

Research Article

Contagion Processes on Time-Varying Networks with Homophily-Driven Group Interactions

Li Ding  and Ping Hu

School of Electrical Engineering and Automation, Wuhan University, Wuhan 430072, China

Correspondence should be addressed to Li Ding; liding@whu.edu.cn

Received 11 June 2019; Revised 11 September 2019; Accepted 27 September 2019; Published 30 October 2019

Academic Editor: Átila Bueno

Copyright © 2019 Li Ding and Ping Hu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The complicated interaction patterns among heterogeneous individuals have a profound impact on the contagion process in the networks. In recent years, there has been increasing evidence for the emergence of many-body interactions between two or more nodes in a wide range of biological and social networks. To encode these multinode interactions explicitly, the simplicial complex is now a popular alternative to simple networks. Meanwhile, the time-varying network has been acknowledged as a key ingredient of the contagion process. In this paper, we consider the connectivity pattern of networks affected by the homophily effect associated with individual attributes and investigate the impact of homophily-driven group interactions on the contagion process in temporal networks. The simplicial complex modeling framework is adopted to capture stochastic interactions between passively selected nodes in the paradigm of activity-driven networks. We study the evolution of infection and the epidemic threshold of the contagion process by both analytical and numerical methods. Our results on statistical topological properties of instantaneous network may shed light on accurately characterizing the evolution curve of infection. Furthermore, we show the impact of the homophily-driven interaction pattern on the epidemic threshold, which generalizes the existing results on both the paradigmatic activity-driven network and the simplicial activity-driven network.

1. Introduction

Network modeling plays a critical role in identifying structural properties and analyzing contagion processes on networks, such as the spreading of epidemic and malware, as well as the diffusion of news and ideas. Meanwhile, it has been acknowledged that the intrinsic characteristic differences among individuals and the complicated connectivity patterns are two main important factors determining the properties of the contagion process [1–3]. For instance, in a social network, individuals are distinguished by their attributes such as gender and age. These, together with biased interaction patterns, induce heterogeneous rates of adoption of an idea [4]. Consequently, there are substantial works dedicated to the comprehension of the typical behavioral characteristics and distinctive connectivity features, such as the homophily [4, 5], reinforcement effect [6, 7], heterogeneous degree distribution [8–10], and community structures [11, 12].

Generally, there is a tacit hypothesis in network modeling approaches that the structure of complex systems is reducible to the pairwise interactions (links) of their entities (nodes) [13]. In recent years, the research community has accumulated overwhelming evidence for the emergence of many-body interaction patterns in a vast body of biological and social systems. For instance, the rich datasets available have revealed that interactions occur between more than two nodes in many systems [14]. These complicated interactions cannot be encoded in the simple network and thus prompt the need for characterizing generalized network structures. Simplicial complexes are a generalization of networks that describe interactions between more than two nodes. They can involve any number of nodes. For instance, simplices of dimension $d = 0, 1, 2,$ and 3 are nodes, links, triangles, and tetrahedra, respectively, while d -dimensional simplices are their d -dimensional generalizations. A simplicial complex is a finite collection of simplices of different dimensions

properly glued together. It has recently been adopted in modeling many complex interacting systems, such as the brain [15], biological protein-interaction systems [16], and social systems [17].

In many cases, the interactions among system entities are rapidly changing and evolving over time. The understanding of time-varying networks and their impacts on the dynamical process has been a long-standing area of research. Recently, an activity-driven (AD) modeling framework has been put forward to address the time-scale separation between the evolving network structure and the dynamical process [18]. In this framework, each node is characterized by a specific activity rate. At each time step, a node may become active and generate finite connections with other nodes randomly selected from the network. The simple formulation of AD networks and its extensions is amenable to analytical treatment and thus facilitates the study of contagion processes on time-varying networks. Examples include epidemic spreading processes with different network structures [19, 20], information diffusion processes with complex adoption models [21, 22], and the interplay between the epidemic spreading and awareness diffusion with different risk perception mechanisms [23–25].

However, the above studies focused on the pairwise interactions between individuals, while the evolving simplicial interactions are rarely considered. This gap is bridged by the simplicial activity-driven (SAD) model proposed in the recent work [26]. Within this framework, a multiagent interaction is depicted by a simplex of nodes, which implies the interaction between any two nodes in a group created by an active node. Although this simplex description is accurate for some systems such as scientific collaboration networks, a more general group interaction may contain different interaction patterns among individuals. Take the micro-blogging network as an example, when a user actively posts a text and interacts with its followers, two or more of its followers may also be connected because they share common interests. Thus, the potential role of possible interactions among nodes connected to a common active node in the epidemic spreading process deserves further discussions.

As mentioned above, the intrinsic characteristic of an individual is critical for shaping the individual interaction pattern. Inspired by the existence of temporal homophily revealed in the temporal motifs of communication networks [27], we consider the case in which the interaction patterns are affected by the individual attribute. We assume that there are two types of nodes in the network depending on their attributes. At each step, an active node randomly chooses its neighbors, and these neighbor nodes would interact at different probabilities determined by their types. In particular, nodes tend to interact with others of the same type. That is to say, the formation of group interactions is homophily-driven. This model extends the paradigms of typical AD models and the SAD model, which either ignore any interaction among neighbor nodes or assume complete interactions among neighbor nodes. We refer to this model as the complex activity-driven (CAD) model since the group interactions facilitated by active nodes are now described by a simplicial complex.

In this paper, the effects of homophily-driven group interactions on the contagion processes in time-varying networks are investigated. We consider the susceptible-infected-susceptible (SIS) model for the epidemic spreading process and provide analytical analysis for both the evolution process and the epidemic threshold. In majority of existing studies, the characterization of both the evolution and the steady state of infection relies on extensive Monte Carlo simulations. Once the epidemics outbreaks, the spreading models based on the classical heterogeneous mean-field (HMF) method can only provide an upper bound for the infection level. Here, we propose a new model by developing the classical HMF method to more accurately predict the evolution of the epidemic curve. To this end, we explore the topological properties of the instantaneous network generated at each step and obtain the degree distribution and the connection correlations among nodes. It is shown that introducing these statistical information of the instantaneous network contributes to accurately characterize the evolution of epidemic. Then, we derive the epidemic threshold with the analytical statistical model. The threshold is related to the connectivity pattern determined by the node's activity potential and the homophily effect, as well as the proportion of two types of nodes. Our analytical results can cover and generalize the existing results in both classical AD and SAD networks.

The rest of this paper is constructed as follows. In Section 2, the basic CAD model is established and the topological properties of the instantaneous network are analyzed. In Section 3, the statistical SIS epidemic model is formulated and the epidemic threshold is discussed. In Section 4, numerical simulations for the epidemic spreading process are presented. Conclusions are drawn in Section 5.

2. Model

This section introduces the CAD model and investigates the topological properties of the instantaneous CAD network.

2.1. Complex Activity-Driven Network. The complex activity-driven network is an extension of the paradigmatic activity-driven network that incorporates group interactions at each step. In the framework of AD models, each node (individual) i in a population of N nodes is endowed with an activity rate a_i , which is taken from a predefined activity probability distribution $F(a)$. a_i is defined as the probability per unit time of node i to create interactions with other nodes. In addition to the heterogeneous activity rates, individuals in the population are distinguished by their attributes. We assume that there are two types of nodes, A and B . The proportions of type A nodes and type B nodes in the network are p and $1 - p$, respectively. For simplicity, we assume that the activity rates of nodes are irrelevant to their types, and thus the activity probability distributions for both type A and B nodes are expected to be $F(a)$. Similar to the approach for static networks, we group nodes into different classes according to their types and activities. Nodes with the same type and the same activity belong to the same class and are

considered to be statistically equivalent. We denote the class \mathcal{N}_a^A as the collection of all nodes of type A with the activity rate a . Other classes are denoted similarly.

The instantaneous network G_t of the CAD model is generated as follows: At each time step t , G_t starts with N disconnected nodes. With probability $a_i \Delta t$, a node i becomes active and generates connection with m other nodes chosen randomly (as shown in Figure 1). All the selected nodes are neighbors of i , which together with the node i , form the group i . There are two different roles in this group: *center* and *neighbor*, which are played by the active node i and all other nodes connected by i , respectively. The *neighbor* nodes may interact with each other, and the probabilities for an interaction between two *neighbor* nodes with the same type and different types are λ_1 and λ_2 , respectively. Since the interactions between *neighbor* nodes are stochastic, there may be different interaction patterns in one group, such as pairwise interactions and interactions involving more than two nodes. With one *center* node and m *neighbor* nodes in each group, the dimension of group interactions may vary from 1 to m . This is different from the SAD model that considers an m -simplex in each group. The instantaneous network G_t is then modeled as a simplicial complex composed of lower dimension simplicial complexes in each group. All these simplicial complexes are deleted at the next time step $t + \Delta t$. Without loss of generality, we set $\Delta t = 1$ in the following.

