + o(\beta^2) \\ &= \beta E^\circ_{\mathbf{s}_0}[D'(\mathbf{S}(t))] + o(\beta^2). \end{aligned} \tag{A8}$$

Substituting Eq. (A8) into Eq. (A6), we obtain the formula of fixation probability under weak selection as

$$\rho_{\mathbf{s}_0} = \hat{s}_0 + \beta \int_0^\infty E^\circ_{\mathbf{s}_0}[D'(\mathbf{S}(t)]\mathrm{d}t + o(\beta^2). \tag{A9}$$

For convenience, we define a new operator $\langle\rangle^\circ_{\mathbf{s}_0}$ to abbreviate the expression of fixed probability. Given any initial state $\mathbf{s}_0$, for any function $g(\mathbf{s})$ of state $\mathbf{s}$ we have

$$\langle g \rangle^\circ_{\mathbf{s}_0} = \int_0^\infty E^\circ_{\mathbf{s}_0}[g(\mathbf{S}(t))]\mathrm{d}t. \tag{A10}$$

Hence Eq. (A9) is abbreviated as

$$\rho_{\mathbf{s}_0} = \hat{s}_0 + \beta \langle D' \rangle^\circ_{\mathbf{s}_0} + o(\beta^2). \tag{A11}$$

In order to derive the formula of fixation probability $\rho_{\mathbf{s}_0}$, we have to calculate the expected instantaneous rate of the change in degree-weighted frequency $D(\mathbf{s})$ for state $\mathbf{s}$ in the weak selection limit. If the imitation event that individual i imitates the opinion of his/her neighbor j occurs, the degree-weighted frequency $\hat{s}$ will be changed by $\pi_i(s_j - s_i)$. Therefore, the expected instantaneous rate of degree-weighted frequency change from state $\mathbf{s}$ is given by

$$D(\mathbf{s}) =$$
$$\sum_i \pi_i \left(-s_i + \sum_j p_{ij} F(s_i, s_j) s_j + \left(1 - \sum_j p_{ij} F(s_i, s_j) \right) s_i \right)$$

$$= \sum_i \pi_i \left(\sum_j p_{ij} F(s_i, s_j) s_j - \left(\sum_j p_{ij} F(s_i, s_j) \right) s_i \right)$$

$$= \sum_i \pi_i \left(\sum_j p_{ij} F(s_i, s_j)(s_j - s_i) \right)$$

$$= \sum_i \pi_i \left(\sum_j p_{ij}(s_j - s_i) \right) \left(\frac{1}{2} - \frac{\beta}{4}(f_i(\mathbf{s}) - f_j(\mathbf{s})) \right) + o(\beta^2)$$

$$= \sum_i \sum_j \pi_i p_{ij} s_j \left(\frac{1}{2} - \frac{\beta}{4}(f_i(\mathbf{s}) - f_j(\mathbf{s})) \right)$$

$$- \sum_i \sum_j \pi_i p_{ij} s_i \left(\frac{1}{2} - \frac{\beta}{4}(f_i(\mathbf{s}) - f_j(\mathbf{s})) \right) + o(\beta^2)$$

$$= \frac{1}{2} \left(\sum_i \sum_j \pi_i p_{ij} s_j - \sum_i \sum_j \pi_i p_{ij} s_i \right)$$

$$+ \frac{\beta}{4} \sum_i \sum_j \pi_i p_{ij} s_i (f_i(\mathbf{s}) - f_j(\mathbf{s}))$$

$$- \frac{\beta}{4} \sum_i \sum_j \pi_i p_{ij} s_j (f_i(\mathbf{s}) - f_j(\mathbf{s})) + o(\beta^2)$$

$$= \frac{\beta}{4} \sum_i \sum_j \pi_i p_{ij} s_i (f_i(\mathbf{s}) - f_j(\mathbf{s}))$$

$$+ \frac{\beta}{4} \sum_i \sum_j \pi_j p_{ji} s_j (f_j(\mathbf{s}) - f_i(\mathbf{s})) + o(\beta^2)$$

$$= \frac{\beta}{2} \left(\sum_i \pi_i s_i \left(f_i(\mathbf{s}) - f_i^{(1)}(\mathbf{s}) \right) \right) + o(\beta^2) \tag{A12}$$

Eq. (A12) implies $D^\circ(\mathbf{s}) = 0$ for all states $\mathbf{s}$ and here the superscript $^\circ$ denotes the case of neutral drift $\beta = 0$. It can be seen that $D(\mathbf{s})$ is related to the total score of individual i and that of his/her one-step neighbor.

