A colored mean-field model for analyzing the effects of awareness on epidemic spreading in multiplex networks

Qingchu Wu[1] and Gaoxi Xiao[2,3]

[1]*College of Mathematics and Information Science, Jiangxi Normal University, Jiangxi 330022, China*
[2]*School of Electrical and Electronic Engineering, Nanyang Technological University, 639798, Singapore*
[3]*Complexity Institute, Nanyang Technological University, 639798, Singapore*
(Dated: September 18, 2018)

We study the impact of susceptible nodes' awareness on epidemic spreading in social systems, where the systems are modelled as multiplex networks coupled with an information layer and a contact layer. We develop a colored heterogeneous mean-field model taking into account the portion of the overlapping neighbors in the two layers. With theoretical analysis and numerical simulations, we derive the epidemic threshold which determines whether the epidemic can prevail in the population, and find that the impacts of awareness on threshold value only depends on epidemic information being available in network nodes' overlapping neighborhood. When there is not any link overlap between the two network layers, the awareness cannot help to raise the epidemic threshold. Such an observation is different from that in a single-layer network, where the existence of awareness almost always helps.

Network epidemiology involves individual contact information and behavioral response. It should be meaningful to study how the epidemic information-based behavioral response affects the epidemic spreading. We can assume that an individual promptly obtains relatively accurate information from the current circumstances and responses to the epidemics via reducing its susceptibility. Generally speaking, the information network may not be exactly the same as the contact network and yet there may exist remarkable link overlap between the information and contact networks. Hence, a multiplex network framework can be used to analyze the influence of awareness on the epidemic spreading. To solve the theoretical analysis problem, we develop an edged-colored heterogeneous mean-field approach, and derive the threshold value of epidemic outbreak. By doing so, we can analyze the effects of the link overlap and inter-layer degree correlation between the information and contact layers. It is found that only epidemic information being available in network nodes' neighborhood overlapping in two layers helps change the epidemic threshold.

I. INTRODUCTION

In recent years, modelling disease spreading in complex networks as a research problem has attracted wide attentions from various research communities in mathematical, computer, physical and biological sciences areas. The developed epidemic spreading theories have found applications in social networks, transportation systems and computer networks [1], etc.

Among such researches, an important topic is to understand the effects of awareness on epidemic spreading, also known as the effects of risk perception. Specifically, many infectious diseases (e.g., SARS and flu) exhibit certain visible symptoms which can be easily observed by ordinary individuals. Further, people may know who have been infected, not necessarily by direct observations, but by information propagating from their friends. With the awareness of the epidemic spreading in their surroundings, healthy individuals may tend to take certain actions (e.g., wearing protective masks, washing hands frequently, or staying home etc.), which generally help reduce their risks of getting infected [2, 3]. Disease-induced awareness therefore plays a critical role in epidemic spreading and control, and cannot be ignored in any serious modeling efforts aiming to help improve real-life applications [4].

To incorporate the effects of awareness into the epidemic spread modeling, a typical approach is to try to reveal the relation between individual susceptibility (denoted by y) and the epidemic information it has (denoted by x), which may be represented by an awareness/response function $y = \phi(x)$. It may be assumed that this function satisfies (i) $\phi(0) = \phi_0$; (ii) $\phi'(x) \leq 0$; and (iii) $0 \leq \phi(x) \leq 1$ [3]. Some researches have been conducted on developing various awareness functions, in linear [3, 5] or nonlinear forms [6–8]. Besides, the system spreading dynamics model with a general awareness function has been investigated for the susceptible-infected-susceptible (SIS) model [9] and the susceptible-infected-recovered/removed (SIR) model [10] respectively, adopting an effective degree (EFD) modeling approach and having managed to achieve a satisfactory accuracy [11, 12].