2.2. Topological Properties of the Instantaneous Network.

To quantitatively characterize the epidemic spreading process, we analyze the degree distribution and the connection correlation for nodes belonging to different classes in the instantaneous network. Since G_t evolves with time, the topological properties of G_t are analyzed in a statistical way. Based on the definition of m , a_i , and $F(a)$, it is expected that there are $N\langle a \rangle$ active nodes with $N\langle a \rangle$ groups in G_t , where $\langle a \rangle = \sum_a a F(a)$. In each group, there are at most $(m(m-1))/2$ edges among m *neighbor* nodes. Since these m *neighbor* nodes are chosen randomly by the *center* node, an interaction between any two *neighbor* nodes occurs with the average probability $\lambda_1(p^2 + (1-p)^2) + 2\lambda_2 p(1-p)$. Thus, it is expected that there are E_g edges in each group, with

$$E_g = m \left[1 + \frac{\lambda_1(p^2 + (1-p)^2) + 2\lambda_2 p(1-p)}{2} (m-1) \right]. \quad (1)$$

From the perspective of directed networks, the number of edges existing in G_t is expected to be $E = 2N\langle a \rangle E_g$.

Next, we discuss the degree distribution of nodes belonging to class \mathcal{N}_a^A in the instantaneous network. At each step, the inactive nodes may be passively connected by the active nodes. The probability $P_{A,a}^{\text{inact}}(k)$ that an inactive node $i \in \mathcal{N}_a^A$ has degree k is

$$P_{A,a}^{\text{inact}}(k) = \sum_{k_1=\lceil k/m \rceil}^k P_a(k_1) P_g(k-k_1), \quad (2)$$

where $\lceil x \rceil$ denotes the minimum integer that is no less than x , $P_a(k_1)$ denotes the probability that the node i is connected by k_1 active nodes in G_t , and $P_g(k-k_1)$ denotes the probability of $k-k_1$ interactions between the node i and other *neighbor* nodes. The expression of $P_a(k_1)$ is the same as that for AD networks [24], which is given by

$$P_a(k_1) = \binom{m\langle a \rangle N}{k_1} \left(\frac{1}{N}\right)^{k_1} \left(1 - \frac{1}{N}\right)^{N-k_1}. \quad (3)$$

In the thermodynamic limit, i.e., $N \rightarrow \infty$, $P_a(k_1)$ can be rewritten as

$$P_a(k_1) = \frac{(m\langle a \rangle)^{k_1}}{k_1!} e^{-m\langle a \rangle}. \quad (4)$$

In addition, the expression of $P_g(k-k_1)$ is

$$P_g(k-k_1) = \binom{k_1(m-1)}{k-k_1} \lambda_A^{k-k_1} (1-\lambda_A)^{k_1 m - k}, \quad (5)$$

with $\lambda_A = \lambda_1 p + \lambda_2(1-p)$ denoting the average probability that the node i of type A interacts with a *neighbor* node in the group.

In equation (2), $\lceil k/m \rceil$ implies that the inactive node i with degree k must be connected by at least $\lceil k/m \rceil$ active nodes. This is due to the fact that a node can have at most m connections in one group. Once the node i is connected by $k_1 \geq \lceil k/m \rceil$ active nodes, there are k_1 edges between the node i and the k_1 *center* nodes in the network. Accordingly, the rest of $k-k_1$ edges of the node i are created by group interactions with *neighbor* nodes. Since the proportions of the nodes of type A and B are, respectively, p and $1-p$, the average probability for an interaction between the node i and a *neighbor* node in the group is expected to be λ_A . Thus, the $k-k_1$ interactions between node i and all other $k_1(m-1)$ *neighbor* nodes included in k_1 groups occur with the probability $P_g(k-k_1)$.

For each active node with degree k , there must be m edges emitted from it. Hence, other $k-m$ edges are generated by other active nodes or through group interactions. Similar to the inactive case, the probability $P_{A,a}^{\text{act}}(k)$ that an active node $i \in \mathcal{N}_a^A$ has degree k is

$$P_{A,a}^{\text{act}}(k) = \begin{cases} 0, & k < m, \\ \sum_{k_1=\lceil (k-m)/m \rceil}^k P_a(k_1) P_g(k-m-k_1), & k \geq m. \end{cases} \quad (6)$$

Combining equations (2) and (6), the instantaneous degree distribution $P_a^A(k)$ of nodes in class \mathcal{N}_a^A is given by

$$P_a^A(k) = a P_{A,a}^{\text{act}}(k) + (1-a) P_{A,a}^{\text{inact}}(k). \quad (7)$$

The analysis for the nodes of type B is analogous to that of type A . For the nodes of type B with activity rate a in class \mathcal{N}_a^B , we can easily derive the instantaneous degree distributions $P_a^B(k)$ as

$$P_a^B(k) = a P_{B,a}^{\text{act}}(k) + (1-a) P_{B,a}^{\text{inact}}(k), \quad (8)$$

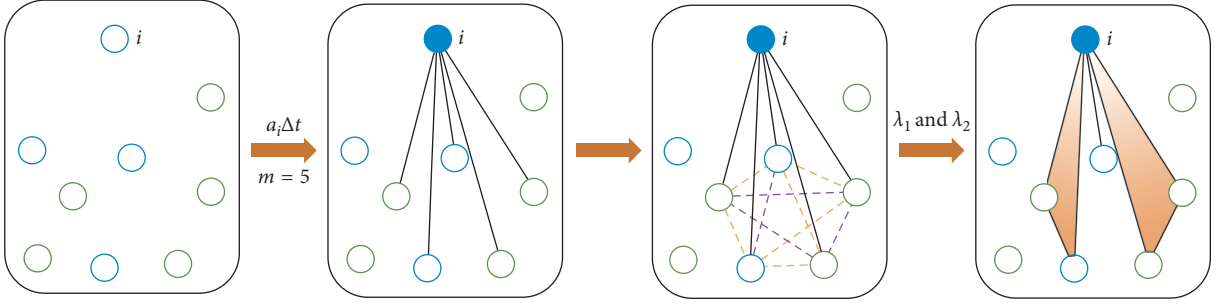


FIGURE 1: The CAD model. In the network, there are two types of nodes, with the blue (green) circles representing type A (B). At each time step, a type A node i becomes active (labeled with solid circle) with the probability $a_i\Delta t$. Once the node i is active, it interacts with m neighbor nodes randomly chosen from all the other nodes in the network. The purple (orange) dotted line denotes the possible interaction between two neighbor nodes with the same type (different types). With the given group interaction probabilities, a simplicial complex with one link and two triangles is generated at this step.

with λ_A in $P_{A,a}^{\text{act}}(k)$ and $P_{A,a}^{\text{inact}}(k)$ replaced by $\lambda_B = \lambda_2 p + \lambda_1(1-p)$ in $P_{B,a}^{\text{act}}(k)$ and $P_{B,a}^{\text{inact}}(k)$.

In Figure 2, we compare the statistical results (7) and (8) with those obtained by Monte Carlo simulations. Due to the heterogeneous activity rates in the real world, we consider different activity rates among individuals in the network. The activity probability distribution is chosen as $F(a) \propto a^{-2}$. For simplicity, we assume that there are n different activity rates in the network and they are uniformly sampled in $(0, 1]$ [24]. The sampling interval is $1/N$, and the minimum activity rate in the network is denoted as a_1 . For the example of $n = 5$, the activity rate is selected from $a \in \{0.2, 0.4, 0.6, 0.8, 1\}$ with $a_1 = 0.2$. In this way, the heterogeneity of node activity rates is adjustable by the value of n . The degree distribution of the nodes with activity rate a_1 is taken as an illustration example and presented in Figure 2. It is shown that the theoretical results of both $P_{a_1}^A(k)$ and $P_{a_1}^B(k)$ are consistent with the simulation results. To further measure the error between the theoretical and simulation results, we define an evaluation parameter, i.e.,

$$\sigma_n = \sum_k \sum_{i=1}^n \left(\left| P_{a_i}^A(k) - \hat{P}_{a_i}^A(k) \right| + \left| P_{a_i}^B(k) - \hat{P}_{a_i}^B(k) \right| \right), \quad (9)$$

with $P_{a_i}^A(k)$ and $P_{a_i}^B(k)$ from equations (7) and (8) and their counterparts $\hat{P}_{a_i}^A(k)$ and $\hat{P}_{a_i}^B(k)$ obtained by simulations. According to the value of σ_n in Figure 2, the absolute error between theoretical and simulation results is very small. This error is reasonable due to the finite network size and independent realizations.

In addition to the degree distribution, we need to analyze the connection correlation among different node classes in the instantaneous network G_t . The connection correlation describes a preference for nodes to interact with others in terms of their class. It is similar to the degree-degree correlation in the static network [8] and the activity-activity correlation in AD networks [24]. Here, we refer to it as the class-class correlation since both the type and the activity rate of a node are considered in each class.