According to Eq. (A7), we have

$$D'(\mathbf{s}) = \frac{1}{2} \left(\sum_i \pi_i s_i (f_i(\mathbf{s}) - f_i^{(1)}(\mathbf{s})) \right). \tag{A13}$$

Since the expected total score of the one-step neighbor of individual i is given by

$$f_i^{(1)}(\mathbf{s}) = \sum_j p_{ij} f_j(\mathbf{s})$$

$$= W \sum_j p_{ij} \pi_j \left((a - b - c + d) s_j s_j^{(1)} + (b - d) s_j \right)$$

$$+ W \sum_j p_{ij} \pi_j \left((c - d) s_j^{(1)} + d \right)$$

$$+ \sum_j p_{ij} s_j \delta_A + \sum_j p_{ij} (1 - s_j) \delta_B, \tag{A14}$$

by substituting Eq. (A14) into Eq. (A13) we have

$$D'(\mathbf{s}) = \frac{1}{2} \left(\sum_i \pi_i s_i \left(f_i(\mathbf{s}) - f_i^{(1)}(\mathbf{s}) \right) \right)$$

$$= \frac{W}{2} (a - b - c + d) \left(\sum_i \pi_i^2 s_i^2 s_i^{(1)} - \sum_i \sum_j \pi_i \pi_j p_{ij} s_i s_j s_j^{(1)} \right)$$

$$+ \frac{W}{2} (b - d) \left(\sum_i \pi_i^2 s_i^2 - \sum_i \sum_j \pi_i \pi_j p_{ij} s_i s_j \right)$$

$$+ \frac{W}{2} (c - d) \left(\sum_i \pi_i^2 s_i s_i^{(1)} - \sum_i \sum_j \pi_i \pi_j p_{ij} s_i s_j^{(1)} \right)$$

$$+ \frac{Wd}{2} \left(\sum_i \pi_i^2 s_i - \sum_i \sum_j \pi_i \pi_j p_{ij} s_i \right)$$

$$+ \frac{(\delta_A - \delta_B)}{2} \left(\sum_i \pi_i s_i^2 - \sum_i \sum_j \pi_i p_{ij} s_i s_j \right). \tag{A15}$$

Since random walks have the reversibility property for each $i, j \in G$ and $\pi_i p_{ij} = \pi_j p_{ji}$, we have

$$D'(\mathbf{s}) = \frac{1}{2} \left(\sum_i \pi_i s_i \left(f_i(\mathbf{s}) - f_i^{(1)}(\mathbf{s}) \right) \right)$$

$$= \frac{W}{\cdot 2} (a - b - c + d) \left(\sum_i \pi_i^2 s_i^2 s_i^{(1)} - \sum_i \sum_j \pi_i^2 p_{ij} s_i s_j s_i^{(1)} \right)$$

$$+ \frac{W}{2} (b - d) \left(\sum_i \pi_i^2 s_i^2 - \sum_i \sum_j \pi_i^2 p_{ij} s_i s_j \right)$$

$$+ \frac{W}{2} (c - d) \left(\sum_i \pi_i^2 s_i s_i^{(1)} - \sum_i \sum_j \pi_i^2 p_{ij} s_j s_i^{(1)} \right)$$

$$+ \frac{Wd}{2} \left(\sum_i \pi_i^2 s_i - \sum_i \sum_j \pi_i^2 p_{ij} s_j \right)$$

$$+ \frac{(\delta_A - \delta_B)}{2} \left(\sum_i \pi_i s_i^2 - \sum_i \sum_j \pi_i p_{ij} s_i s_j \right), \tag{A16}$$

where $s_i^{(1)} = \sum_k p_{ik} s_k$ denotes the expected type of the neighbor of individual i. Hence

$$D'(\mathbf{s}) = \frac{1}{2} \left(\sum_{i,j,k} \pi_i s_i \left(f_i(\mathbf{s}) - f_i^{(1)}(\mathbf{s}) \right) \right)$$