An important topic for studying the effects of awareness is to evaluate how risk perception may affect [3, 5–10] epidemic threshold, infections with a transmissibility higher than which may have a good chance to prevail. This is highly relevant to epidemic control for public health. It is found that epidemic threshold may be strongly affected by local behavioral responses based on individuals' situation awareness. Since the epidemic information has many resources, it is of significance to study which type of the epidemic information spreading and the corresponding responses may affect epidemic threshold. In [3], we considered the case where there exists spreading of two types of information, namely the infection information in the neighborhood and in the whole network respectively. The former is called the *local* information, which is obtained

by individual observations or communications. The latter is called the *global* information, which may be obtained from the mass media (e.g., television or radio broadcasts regarding the spread of any disease). It is shown that the epidemic threshold depends on the response to the local (but not global) information. Yan *et al.* [5] analyzed the effects of epidemic information spreading from two different types of information resources: the mildly infected neighbors and severely infected neighbors. The former still can infect susceptible nodes and are called the *direct* information resources. The latter cannot infect susceptible nodes and are called the *indirect* information resources. It is revealed that the epidemic threshold can only be affected by information from direct information resources. In [10], we studied the SIR epidemic model with a general awareness function and obtained that the infection information from the infected neighbors can affect the epidemic threshold, while the recovery information from the recovered neighbors does not help. For a susceptible node, the information about the infected nodes that can potentially transmit the disease to the susceptible node is referred to as the *contact-based epidemic information*. Overall, one can conclude that only the contact-based epidemic information affects the epidemic threshold.

The above studies have all adopted the assumption that epidemic information propagation and disease spreading happen in the same network. An important recent development is introducing *multiplex network* model, also known as multilayer network model, into the research on epidemic spreading in social systems. In a multiplex network, there exist two different layers which may not necessarily be of the same topology or having the same set of nodes: awareness spreads on the *information layer* (or say subnetwork) and the epidemic spreads on the *contact layer*. From a realistic point of view, people may get epidemic information from their friends via mobile phones or online platforms. The friendship network may not be exactly the same as the contact network [13], yet there may exist remarkable link overlap between the friendship and contact networks. In this sense, the framework of multiplex/multilayer networks can be effective in analyzing the dynamical interplay between individual behaviors and spreading dynamics [14–16].

While extensive studies have been carried out on epidemic spreading in multiplex networks with a strong sense on analyzing awareness spreading dynamics (e.g., [16, 17]), studies focusing on the effects of risk perception in multiplex networks models are still relatively limited. Massaro and Bagnoli [18] were the first to consider a discrete-time model with risk perception in a multiplex network. They developed an efficient iteration method to compute the epidemic threshold. Their results showed that the epidemic threshold can be raised higher by increasing the structure similarity between the information and contact layers. However, this study did not provide a general dynamical model or derive an explicit expression of epidemic threshold in multiplex networks with general degree distribution and link overlap.

There are some important research problems remain largely open. First, how could a continuous-time dynamical evolution model be built to predict the epidemic spreading dynamics and

further to derive the explicit expression of epidemic threshold? Secondly, how does the epidemic threshold depend on the inter-layer degree correlation (i.e., the correlation between the node degrees in different layers [19–21]) and the portion of overlapping links in the whole network [22–25]? Thirdly, what kinds of epidemic information may affect the epidemic threshold as that in the single network?

To help answer these questions, in this paper we consider the effect of awareness on epidemic spreading in multiplex networks with general degree distribution and awareness function. We shall develop a novel generalized form of heterogeneous mean-field (HMF) approach to build a dynamical model and derive an explicit expression of epidemic threshold. Further, we will analyze the influences of link overlap, inter-layer correlation and behavioral response on the epidemic threshold and epidemic prevalence. These theoretical arguments are also verified by the continuous-time stochastic simulations.

The remainder of this paper is organized as follows: In Section II, we propose the multiplex network modelling framework used in the theoretical analysis; then in Section III, we build the dynamical model and derive the condition of an epidemic outbreak. In Section IV, we present numerical simulations to verify the theoretical results, and to analyze the influence of link overlap and behavioral response on the epidemic spreading. Finally, in Section V we conclude the paper.

II. MULTIPLEX NETWORK MODELLING FOR STUDYING EFFECTS OF SITUATION AWARENESS

A multiplex network is a network with multiple types of links connecting the same set of nodes on multiple different layers, composing into a subnetwork on each of these layers. A general multiplex network model may consist of M layers and N nodes and each layer can be defined by a subgraph $G_x = (V_x, E_x)$, where V_x is the vertex set and E_x is the edge set of layer x. We consider the case where $M = 2$, i.e, the duplex network with two layers, denoted as a and b respectively. We define that the layer a denotes the information layer, on which a node may spread and receive epidemic information, and the layer b represents the contact layer, on which the epidemic spreads. Since the whole network has a finite size, there exist a maximal degree M_a for layer a and a maximal degree M_b for layer b, respectively.