To this end, we define $F(\mathcal{N}_a^A, \mathcal{N}_{a'}^A)$ as the probability of finding a node in class \mathcal{N}_a^A and another node in class $\mathcal{N}_{a'}^A$ at two ends of an edge randomly chosen in G_t . Also, $F_n(\mathcal{N}_a^A)$ is

defined as the probability that there is a node belonging to the class \mathcal{N}_a^A at the end of a randomly selected edge in G_t . Obviously, $F(\mathcal{N}_a^A, \mathcal{N}_{a'}^A)$ represents the ratio of the number of edges $E(\mathcal{N}_a^A, \mathcal{N}_{a'}^A)$ created by one node in class \mathcal{N}_a^A and the other node in $\mathcal{N}_{a'}^A$, to the total number of edges E in the instantaneous network G_t . According to the generating rules of G_t , we can obtain

$$\begin{aligned} E(\mathcal{N}_a^A, \mathcal{N}_{a'}^A) &= pNF(a)ma pF(a') + pNF(a')ma' pF(a) \\ &\quad + N\langle a \rangle m pF(a)(m-1)pF(a')\lambda_1 \\ &= p^2NmF(a)F(a')[a + a' + \lambda_1\langle a \rangle(m-1)]. \end{aligned} \quad (10)$$

The first (second) term corresponds to an active node in class a (a') that becomes active and connects to a node in class a' (a). The third term describes the case where both a node in class a and a node in class a' are chosen as neighbor nodes of an active node. Similarly, the number of edges $E(\mathcal{N}_a^A, \mathcal{N}_{a'}^B)$ related to classes \mathcal{N}_a^A and $\mathcal{N}_{a'}^B$ can be written as

$$E(\mathcal{N}_a^A, \mathcal{N}_{a'}^B) = p(1-p)NmF(a)F(a')[a + a' + \lambda_2\langle a \rangle(m-1)], \quad (11)$$

with equations (1), (10), and (11), and the explicit expressions of $F(\mathcal{N}_a^A, \mathcal{N}_{a'}^A) = E(\mathcal{N}_a^A, \mathcal{N}_{a'}^A)/E$ and $F(\mathcal{N}_a^A, \mathcal{N}_{a'}^B) = E(\mathcal{N}_a^A, \mathcal{N}_{a'}^B)/E$ can be easily obtained. The analysis of $F(\mathcal{N}_a^B, \mathcal{N}_{a'}^A)$ and $F(\mathcal{N}_a^B, \mathcal{N}_{a'}^B)$ is carried out in a similar way. Obviously, $F(\mathcal{N}_a^A, \mathcal{N}_{a'}^A) = F(\mathcal{N}_{a'}^A, \mathcal{N}_a^A)$, $F(\mathcal{N}_a^B, \mathcal{N}_{a'}^A) = F(\mathcal{N}_{a'}^B, \mathcal{N}_a^A) = F(\mathcal{N}_a^A, \mathcal{N}_{a'}^B) = F(\mathcal{N}_{a'}^A, \mathcal{N}_a^B)$, $F(\mathcal{N}_a^B, \mathcal{N}_{a'}^B) = F(\mathcal{N}_{a'}^B, \mathcal{N}_a^B)$, $\forall a, a'$. In addition, the sum rule of $\sum_{a,a'} F(\mathcal{N}_a^A, \mathcal{N}_{a'}^A) + 2F(\mathcal{N}_a^A, \mathcal{N}_{a'}^B) + F(\mathcal{N}_a^B, \mathcal{N}_{a'}^B) = 1$ is satisfied.

According to the definition of $F_n(\mathcal{N}_a^A)$, $F_n(\mathcal{N}_a^A)$ equals the ratio of the number of edges $E(\mathcal{N}_a^A)$ connected to the nodes in class \mathcal{N}_a^A to the total number of edges E in the instantaneous network G_t . That is,

$$\begin{aligned} E(\mathcal{N}_a^A) &= pNF(a)ma + pNF(a)m\langle a \rangle \\ &\quad \cdot [1 + (m-1)(\lambda_1 p + \lambda_2(1-p))] \\ &= pNmF(a)[a + \langle a \rangle + (m-1)\langle a \rangle(\lambda_1 p + \lambda_2(1-p))]. \end{aligned} \quad (12)$$

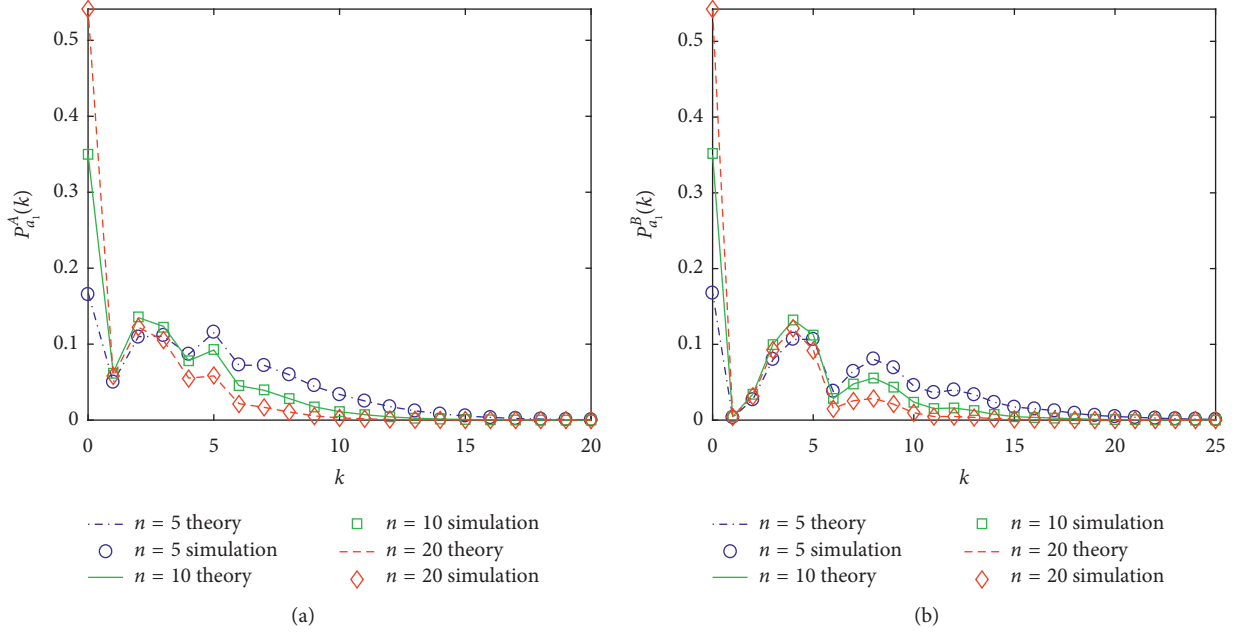


FIGURE 2: The degree distribution of the nodes with activity a_1 in the instantaneous network G_t . Node types in (a) and (b) are, respectively, A and B. The dash-dotted line (circle), solid line (square), and dotted line (diamond) represent the theoretical (simulation) results for $n = 5$, $n = 10$, and $n = 20$, respectively. For the three candidates of n , the values of σ_n are 0.0572, 0.1736, and 0.5779, respectively. Other parameters are set to be $N = 10,000$, $p = 0.3$, $m = 5$, $\lambda_1 = 0.9$, and $\lambda_2 = 0.1$. The simulation results are averaged over 1,000 realizations.

The first term corresponds to an active node in class \mathcal{N}_a^A that generates m edges connected to m other randomly selected nodes. The second term stems from the fact that an inactive node i in class \mathcal{N}_a^A is connected by an active node in G_t , and as the role of *neighbor* in the group, node i interacts with other *neighbor* nodes with certain probabilities. Moreover, the explicit expression of $F_n(\mathcal{N}_a^A)$ can be obtained with equations (1) and (12). Based on equations (10)–(12), it is easily to prove that the quantities $F(\mathcal{N}_a^A, \mathcal{N}_a^A)$, $F(\mathcal{N}_a^A, \mathcal{N}_a^B)$, and $F_n(\mathcal{N}_a^A)$ are satisfied with $F_n(\mathcal{N}_a^A) = \sum_{a'} F(\mathcal{N}_a^A, \mathcal{N}_{a'}^A) + F(\mathcal{N}_a^A, \mathcal{N}_{a'}^B)$.

We define $F(\mathcal{N}_{a'}^A | \mathcal{N}_a^A)$ as the conditional probability that a randomly selected edge of the nodes in class \mathcal{N}_a^A points to a node in the class $\mathcal{N}_{a'}^A$. Then, combining equations (10)–(12), $F(\mathcal{N}_{a'}^A | \mathcal{N}_a^A)$ can be written as

$$\begin{aligned} F(\mathcal{N}_{a'}^A | \mathcal{N}_a^A) &= \frac{F(\mathcal{N}_a^A, \mathcal{N}_{a'}^A)}{F_n(\mathcal{N}_a^A)} \\ &= \frac{pF(a')[a + a' + \lambda_1 \langle a \rangle (m - 1)]}{a + \langle a \rangle + (m - 1) \langle a \rangle (\lambda_1 p + \lambda_2 (1 - p))}, \end{aligned} \quad (13)$$

which indicates a connection correlation between the class \mathcal{N}_a^A and $\mathcal{N}_{a'}^A$ in the instantaneous network G_t . Following the analysis steps of equation (13), we can derive the following conditional probabilities for other three cases:

$$\begin{aligned} F(\mathcal{N}_{a'}^B | \mathcal{N}_a^A) &= \frac{(1 - p)F(a')[a + a' + \lambda_2 \langle a \rangle (m - 1)]}{a + \langle a \rangle + (m - 1) \langle a \rangle (\lambda_1 p + \lambda_2 (1 - p))}, \\ F(\mathcal{N}_{a'}^A | \mathcal{N}_a^B) &= \frac{pF(a')[a + a' + \lambda_2 \langle a \rangle (m - 1)]}{a + \langle a \rangle + (m - 1) \langle a \rangle (\lambda_2 p + \lambda_1 (1 - p))}, \\ F(\mathcal{N}_{a'}^B | \mathcal{N}_a^B) &= \frac{(1 - p)F(a')[a + a' + \lambda_1 \langle a \rangle (m - 1)]}{a + \langle a \rangle + (m - 1) \langle a \rangle (\lambda_2 p + \lambda_1 (1 - p))}. \end{aligned} \quad (14)$$