$$= \frac{W}{2} (a - b - c + d) \left(\sum_{i,j,k} \pi_i^2 p_{ij} p_{ik} \left(s_i^2 s_k - s_i s_j s_k \right) \right)$$

$$+\frac{W}{2}(b-d)\left(\sum_{i,j}\pi_i^2 p_{ij}\left(s_i^2-s_is_j\right)\right)$$

$$+\frac{W}{2}(c-d)\left(\sum_{i,j,k}\pi_i^2 p_{ij}p_{ik}\left(s_is_k-s_js_k\right)\right)$$

$$+\frac{Wd}{2}\left(\sum_{i,j}\pi_i^2 p_{ij}\left(s_i-s_j\right)\right)$$

$$+\frac{(\delta_A-\delta_B)}{2}\left(\sum_{i,j}\pi_i p_{ij}\left(s_i^2-s_is_j\right)\right). \tag{A17}$$

Accordingly, the fixation probability of A is given by

$$\rho_{\mathbf{s}_0}=\hat{s}_0+\beta\langle D'\rangle^\circ_{\mathbf{s}_0}+o(\beta^2). \tag{A18}$$

We have obtained $\rho_{\mathbf{s}_0}$ for any initial state $\mathbf{s}_0$. Here we concentrate on the initial state where there is only one individual choosing A in the population. We mainly compute the fixation probability for such a special initial state $\mathbf{s}_0$ in which $s_i=1$ and $s_j=0$ for all $j\neq i$. Let $\mathbf{u}$ be the probability distribution over all states $\mathbf{s}$, assigning probability $\frac{1}{N}$ to states in which there is only one vertex with opinion A, and probability zero to all other states. Hence, when the initial state $\mathbf{s}_0$ of the *evolutionary Markov chain* is sampled from $\mathbf{u}$, the fixation probability of opinion A is given as

$$\rho_A=\frac{1}{N}+\beta\langle D'\rangle^\circ_{\mathbf{u}}+o(\beta^2), \tag{A19}$$

where

$$\begin{aligned}&\langle D'(\mathbf{s})\rangle^\circ_{\mathbf{u}}\\&=\frac{W}{2}(a-b-c+d)\left(\sum_{i,j,k}\pi_i^2 p_{ij}p_{ik}\left\langle s_i^2s_k-s_is_js_k\right\rangle^\circ_{\mathbf{u}}\right)\\&+\frac{W}{2}(b-d)\left(\sum_{i,j}\pi_i^2 p_{ij}\left\langle s_i^2-s_is_j\right\rangle^\circ_{\mathbf{u}}\right)\\&+\frac{W}{2}(c-d)\left(\sum_{i,j,k}\pi_i^2 p_{ij}p_{ik}\left\langle s_is_k-s_js_k\right\rangle^\circ_{\mathbf{u}}\right)\\&+\frac{Wd}{2}\left(\sum_{i,j}\pi_i^2 p_{ij}\left\langle s_i-s_j\right\rangle^\circ_{\mathbf{u}}\right)\\&+\frac{(\delta_A-\delta_B)}{2}\left(\sum_{i,j}\pi_i p_{ij}\left\langle s_i^2-s_is_j\right\rangle^\circ_{\mathbf{u}}\right).\end{aligned} \tag{A20}$$

Here

$$\langle s_i\rangle^\circ_{\mathbf{u}}=\int_0^\infty E^\circ_{\mathbf{u}}[S_i(t)]\mathrm{d}t, \tag{A21}$$

$$\langle s_is_j\rangle^\circ_{\mathbf{u}}=\int_0^\infty E^\circ_{\mathbf{u}}[S_i(t)S_j(t)]\mathrm{d}t, \tag{A22}$$

and

$$\langle s_is_js_k\rangle^\circ_{\mathbf{u}}=\int_0^\infty E^\circ_{\mathbf{u}}[S_i(t)S_j(t)S_k(t)]\mathrm{d}t, \tag{A23}$$

where the subscript $\mathbf{u}$ is used to denote the expectation of a quantity, when the initial state of the *evolutionary Markov chain* is sampled from $\mathbf{u}$.