Pastor-Satorras and Vespignani developed a heterogenous mean-field (HMF) approach to analyze the SIS epidemic spreading in heterogenous networks with a given degree distribution $P(k)$ [26]. In the HMF modeling framework, network nodes are grouped by their degrees and dynamical states. Such an approach is very effective in analyzing the multi-state epidemic model in complex networks [27] and scaling properties [28]. When considering multiplex/mutilayer networks, a natural generalization of the HMF model is to introduce the vector degree $\vec{k} = (k_a, k_b)$ to denote a class of nodes with degree k_a in layer a and degree k_b in layer b [29]. By using the generalized HMF approach, Sanz *et al.* [30] studied the dynamics of two interacting pathogens in multiplex networks. In [21], we studied the epidemic spreading of SIS model in

multiplex networks and the immunization.

The generalized HMF approach however is not effective in analyzing the coupling effects between information and contact subnetworks, as it does not fully utilize the information of overlapping degree in the duplex network. Such an observation will become clearer in the following subsections. To reflect subnetwork overlap in more details, we here introduce two types of vector degrees: (i) $\vec{k} = (k_a, k_b, k_o)$ following the joint probability distribution $p(k_a, k_b, k_o)$; and (ii) $\vec{k^r} = (k_a^r, k_b^r, k_o)$ following the joint probability distribution $q(k_a^r, k_b^r, k_o)$. Here k_o denotes the number of overlapping neighbors in two layers, named as the *overlapping degree* [22]. k_a^r (k_b^r) denotes the the number of neighbors in layer a (b) except for the overlapping neighbors. In other words, $k_a^r = k_a - k_o$ and $k_b^r = k_b - k_o$. We call them as the *remaining degrees* in layers a and b, respectively. It is easy to see that the relation between the two distributions satisfies

$$q(k_a^r, k_b^r, k_o) = p(k_a^r + k_o, k_a^r + k_o, k_o). \tag{1}$$

As a generalization of the configuration model in a single network [31], one can generate the random duplex network with a given joint degree probability $p(\vec{k})$. To be specific, three degree sequences corresponding to three kinds of degrees are firstly generated according to $q(k_a^r, k_b^r, k_o)$ obtained by Eq. (1); then we may construct three randomly connected subnetworks by using the generating algorithm of the standard configuration model, avoiding the link overlap across each layer. Finally, one link between a node and its k_o neighbors in the third subnetwork is copied and changed into an edge between the same pair of nodes in both the first and the second subnetworks.

For the joint vector degree distribution $p(k_a, k_b, k_o)$, the respective marginal probability distributions of node degree $\vec{k}$ read as

$$p(k_a) = \sum_{k_b, k_o} p(k_a, k_b, k_o),$$

$$p(k_b) = \sum_{k_a, k_o} p(k_a, k_b, k_o),$$

$$p(k_o) = \sum_{k_a, k_b} p(k_a, k_b, k_o).$$

Moreover, its n-th order moment can be written as

$$\langle k_a^n \rangle = \sum_{k_a, k_b, k_o} k_a^n p(k_a, k_b, k_o) = \sum_{k_a} k_a^n p(k_a),$$

$$\langle k_b^n \rangle = \sum_{k_a, k_b, k_o} k_b^n p(k_a, k_b, k_o) = \sum_{k_b} k_b^n p(k_b),$$

and

$$\langle k_o^n \rangle_p = \sum_{k_a, k_b, k_o} k_o^n p(k_a, k_b, k_o) = \sum_{k_o} k_o^n p(k_o).$$

For the remaining degree distribution, its marginal probabilities (i.e., $q(k_a^r)$, $q(k_b^r)$, $q(k_o)$) and the n-th order moment of

FIG. 1: An example of the neighborhood structure of a node in a two-layer network where links in layer a (b) are denoted by solid (dotted) lines. For this node, $k_a = 7, k_b = 5, k_o = 3$ and $k_a^r = k_a - k_o = 4, k_b^r = k_b - k_o = 2$. Additionally, its infectious degrees $n_a = 3, n_b = 2, n_o = 1$ and its remaining infectious degrees $n_a^r = n_a - n_o = 2$ and $n_b^r = n_b - n_o = 1$.

the joint probability (i.e., $\langle k_{a^r}^n \rangle, \langle k_{b^r}^n \rangle, \langle k_o^n \rangle_q$ can also be obtained similarly.