The combination of any two activity rates a and a' can be analyzed by taking similar steps (10)–(14):

In Figure 3, the conditional probabilities of nodes with the minimum activity rate a_1 are presented. Here, $F(\mathcal{N}_{a_1}^{A/B} | \mathcal{N}_{a_1}^A)$ ($F(\mathcal{N}_{a_1}^{A/B} | \mathcal{N}_{a_1}^B)$) denotes the conditional probability with which an edge of nodes in the class $\mathcal{N}_{a_1}^A$ ($\mathcal{N}_{a_1}^B$) is connected to a node in the class $\mathcal{N}_{a_1}^A$ or $\mathcal{N}_{a_1}^B$. It can be found that the simulation results match well with the theoretical results (13) and (14). Also, it is observed that the homophily-driven group interactions play a significant role in shaping the connection pattern of nodes. Particularly, as shown in Figure 3(a), nodes in class $\mathcal{N}_{a_1}^A$ tend to interact with nodes in class $\mathcal{N}_{a_1}^A$ rather than nodes in class $\mathcal{N}_{a_1}^B$, even though the latter is dominant in quantity. Similarly, we introduce an evaluation parameter ϵ_n to further measure the error between theoretical and simulation results. That is,

$$\begin{aligned} \varepsilon_n = \sum_{i=1}^n \sum_{j=1}^n & \left(\Delta F(\mathcal{N}_{a_j}^A | \mathcal{N}_{a_i}^A) + \Delta F(\mathcal{N}_{a_j}^B | \mathcal{N}_{a_i}^A) \right) \\ & + \Delta F(\mathcal{N}_{a_j}^A | \mathcal{N}_{a_i}^B) + \Delta F(\mathcal{N}_{a_j}^B | \mathcal{N}_{a_i}^B), \end{aligned} \quad (15)$$

with $\Delta F(\mathcal{N}_{a_j}^C | \mathcal{N}_{a_i}^D) = |F(\mathcal{N}_{a_j}^C | \mathcal{N}_{a_i}^D) - \widehat{F}(\mathcal{N}_{a_j}^C | \mathcal{N}_{a_i}^D)|$, $C, D = \{A, B\}$ and $F(\mathcal{N}_{a_j}^C | \mathcal{N}_{a_i}^D)$ from equations (13) and (14) and its counterpart $\widehat{F}(\mathcal{N}_{a_j}^C | \mathcal{N}_{a_i}^D)$ obtained by simulations. According to Figure 3, ε_n is small under the three candidates of n .

3. Analysis

In this section, we characterize the paradigmatic SIS model with the statistical properties of the instantaneous network obtained in Section 2. Based on the statistical SIS model, we discuss the epidemic threshold in the CAD networks.

3.1. Statistical SIS Model. In the SIS model, each node can be either susceptible (S) or infectious (I). A susceptible node in contact with an infectious node becomes infected with an infection rate β , while each infectious node recovers from infection with a recovery rate μ .

At the time step t , we denote the fractions of susceptible and infectious nodes in the class $\mathcal{N}_{a,a}^A$ as $S_{A,a}^t$ and $I_{A,a}^t$, respectively. Then, it gives $S_{A,a}^t + I_{A,a}^t = 1$. To analyze the contagion process quantitatively, we define $\omega_{A,a}^t$ as the probability that an edge of nodes in class $\mathcal{N}_{a,a}^A$ points to an infectious node in G_t . Obviously, $\omega_{A,a}^t$ is dependent on both the class-class connection correlation property of G_t and the infection level in each node class. Based on the analysis in Section 2, we have

$$\omega_{A,a}^t = \sum_{a'} F(\mathcal{N}_{a'}^A | \mathcal{N}_a^A) I_{A,a'}^t + F(\mathcal{N}_{a'}^B | \mathcal{N}_a^A) I_{B,a'}^t, \quad (16)$$

with $F(\mathcal{N}_{a'}^A | \mathcal{N}_a^A)$ and $F(\mathcal{N}_{a'}^B | \mathcal{N}_a^A)$ given by equations (13) and (14), respectively.

For a node i in the class \mathcal{N}_a^A , the probability of node i with degree k is given by the degree distribution $P_a^A(k)$. Based on equation (7), the fractions of susceptible and infectious nodes in the class \mathcal{N}_a^A at next time step $t+1$ are given by

$$\begin{cases} S_{A,a}^{t+1} = \mu I_{A,a}^t - S_{A,a}^t \sum_k \left[1 - (1 - \beta \omega_{A,a}^t)^k \right] P_a^A(k), \\ I_{A,a}^{t+1} = (1 - \mu) I_{A,a}^t + S_{A,a}^t \sum_k \left[1 - (1 - \beta \omega_{A,a}^t)^k \right] P_a^A(k). \end{cases} \quad (17)$$

It is noted that the statistical properties of the instantaneous network G_t , $P_a^A(k)$ and $F(\mathcal{N}_{a'}^A | \mathcal{N}_a^A)$, are included in (17). Since $P_a^A(k)$ and $F(\mathcal{N}_{a'}^A | \mathcal{N}_a^A)$ are related to the parameters p , λ_1 and λ_2 , the potential impacts of homophily-driven group interactions on the contagion process are reflected in the statistical properties of G_t . We refer to the SIS model described by equation (17) as our statistical model. This model is different from models

derived by the classical HMF theory in most of existing studies within the framework of AD models. Leveraging the HMF theory, the SIS model on the CAD networks is derived as

$$\begin{aligned} \widehat{I}_{A,a}^{t+1} &= (1 - \mu) \widehat{I}_{A,a}^t + \beta \widehat{S}_{A,a}^t m a \int \frac{\widehat{I}_{A,a'}^t + \widehat{I}_{B,a'}^t}{N} da' + \beta \widehat{S}_{A,a}^t \\ &\cdot \int \frac{\widehat{I}_{A,a'}^t + \widehat{I}_{B,a'}^t}{N} m a' da' + \beta \widehat{S}_{A,a}^t \int m a' da' \\ &\times \int \frac{\lambda_1 \widehat{I}_{A,a''}^t + \lambda_2 \widehat{I}_{B,a''}^t}{N} (m - 1) da'', \end{aligned} \quad (18)$$

where $\widehat{S}_{A,a}^t$ and $\widehat{I}_{A,a}^t$, respectively, denote the number of susceptible and infectious nodes in class \mathcal{N}_a^A at the time step t . With $\widehat{S}_{A,a}^t = p N F(a) S_{A,a}^t$ and $\widehat{I}_{A,a}^t = p N F(a) I_{A,a}^t$, equation (18) can be rewritten as

$$\begin{aligned} I_{A,a}^{t+1} &= (1 - \mu) I_{A,a}^t + \beta S_{A,a}^t m a (p I_A^t + (1 - p) I_B^t) \\ &+ \beta S_{A,a}^t m (p \theta_A^t + (1 - p) \theta_B^t) + \beta S_{A,a}^t m \langle a \rangle (m - 1) \\ &\cdot (\lambda_1 p I_A^t + \lambda_2 (1 - p) I_B^t), \end{aligned} \quad (19)$$

with $I_A^t = \sum_a F(a) I_{A,a}^t$, $I_B^t = \sum_a F(a) I_{B,a}^t$, $\theta_A^t = \sum_a a F(a) I_{A,a}^t$, and $\theta_B^t = \sum_a a F(a) I_{B,a}^t$. To distinguish with the statistical model, the model (19) is referred to as the classical MF model. This model is introduced as a comparison model to measure the performance of the proposed statistical model that contains more topological properties of the instantaneous network.

3.2. Epidemic Threshold. To obtain the epidemic threshold, we analyze the stability around $I_{A,a}^t \rightarrow 0$. In this case, we have the approximation $1 - (1 - \beta \omega_{A,a}^t)^k \approx k \beta \omega_{A,a}^t$. Then, it gives

$$\sum_k \left[1 - (1 - \beta \omega_{A,a}^t)^k \right] P_a^A(k) = \sum_k k \beta \omega_{A,a}^t P_a^A(k) = \frac{E(\mathcal{N}_a^A)}{p N F(a)} \beta \omega_{A,a}^t. \quad (20)$$

By introducing equations (13) and (14) into $\omega_{A,a}^t$, the fraction of infectious nodes $I_{A,a}^{t+1}$ in equation (17) can be rewritten as

$$\begin{aligned} I_{A,a}^{t+1} &= (1 - \mu) I_{A,a}^t + \beta S_{A,a}^t m \left[p I_A^t (a + \lambda_1 \langle a \rangle (m - 1)) + (1 - p) I_B^t \right. \\ &\left. (a + \lambda_2 \langle a \rangle (m - 1)) + (p \theta_A^t + (1 - p) \theta_B^t) \right], \end{aligned} \quad (21)$$

which is exactly equation (19) derived by the classical HMF theory. This fact then implies that the classical MF model (19) is an approximation of the statistical model (17). Meanwhile, the result of approximation is accurate only when the infection level is very low. This will be further discussed in the following numerical simulations.