ACKNOWLEDGMENTS

This research was supported by the National Natural Science Foundation of China (Grant Nos. 61976048, 62036002, and 61773121) and the Fundamental Research Funds of the Central Universities of China.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

[1] D. Acemoglu and A. Ozdaglar, *Dyn. Games Appl.* **1**, 3-49 (2011).
[2] E. Yildiz, A. Ozdaglar, D. Acemoglu, et al., *ACM Trans. Econ. Comput.* **1**, 1-30 (2013).
[3] A. V. Proskurnikov and R. Tempo, *Annu. Rev. Control* **43**, 65-79 (2017).
[4] X. Lin, Q. Jiao, and L. Wang, *IEEE Trans. Autom. Control* **64**, 3431-3438 (2018).
[5] A. V. Proskurnikov, A. S. Matveev, and M. Cao, *IEEE Trans. Autom. Control* **61**, 1524-1536 (2015).
[6] J. Ghaderi and R. Srikant, *Automatica* **50**, 3209-3215 (2014).
[7] F. Battiston, G. Cencetti, I. Iacopini, et al., *Phys. Rep.* **874**, 1-92 (2020).
[8] M. H. DeGroot, *J. Am. Stat. Assoc.* **69**, 118-121 (1974).
[9] N. E. Friedkin and E. C. Johnsen, *J. Math. Sociol.* **15**, 193-206 (1990).