Let $N_{\vec{k}}$ ($N_{\vec{k^r}}$) denote the number of nodes in class with degree $\vec{k}$ ($\vec{k^r}$). Clearly, $N_{\vec{k}} = Np(k_a, k_b, k_o)$ and $N_{\vec{k^r}} = Nq(k_a^r, k_b^r, k_o)$.

Denote $\mathcal{N}_x(i)$ as the neighborhood of node i at the layer x for $x = a, b$. Furthermore, we define the overlapping ratio for node i as follows

$$\sigma(i) = \frac{|\mathcal{N}_a(i) \cap \mathcal{N}_b(i)|}{\min\{|\mathcal{N}_a(i)|, |\mathcal{N}_b(i)|\}}.$$

Then, the mean overlapping ratio is given by

$$\bar{\sigma} = \frac{\sum_{i=1}^N \sigma(i)}{N}.$$

Given the joint probability $q(k_a^r, k_b^r, k_o)$, we shall then have

$$\begin{aligned} \bar{\sigma} &= \sum_{i=1}^N \frac{k_o(i)}{\min\{k_a(i), k_b(i)\}} \\ &= \sum_{k_a^r, k_b^r, k_o} \frac{k_o q(k_a^r, k_b^r, k_o)}{\min\{k_a^r + k_o, k_b^r + k_o\}}. \end{aligned}$$

Another measure of multiplex network is the clustering coefficient. Baxter *et al.* [32] gave a general formula for computing the clustering coefficient (both global clustering coefficient in the whole network and partial clustering coefficient in a given layer or several given layers) in a configuration multiplex network defined by joint multi-degree distribution. They showed that the cluster coefficient is very small for a large sparse random network. The clustering coefficient or the number of cycles in a graph is relevant for accurate epidemic modelling [31]. In our model, the mean clustering coefficient C of

the aggregated network shall be used to characterize the transitivity of the multiplex network, as will be seen in the next section.

III. THEORETICAL ANALYSIS

A. Notations and assumptions

In this paper we investigate the effect of situation awareness using the SIS model, which is widely applicable and may be adopted for modelling some epidemic diseases such as meningitis and gonorrhea [33]. The standard network SIS spreading model assumes that a susceptible node is infected at a probability $\beta \Delta t$ by each of it infected neighbors, and each infected node recovers with a probability $\mu \Delta t$ and moves to the susceptible state again within a time Δt.

Individual awareness may reshape the function of infection rate, which is dependent of the number of the infected neighbors. Denote $n_x(i)$ as the number of infected neighbors of node i at the layer x and $n_o(i)$ the number of the infected neighbors in the overlapping region of node i. These values form a vector $\vec{n} = (n_a, n_b, n_o)$. Meanwhile, we define $n_x^r(i)$ be the number of remaining infected neighbors of node i at the layer x and $n_o^r(i) = n_o(i)$. These values also can be written into a vector $\vec{n^r} = (n_a^r, n_b^r, n_o)$ (see Fig. 1). With these notations, β is replaced by $\beta\psi(\vec{k^r}, \vec{n^r})$, which is referred to as the awareness function and it must satisfy: (i) $\psi(\vec{k^r}, \vec{0}) = g(\vec{k^r})$; (ii) $\psi(\vec{k^r}, \vec{n_1^r}) \leq \psi(\vec{k^r}, \vec{n_2^r})$ if $\vec{n_1^r} \geq \vec{n_2^r}$; (iii) $\psi(\vec{k^r}, \vec{n^r}) \geq 0$; and (iv) $\vec{0} \leq \vec{n^r} \leq \vec{k^r}$.