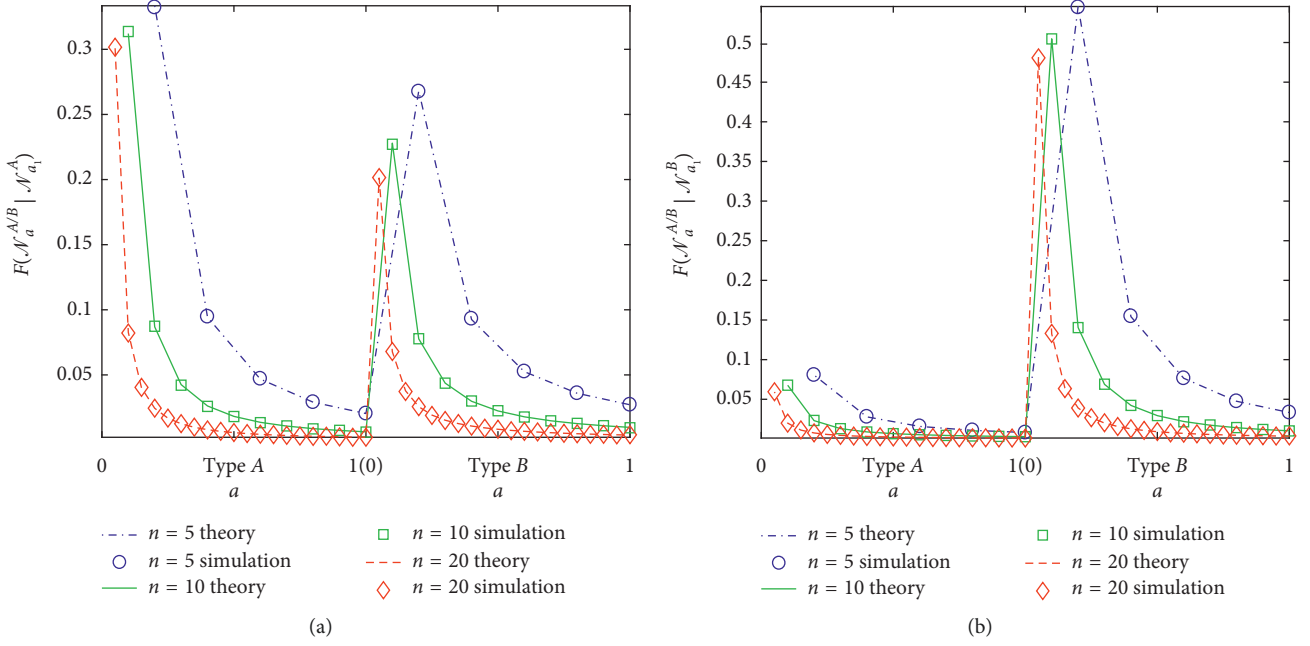


FIGURE 3: The conditional probability corresponding to the nodes with activity a_1 in the instantaneous network G_t . In (a) and (b), we show $F(\mathcal{N}_a^{A/B} | \mathcal{N}_{a_1}^A)$ and $F(\mathcal{N}_a^{A/B} | \mathcal{N}_{a_1}^B)$ for the type A nodes and type B nodes, respectively. Theoretical and simulation results are labeled in the same manner as in Figure 2. For $n = 5, 10$, and 20 , $\varepsilon_n = 0.015, 0.0702$, and 0.3981 , respectively. Other parameters are set to be the same as those in Figure 2.

By integrating equation (21) over all the node classes and ignoring the second-terms in I_A^t , we can obtain the expression of I_A^{t+1} :

$$I_A^{t+1} = (1 - \mu)I_A^t + \beta m \left[pI_A^t \langle a \rangle (1 + \lambda_1(m-1)) + (1-p)I_B^t \langle a \rangle (1 + \lambda_2(m-1)) + (p\theta_A^t + (1-p)\theta_B^t) \right]. \quad (22)$$

Moreover, the closed expression for θ_A^{t+1} is derived by multiplying both sides of equation (21) by a and integrating over all classes spectrum. That is,

$$\theta_A^{t+1} = (1 - \mu)\theta_A^t + p\beta m I_A^t \left[\langle a^2 \rangle + \lambda_1 \langle a \rangle^2 (m-1) \right] + (1-p)\beta m I_B^t \left[\langle a^2 \rangle + \lambda_2 \langle a \rangle^2 (m-1) \right] + \beta m \langle a \rangle \left[p\theta_A^t + (1-p)\theta_B^t \right]. \quad (23)$$

Similarly, the expressions of I_B^{t+1} and θ_B^{t+1} that correspond to the nodes of type B can be written as

$$\begin{aligned} I_B^{t+1} &= (1 - \mu)I_B^t + \beta m \left[pI_A^t \langle a \rangle (1 + \lambda_2(m-1)) + (1-p)I_B^t \langle a \rangle (1 + \lambda_1(m-1)) + (p\theta_A^t + (1-p)\theta_B^t) \right], \\ \theta_B^{t+1} &= (1 - \mu)\theta_B^t + p\beta m I_A^t \left[\langle a^2 \rangle + \lambda_2 \langle a \rangle^2 (m-1) \right] + (1-p)\beta m I_B^t \left[\langle a^2 \rangle + \lambda_1 \langle a \rangle^2 (m-1) \right] + \beta m \langle a \rangle \left[p\theta_A^t + (1-p)\theta_B^t \right]. \end{aligned} \quad (24)$$

Thus, we obtain the following closed system of the master equations for I_A^t , I_B^t , θ_A^t , and θ_B^t :

$$\begin{aligned}
\partial_t I_A^t &= -\mu I_A^t + p\beta m \langle a \rangle (1 + \lambda_1 (m-1)) I_A^t + (1-p)\beta m \langle a \rangle (1 + \lambda_2 (m-1)) I_B^t + \beta m (p\theta_A^t + (1-p)\theta_B^t), \\
\partial_t I_B^t &= -\mu I_B^t + p\beta m \langle a \rangle (1 + \lambda_2 (m-1)) I_A^t + (1-p)\beta m \langle a \rangle (1 + \lambda_1 (m-1)) I_B^t + \beta m (p\theta_A^t + (1-p)\theta_B^t), \\
\partial_t \theta_A^t &= -\mu \theta_A^t + p\beta m [\langle a^2 \rangle + \lambda_1 \langle a \rangle^2 (m-1)] I_A^t + (1-p)\beta m [\langle a^2 \rangle + \lambda_2 \langle a \rangle^2 (m-1)] I_B^t + \beta m \langle a \rangle (p\theta_A^t + (1-p)\theta_B^t), \\
\partial_t \theta_B^t &= -\mu \theta_B^t + p\beta m [\langle a^2 \rangle + \lambda_2 \langle a \rangle^2 (m-1)] I_A^t + (1-p)\beta m [\langle a^2 \rangle + \lambda_1 \langle a \rangle^2 (m-1)] I_B^t + \beta m \langle a \rangle (p\theta_A^t + (1-p)\theta_B^t),
\end{aligned} \tag{25}$$

whose Jacobian matrix is given by

$$J = \begin{bmatrix} -\mu + p\beta m \langle a \rangle (1 + \lambda_1 (m-1)) & (1-p)\beta m \langle a \rangle (1 + \lambda_2 (m-1)) & p\beta m & (1-p)\beta m \\ p\beta m \langle a \rangle (1 + \lambda_2 (m-1)) & -\mu + (1-p)\beta m \langle a \rangle (1 + \lambda_1 (m-1)) & p\beta m & (1-p)\beta m \\ p\beta m [\langle a^2 \rangle + \lambda_1 \langle a \rangle^2 (m-1)] & (1-p)\beta m [\langle a^2 \rangle + \lambda_2 \langle a \rangle^2 (m-1)] & -\mu + p\beta m \langle a \rangle & (1-p)\beta m \langle a \rangle \\ p\beta m [\langle a^2 \rangle + \lambda_2 \langle a \rangle^2 (m-1)] & (1-p)\beta m [\langle a^2 \rangle + \lambda_1 \langle a \rangle^2 (m-1)] & p\beta m \langle a \rangle & -\mu + (1-p)\beta m \langle a \rangle \end{bmatrix}. \tag{26}$$

The critical condition for the epidemic threshold is determined by the largest eigenvalue Λ_{\max} of J . That is, if

$$\Lambda_{\max}(J) > 0, \tag{27}$$

the epidemics will outbreak in the network. Here, we cannot get the explicit expression of equation (27) for general

systems. In the following, we will discuss two special cases of the system.