[10] C. Castellano, S. Fortunato, and V. Loreto, *Rev. Mod. Phys.* **81**, 591-646 (2009).
[11] G. Weisbuch, *Eur. Phys. J. B* **38**, 339-343 (2004).
[12] R. G. Holcombe, *Public Choice* **61**, 115-125 (1989).
[13] S. Galam, *Eur. Phys. J. B* **25**, 403-406 (2002).
[14] K. Sznajd-Weron and J. Sznajd, *Int. J. Mod. Phys. C* **11**, 1157-1165 (2000).
[15] G. Deffuant, D. Neau, F. Amblard, and G. Weisbuch, *Adv. Complex Syst.* **3**, 87-98 (2000).
[16] R. Hegselmann and U. Krause, *J. Artif. Soc. Soc. Simul.* **5**, 1-33 (2002).
[17] R. Hegselmann and U. Krause, *Comput. Econ.* **25**, 381-405 (2005).
[18] C. Jiang, Y. Chen, and K. J. R. Liu, *IEEE Tran. Signal Proces.* **62**, 4573-4586 (2014).
[19] H. Zhang, Y. Li, Y. Chen, and H. V. Zhao, *IEEE Tran. Info. Fore. Secu.* **16**, 1203-1217 (2021).
[20] A. Di Mare and V. Latora, *Int. J. Mod. Phys. C* **18**, 1377-1395 (2007).
[21] F. Ding, Y. Liu, and Y. Li, *Int. J. Mod. Phys. C* **20**, 479-490 (2009).
[22] F. Ding, Y. Liu, B. Shen, and X. M. Si, *Phys. A* **389**, 1745-1752 (2010).
[23] H.-X. Yang, *EPL* **115**, 40007 (2016).
[24] J. L. Zhou, C. W. Huang, and Q. L. Dai, *EPL* **123**, 30004 (2018).
[25] L. Zino, M. Ye, and M. Cao, 59th IEEE Conference on Decision and Control, 1110-1115 (2020).
[26] B. Wu, J. Du, and L. Wang, *Proceedings of the 39th Chinese Control Conference*, 6707-6714 (2020).
[27] A. Traulsen, J. M. Pacheco, and M. A. Nowak, *J. Theor. Biol.* **246**, 522-529 (2007).
[28] M. Rabin, *Am. Econ. Rev.* **83**, 1281-1302 (1993).
[29] L. Cao and X. Li, *Phys. Rev. E* **77**, 016108 (2008).
[30] A. Szolnoki and M. Perc, *J. R. Soc. Interface* **12**, 20141299 (2015).
[31] A. Szolnoki and X. Chen, *New J. Phys.* **20**, 093008 (2018).
[32] J. W. Weibull, Evolutionary game theory, (MIT Press, 1997).
[33] J. Hofbauer and K. Sigmund, Evolutionary games and population dynamics, (Cambridge University Press, 1998).
[34] K. Sigmund and M. A. Nowak, *Curr. Biol.* **9**, 503-505 (1999).
[35] S. Tadelis, Game theory: an introduction, (Princeton University Press, 2013).
[36] H. G. Peng, X. K. Wang, T. L. Wang, and J. Q. Wang, *Appl. Soft. Comput.* **74**, 451-465 (2019).
[37] J. F. C. Kingman, *Stochastic Process. Appl.* **13**, 235-248 (1982).
[38] J. Wakeley, Coalescent theory: an introduction, (Roberts and Company Publishers, 2009).
[39] R. Oliveira, *Trans. Am. Math. Soc.* **364**, 2109-2128 (2012).
[40] B. Allen, G. Lippner, Y. T. Chen, B. Fotouhi, N. Momeni, S. T. Yau, and M. A. Nowak, *Nature* **544**, 227-230 (2017).
[41] G. Szabó and C. Tőke, *Phys. Rev. E* **58**, 69 (1998).
[42] H. Ohtsuki, C. Hauert, E. Lieberman, and M. A. Nowak, *Nature* **441**, 502-505 (2006).
[43] H. Ohtsuki and M. A. Nowak, *J. Theor. Biol.* **243**, 86-97 (2006).
[44] H. Ohtsuki and M. A. Nowak, *J. Theor. Biol.* **251**, 698-707 (2008).
[45] T. Konno, *J. Theor. Biol.* **269**, 224-233 (2011).
[46] C. Sample and B. Allen, *J. Math. Biol.* **75**, 1285-1317 (2017).
[47] Y. T. Chen, *Ann. Appl. Probab.* **23**, 637-664 (2013).
[48] F. C. Santos, J. M. Pacheco, and T. Lenaerts, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 3490-3494 (2006).
[49] D. Fudenberg, M. A. Nowak, C. Taylor, and L. A. Imhof, *Theor. Popul. Biol.* **70**, 352-363 (2006).
[50] F. C. Santos, M. D. Santos, and J. M. Pacheco, *Nature* **454**, 213-216 (2008).
[51] F. C. Santos, F. L. Pinheiro, T. Lenaerts, and J. M. Pacheco, *J. Theor. Biol.* **299**, 88-96 (2012).
[52] B. Allen and C. E. Tarnita, *J. Math. Biol.* **68**, 109-143 (2014).
[53] J. T. Cox, *Ann. Appl. Probab.* **17**, 1333-1366 (1989).
[54] M. E. Newman and D. J. Watts, *Phys. Lett. A* **263**, 341-346 (1999).
[55] A. L. Barabási and R. Albert, *Science* **286**, 509-512 (1999).
[56] S. Auer, J. Heitzig, U. Kornek, E. Schöll, and J. Kurths, *Sci. Rep.* **5**, 13386 (2015).
[57] T. Evans and F. Fu, *R. Soc. Open Sci.* **5**, 181122 (2018).
[58] G. F. de Arruda, F. A. Rodrigues, and Y. Moreno, *Phys. Rep.* **756**, 2-59 (2018).
[59] M. A. Amaral, W. G. Dantas, and J. J. Arenzon, *Phys. Rev. E* **101**, 062418 (2020).
[60] M. C. Gimenez, L. Reinaudi, A. P. Paz-García, P. M. Centres, and A. J. Ramirez-Pastor, *Eur. Phys. J. B* **94**, 1-11 (2021).
[61] L. Hindersin and A. Traulsen, *J. R. Soc. Interface* **11**, 20140606 (2014).
[62] M. Askari and K. A. Samani, *Phys. Rev. E* **92**, 042707 (2015).
[63] S. Farhang-Sardroodi, A. H. Darooneh, M. Nikbakht, N. L. Komarova, and M. Kohandel, *PLoS Comp. Biol.* **13**, e1005864 (2017).
[64] D. Hathcock and S. H. Strogatz, *Phys. Rev. E* **100**, 012408 (2019).
[65] M. Möller, L. Hindersin, and A. Traulsen, *Commun. Biol.* **2**, 1-9 (2019).
[66] J. Tkadlec, A. Pavlogiannis, K. Chatterjee, and M. A. Nowak, *Commun. Biol.* **2**, 1-8 (2019).
[67] S. Lessard and V. Ladret, *J. Math. Biol.* **54**, 721-744 (2007).
[68] W. Maciejewski, *J. Theor. Biol.* **340**, 285-293 (2014).
[69] C. E. Tarnita and P. D. Taylor, *Am. Nat.* **184**, 477-488 (2014).
[70] A. McAvoy and B. Allen, *J. Math. Biol.* **82**, 1-41 (2021).