If $\lambda_1 = \lambda_2 = \lambda$, there is no homophily effect on interactions between *neighbor* nodes in any groups. Thus, all nodes in the network can be considered to be of the same type. In this case, the master equations of the system can be rewritten as

$$\begin{aligned}
\partial_t I_A^t &= -\mu I_A^t + \beta m \langle a \rangle (1 + \lambda (m-1)) (pI_A^t + (1-p)I_B^t) + \beta m (p\theta_A^t + (1-p)\theta_B^t), \\
\partial_t I_B^t &= -\mu I_B^t + \beta m \langle a \rangle (1 + \lambda (m-1)) (pI_A^t + (1-p)I_B^t) + \beta m (p\theta_A^t + (1-p)\theta_B^t), \\
\partial_t \theta_A^t &= -\mu \theta_A^t + \beta m [\langle a^2 \rangle + \lambda \langle a \rangle^2 (m-1)] (pI_A^t + (1-p)I_B^t) + \beta m \langle a \rangle (p\theta_A^t + (1-p)\theta_B^t), \\
\partial_t \theta_B^t &= -\mu \theta_B^t + \beta m [\langle a^2 \rangle + \lambda \langle a \rangle^2 (m-1)] (pI_A^t + (1-p)I_B^t) + \beta m \langle a \rangle (p\theta_A^t + (1-p)\theta_B^t),
\end{aligned} \tag{28}$$

which can be further reduced to be

$$\begin{aligned}
\partial_t I^t &= -\mu I^t + \beta m \langle a \rangle (1 + \lambda (m-1)) I^t + \beta m \theta^t, \\
\partial_t \theta^t &= -\mu \theta^t + \beta m [\langle a^2 \rangle + \lambda \langle a \rangle^2 (m-1)] I^t + \beta m \langle a \rangle \theta^t,
\end{aligned} \tag{29}$$

with $I^t = pI_A^t + (1-p)I_B^t$ and $\theta^t = p\theta_A^t + (1-p)\theta_B^t$. The corresponding Jacobian matrix J is given by

$$J = \begin{bmatrix} -\mu + \beta m \langle a \rangle (1 + \lambda (m-1)) & \beta m \\ \beta m [\langle a^2 \rangle + \lambda \langle a \rangle^2 (m-1)] & -\mu + \beta \langle a \rangle m \end{bmatrix}, \tag{30}$$

and the two eigenvalues of J are

$$\Lambda_{(1,2)} = -\mu + \frac{\beta m (\lambda (m-1) + 2) \langle a \rangle \pm \beta m \sqrt{(\lambda (m-1) + 2)^2 \langle a \rangle^2 + 4(\langle a^2 \rangle - \langle a \rangle^2)}}{2}. \tag{31}$$

Thus, the critical condition (27) for the presence of an endemic state is

$$\frac{\beta}{\mu} > \frac{2}{m (\lambda (m-1) + 2) \langle a \rangle + m \sqrt{(\lambda (m-1) + 2)^2 \langle a \rangle^2 + 4(\langle a^2 \rangle - \langle a \rangle^2)}}. \tag{32}$$

Particularly, with $\lambda = \lambda_1$, the results for the cases of $p = 0$ and $p = 1$ are also included in (32).

Furthermore, if $\lambda = 0$, no interaction between *neighbor* nodes will be allowed and equation (32) will recover the result $(\beta/\mu) > (1/(m(\langle a \rangle + \sqrt{\langle a^2 \rangle})))$ for AD networks. In addition, if $\lambda = 1$, an interaction between any two *neighbor* nodes can be expected in each group. Then, equation (32) gives $(\beta/\mu) > (2/(m(m+1)\langle a \rangle + m\sqrt{(m+1)^2\langle a \rangle^2 + 4(\langle a^2 \rangle - \langle a \rangle^2)}))$, which is degenerated to the results of the SAD model in [26].

Next, we discuss the case where the number of nodes of both types is the same. With $p = 0.5$, the master equation (25) can be reduced as

$$\begin{aligned}\partial_t I^t &= -\mu I^t + \beta m \langle a \rangle \left[1 + \frac{\lambda_1 + \lambda_2}{2} (m-1) \right] I^t + \beta m \theta^t, \\ \partial_t \theta^t &= -\mu \theta^t + \beta m \left[\langle a^2 \rangle + \frac{\lambda_1 + \lambda_2}{2} \langle a \rangle^2 (m-1) \right] I^t + \beta m \langle a \rangle \theta^t.\end{aligned}\quad (33)$$

Following the analysis steps of equation (29), we can derive the eigenvalues of J determined by system (33). That is,

$$\Lambda_{(1,2)} = -\mu + \frac{\beta m \left(((\lambda_1 + \lambda_2)/2)(m-1) + 2 \right) \langle a \rangle \pm \beta m \sqrt{\left(((\lambda_1 + \lambda_2)/2)(m-1) + 2 \right)^2 \langle a \rangle^2 + 4(\langle a^2 \rangle - \langle a \rangle^2)}}{2}.\quad (34)$$

Thus, the critical condition (27) is given by

$$\frac{\beta}{\mu} > \frac{2}{m \left(((\lambda_1 + \lambda_2)/2)(m-1) + 2 \right) \langle a \rangle + m \sqrt{\left(((\lambda_1 + \lambda_2)/2)(m-1) + 2 \right)^2 \langle a \rangle^2 + 4(\langle a^2 \rangle - \langle a \rangle^2)}}.\quad (35)$$

Comparing equations (32) and (35), we can see that the case of a uniform mixing of two types of nodes is equivalent to the case of $\lambda = ((\lambda_1 + \lambda_2)/2)$ without homophily effect.

4. Numerical Simulations

In this section, we perform numerical simulations to investigate the epidemic spreading process on CAD networks. For simplicity, we consider five activity rates of the nodes, i.e., $a \in \{0.5, 0.6, 0.7, 0.8, 0.9\}$ [24], and the activity rate distribution obeys $F(a) \propto a^{-2}$. Accordingly, the proportions of nodes with the activity rate 0.5, 0.6, 0.7, 0.8, and 0.9 are 0.344, 0.239, 0.176, 0.135, and 0.106, respectively. At each time step, the number of edges generated by an active node is chosen to be $m = 5$. In the following simulations, the size of network, the recovery rate, and the initial fraction of randomly chosen infectious nodes in the population are set to be $N = 5,000$, $\mu = 0.5$, and $I^0 = 0.1$, respectively. If not specified, we assume the proportion of type A nodes, the interaction probabilities for two *neighbor* nodes of the same type, and different types to be $p = 0.6$, $\lambda_1 = 0.5$, and $\lambda_2 = 0.1$, respectively. All the Monte Carlo simulation results are averaged over 40 independent realizations.

Firstly, we compare the performance of the statistical model and the classical MF model on characterizing the SIS epidemic spreading process. To this end, we depict the time evolution of the fraction of infectious nodes in the network. Under different infection rate β , the evolution curves based on the statistical model (17), the classical MF model (19) and the Monte Carlo simulations are recorded. The results are presented in Figure 4. It is observed that the statistical model captures the evolution

curve for all four values of β , while the classical MF model approximates the simulation results only for the case of $\beta = 0.01$. This is consistent with the analysis in Section 2, as the approximation adopted in the classical MF model is accurate only when the effective infection rate is below the epidemic threshold. Furthermore, as shown in Figures 4(b)–4(d), the classical MF model provides an upper bound for the infection level in simulations. The difference between this upper bound and the actual infection level increases as the infection rate β increases. According to the results of Figure 4, we can conclude that analyzing the statistical properties of the instantaneous network is beneficial to accurately characterize the contagion process in the time-varying networks.

Next, we investigate the impact of homophily-driven group interactions on the final steady state of the epidemic spreading. The growth of the final fraction of infectious nodes I^∞ versus the infection rate β is presented in Figure 5. It is shown that I^∞ decreases with the increasing p from 0 to 0.5. Since p describes the proportion of type A nodes in the network, an increasing $p \leq 0.5$ indicates more uniform mixing of both type nodes. This implies that any two *neighbor* nodes in any groups are more likely to be of different types. Due to the homophily effect, it is expected that there are less group interactions among nodes. Consequently, as shown in Figure 5, the contagion process in the network is impeded and the epidemic threshold is enhanced with an increasing p . Furthermore, noting that the definition of type A and type B is exchangeable, the effect of a varying p from 0.5 to 1 is the same as a varying $1 - p$ from 0 to 0.5. This fact, together with the observation in Figure 5, demonstrates that the impact of homophily effect on the contagion process is the greatest

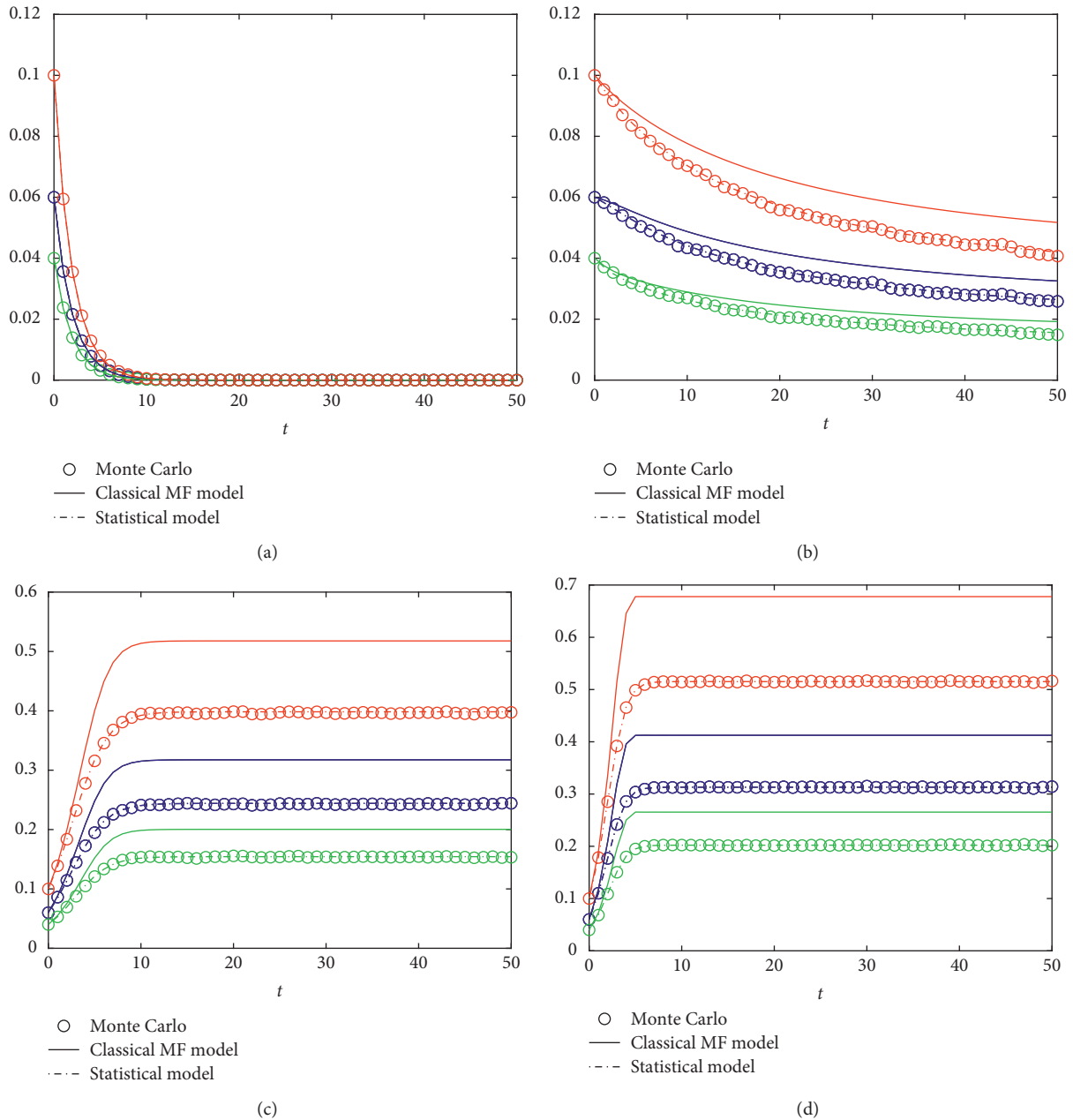


FIGURE 4: The time evolution of the fraction of infectious nodes of the type A (marked with blue), type B (marked with green), and both two types (marked with red). The circles, solid lines, and dotted lines represent the results of Monte Carlo simulations, the classical MF model, and the statistical model, respectively. (a) $\beta = 0.01$. (b) $\beta = 0.05$. (c) $\beta = 0.1$. (d) $\beta = 0.15$.

when $p = 0.5$. In fact, this result can be obtained by analyzing the average probability of an interaction between any two *neighbor* nodes. That is, for the given λ_1 and λ_2 , the value of $\lambda_1(p^2 + (1-p)^2) + 2\lambda_2p(1-p)$ reaches its minimum with $p = 0.5$. In addition, as illustrated in Figure 5, the epidemic thresholds suggested by the Monte Carlo simulations and the statistical model are close to the theoretical values given by equation (27). This observation, in turn, validates our analytical results presented in Section 3.

Furthermore, we consider the contagion processes on the AD networks, the SAD networks, and the proposed CAD

networks. According to the analysis in Section 3, the AD model and the SAD model are indeed two special cases of the CAD model and they are obtained by letting $\lambda_1 = \lambda_2 = 0$ and $\lambda_1 = \lambda_2 = 1$, respectively.

In Figure 6, we depict the growth of the final fraction of infectious nodes I^∞ versus the infection rate β in the three networks. Obviously, the group interactions between two *neighbor* nodes promote the contagion process in the network. Compared with the AD network, the epidemic threshold decreases in both CAD and SAD networks. Specifically, the critical infection rates corresponding to the

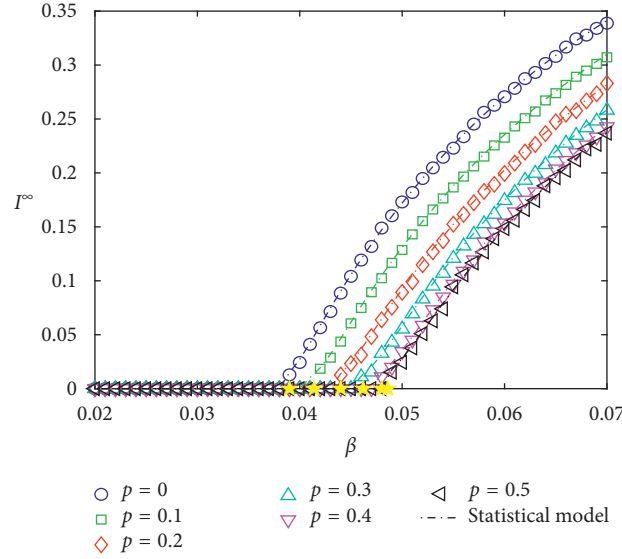


FIGURE 5: The final fraction of infectious nodes as a function of the infection rate β for different proportions of type A nodes, p . The epidemic thresholds suggested by the Monte Carlo simulation and the statistical model are recorded as soon as I^∞ is slightly than 0.01 (the critical value above which the realization is considered as endemic). The analytical values of the epidemic threshold given by equation (27) are labeled by yellow pentagons.

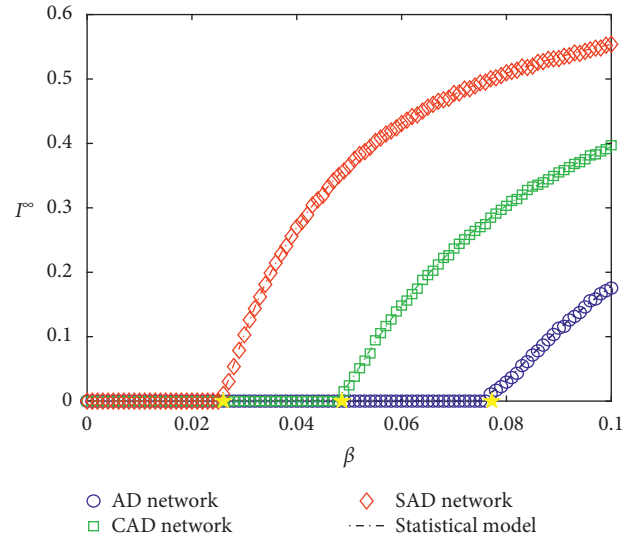


FIGURE 6: The final fraction of infectious nodes as a function of the infection rate β in AD networks, CAD networks, and SAD networks. Here, we chose $p = 0.5$ for the CAD networks to illustrate the theoretical results of equation (35). The epidemic thresholds suggested by the Monte Carlo simulations, the statistical model, and the theoretical analysis are recorded and labeled following the same step as in Figure 5.

epidemic threshold are satisfied with $\beta_c^{\text{AD}} > \beta_c^{\text{CAD}} > \beta_c^{\text{SAD}}$. In addition, the theoretical results of the epidemic threshold given by equations (32) and (35) are approximate to the simulation results.

Finally, we investigate the impacts of the infection rate β and the group interaction probabilities λ_1 and λ_2 on the epidemic threshold. Figure 7 presents a comparison between the theoretical and numerical values of the epidemic threshold for different parameter combinations, where the numerically obtained steady-state infection density is color coded and shown as the background. Here, we assume λ_2 is proportional to λ_1 , i.e., $\lambda_2 = \tau\lambda_1$, and the proportionality factors τ are

chosen as 0.1, 0.5, and 1 in Figures 7(a)–7(c), respectively. For a fixed parameter set, the dashed curve represents the theoretical value of λ_1^c as a function of the infection rate β based on equation (27). The markers are the corresponding numerical results of λ_1^c . From Figure 7, the theoretical results agree with the simulation results for all three cases of τ . For each τ , there are two critical values of β , i.e., β_c^1 and β_c^2 . When $\beta < \beta_c^1$ or $\beta > \beta_c^2$, the epidemic threshold is not related to λ_1 . Specifically, the epidemics will vanish from the network with $\beta < \beta_c^1$ or it breaks out with $\beta > \beta_c^2$. When $\beta_c^1 < \beta < \beta_c^2$, it is observed that λ_1^c decreases with an increase of β . Here, β_c^1 is obtained under $\lambda_1 = 1$ and is determined by the value of λ_2 .

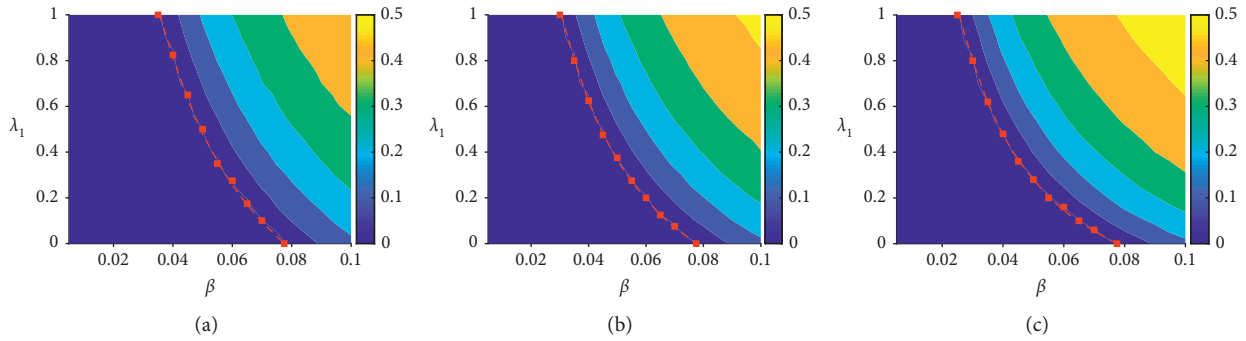


FIGURE 7: The final fraction of infectious nodes as a function of the infection rate β and the group interaction probabilities λ_1 and λ_2 . In each subgraph, the curve with markers represents the dependence of the predicted critical value λ_1^c for the onset of the epidemic on the infection rate β . λ_1^c suggested by Monte Carlo simulations is obtained in the same way as deriving the critical value of β in Figure 5. (a) $\lambda_2 = 0.1\lambda_1$; (b) $\lambda_2 = 0.5\lambda_1$; (c) $\lambda_2 = \lambda_1$.

Given the three candidates of τ in Figure 7, we can see that β_c^1 decreases with an increase of τ due to more group interactions between nodes with different types. Moreover, since the CAD network turns into the AD network with $\lambda_1 = \lambda_2 = 0$, β_c^2 is exactly the epidemic threshold in the AD network.

5. Conclusions

In this paper, we investigate the effect of group interactions involving more than two individuals on the contagion process in the time-varying networks. We consider that the interactions are affected by the individual attribute. The homophily effects on interaction patterns are studied. A new network model that extends the paradigmatic activity-driven model to the framework of simplicial complex networks is presented. The statistical properties of the instantaneous network based on this model are explored and incorporated in characterizing the epidemic spreading process. Results show that these properties provide a more accurate description of the evolution of epidemic. Furthermore, it is demonstrated that the homophily-driven group interactions have a profound impact on the epidemic threshold. Our results generalize the existing results of the epidemic threshold on two paradigms of activity-driven network models. However, the proposed model also comes with some limitations. Here, we consider that the activity rate of a node is independent of its type. Future work may explore a general activity probability distribution that incorporates the effect of different node types. Moreover, we assume that the infection rates through pairwise interaction and multinode interaction are the same. This model can be further developed by exploring different infection rates for different interaction patterns.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant no. 61873194.

References

- [1] G. Cimini, T. Squartini, F. Saracco, D. Garlaschelli, A. Gabrielli, and G. Caldarelli, "The statistical physics of real-world networks," *Nature Reviews Physics*, vol. 1, no. 1, pp. 58–71, 2019.
- [2] L. Huo, F. Ding, C. Liu, and Y. Cheng, "Dynamical analysis of rumor spreading model considering node activity in complex networks," *Complexity*, vol. 2018, Article ID 1049805, 10 pages, 2018.
- [3] P. Hu, L. Ding, and T. Hadzibeganovic, "Individual-based optimal weight adaptation for heterogeneous epidemic spreading networks," *Communications in Nonlinear Science and Numerical Simulation*, vol. 63, pp. 339–355, 2018.
- [4] M. O. Jackson and D. López-pintado, "Diffusion and contagion in networks with heterogeneous agents and homophily," *Network Science*, vol. 1, no. 1, pp. 49–67, 2013.
- [5] G. Kossinets and D. J. Watts, "Origins of homophily in an evolving social network," *American Journal of Sociology*, vol. 115, no. 2, pp. 405–450, 2009.
- [6] W. Wang, M. Tang, H. F. Zhang, and Y. C. Lai, "Dynamics of social contagions with memory of nonredundant information," *Physical Review E*, vol. 92, no. 1, Article ID 012820, 2015.
- [7] R.-J. Mei, L. Ding, X.-M. An, and P. Hu, "Modeling for heterogeneous multi-stage information propagation networks and maximizing information," *Chinese Physics B*, vol. 28, no. 2, Article ID 028701, 2019.
- [8] M. E. Newman, "Assortative mixing in networks," *Physical Review Letters*, vol. 89, no. 20, Article ID 208701, 2002.
- [9] X. Ma and Y. Ma, "The local triangle structure centrality method to rank nodes in networks," *Complexity*, vol. 2019, Article ID 9057194, 16 pages, 2019.
- [10] X. Chen, R. Wang, M. Tang, S. Cai, H. E. Stanley, and L. A. Braunstein, "Suppressing epidemic spreading in multiplex networks with social-support," *New Journal of Physics*, vol. 20, no. 1, Article ID 013007, 2018.
- [11] J. M. Kumpula, J. P. Onnela, J. Saramäki, K. Kaski, and J. Kertész, "Emergence of communities in weighted networks," *Physical Review Letters*, vol. 99, no. 22, Article ID 228701, 2007.

- [12] G. Palla, I. Derényi, I. Farkas, and T. Vicsek, “Uncovering the overlapping community structure of complex networks in nature and society,” *Nature*, vol. 435, no. 7043, pp. 814–818, 2005.
- [13] J. G. Young, G. Petri, F. Vaccarino, and A. Patania, “Construction of an efficient sampling from the simplicial configuration model,” *Physical Review E*, vol. 96, no. 3, Article ID 032312, 2017.
- [14] O. T. Courtney and G. Bianconi, “Weighted growing simplicial complexes,” *Physical Review E*, vol. 95, no. 6, Article ID 062301, 2017.
- [15] C. Giusti, E. Pastalkova, C. Curto, and V. Itskov, “Clique topology reveals intrinsic geometric structure in neural correlations,” *Proceedings of the National Academy of Sciences*, vol. 112, no. 44, pp. 13455–13460, 2015.
- [16] C. Giusti, R. Ghrist, and D. S. Bassett, “Two’s company, three (or more) is a simplex: algebraic-topological tools for understanding higher-order structure in neural data,” *Journal of Computational Neuroscience*, vol. 41, no. 1, pp. 1–14, 2016.
- [17] S. Maletić and M. Rajković, “Consensus formation on a simplicial complex of opinions,” *Physica A: Statistical Mechanics and Its Applications*, vol. 397, pp. 111–120, 2014.
- [18] N. Perra, B. Gonçalves, R. Pastor-Satorras, and A. Vespignani, “Activity driven modeling of time varying networks,” *Scientific Reports*, vol. 2, no. 1, p. 469, 2012.
- [19] Y. Lei, X. Jiang, Q. Guo, Y. Ma, M. Li, and Z. Zheng, “Contagion processes on the static and activity-driven coupling networks,” *Physical Review E*, vol. 93, no. 3, Article ID 032308, 2016.
- [20] Q.-H. Liu, X. Xiong, Q. Zhang, and N. Perra, “Epidemic spreading on time-varying multiplex networks,” *Physical Review E*, vol. 98, no. 6, 2018.
- [21] D. Li, D. Han, J. Ma et al., “Opinion dynamics in activity-driven networks,” *EPL (Europhysics Letters)*, vol. 120, no. 2, Article ID 28002, 2017.
- [22] M. X. Liu, W. Wang, Y. Liu, M. Tang, S. M. Cai, and H. F. Zhang, “Social contagions on time-varying community networks,” *Physical Review E*, vol. 95, no. 5, Article ID 052306, 2017.
- [23] A. Moinet, R. Pastor-Satorras, and A. Barrat, “Effect of risk perception on epidemic spreading in temporal networks,” *Physical Review E*, vol. 97, no. 1, Article ID 012313, 2018.
- [24] P. Hu, L. Ding, and X. An, “Epidemic spreading with awareness diffusion on activity-driven networks,” *Physical Review E*, vol. 98, no. 6, 2018.
- [25] H. Yang, C. Gu, M. Tang, S.-M. Cai, and Y.-C. Lai, “Suppression of epidemic spreading in time-varying multiplex networks,” *Applied Mathematical Modelling*, vol. 75, pp. 806–818, 2019.
- [26] G. Petri and A. Barrat, “Simplicial activity driven model,” *Physical Review Letters*, vol. 121, no. 22, Article ID 228301, 2018.
- [27] L. Kovanen, K. Kaski, J. Kertesz, and J. Saramaki, “Temporal motifs reveal homophily, gender-specific patterns, and group talk in call sequences,” *Proceedings of the National Academy of Sciences*, vol. 110, no. 45, pp. 18070–18075, 2013.




Hindawi

Submit your manuscripts at
www.hindawi.com