Clearly, $\psi(\vec{k^r}, \vec{n^r})$ is defined as a generalization of the function $\phi(x)$ and it comprises three kinds of information sources: (i) $\Omega_a \setminus \Omega_b$, i.e., the neighborhood on only layer a; (ii) $\Omega_b \setminus \Omega_a$, i.e., the neighborhood on only layer b; and (iii) $\Omega_a \cap \Omega_b$ i.e., the neighborhood on both of two layers a, b. To evaluate the impact of epidemic-information based awareness on the epidemic spreading, we assume that $g(\vec{k^r}) \equiv 1$ and $\psi(\vec{k^r}, \vec{n^r})$ satisfies $n_b^r \equiv 0$ since the awareness depends on the information subnetwork. That is to say, $\psi(\vec{k^r}, \vec{n^r}) = \psi(k_a^r, 0, k_o, n_a^r, 0, n_o)$.

In epidemiology, the mathematical model provides a powerful tool to analyze the disease dynamics. Due to the merits of HMF approach in analyzing recurrence epidemic spreading [12, 34], we adopt the HMF approach to model the dynamical evolution. Note that the previous modelling approach based on vector degree (k_a, k_b) is not suitable for the present model because the mean-field rate equations based on this classification strategy cannot be closed and be further analyzed. To tackle the problem, we use the vector degree $\vec{k^r}$ to build the dynamical models.

Specifically, denote $\rho_{\vec{k^r}}(t)$ $(s_{\vec{k^r}}(t))$ as the fraction of infected (susceptible) nodes of vector degree $\vec{k^r}$. Obviously, $\rho_{\vec{k^r}}(t) + s_{\vec{k^r}}(t) = 1$. Let $I_{\vec{k^r}}(t)$ represent the number of infected nodes at time t in class with degree $\vec{k}$ ($\vec{k^r}$). We have $\rho_{\vec{k^r}}(t) = I_{\vec{k^r}}(t)/N_{\vec{k^r}}$.

B. A colored HMF approach

The mean-field rate equations depend on the transition probabilities from susceptible to infected and from infected to susceptible respectively, i.e., $W_{s\rho}^{\vec{k^r}}(t) : s_{\vec{k^r}}(t) \to \rho_{\vec{k^r}}(t)$ and $W_{\rho s}^{\vec{k^r}}(t) : \rho_{\vec{k^r}}(t) \to s_{\vec{k^r}}(t)$. In an infinitesimal interval $(t, t + \Delta t]$, the dynamical equations can be presented as

$$\rho_{\vec{k^r}}(t + \Delta t) = \rho_{\vec{k^r}}(t)\left[1 - W_{\rho s}^{\vec{k^r}}(t)\right] + \left[1 - \rho_{\vec{k^r}}(t)\right] W_{s\rho}^{\vec{k^r}}(t).$$
(2)

This mean-field equation does not consider dynamical correlations between two connected nodes [35], meaning that the states of each node's neighbors are independent with the state of that node. Following the definition of the stochastic model, it is easy to know that $W_{\rho s}^{\vec{k^r}}(t) = \gamma \Delta t$. However, it is not straightforward to compute the value of $W_{s\rho}^{\vec{k^r}}(t)$. Using the law of total probability, we can get that

$$
\begin{aligned}
W_{s\rho}^{\vec{k^r}}(t) &= \mathbb{P}\left[s_{\vec{k^r}}(t) \to \rho_{\vec{k^r}}(t)\right] \\
&= \sum_{n_a^r, n_b^r, n_o} \mathbb{P}\left[s_{\vec{k^r}}(t) \to \rho_{\vec{k^r}}(t)|n_a^r, n_b^r, n_o\right]\mathbb{P}\left[n_a^r, n_b^r, n_o\right] \\
&= \sum_{n_a^r, n_b^r, n_o} \left\{1 - \left[1 - \beta\psi(\vec{k^r}, \vec{n^r})\Delta t\right]^{n_b}\right\}\mathbb{P}\left[n_a^r, n_b^r, n_o\right]
\end{aligned}
$$

where $n_b = n_b^r + n_0$. After substituting this into the equation (2) and performing a simple transformation, one can obtain the rate equation that

$$
\begin{aligned}
\frac{d}{dt}\rho_{\vec{k^r}}(t) &= -\gamma\rho_{\vec{k^r}}(t) \\
&+ \left[1 - \rho_{\vec{k^r}}(t)\right]\beta \sum_{n_a^r, n_b^r, n_o} \psi(\vec{k^r}, \vec{n^r})n_b\mathbb{P}[n_a^r, n_b^r, n_o].
\end{aligned}
$$
(3)

Here, the summation computation is employed by taking $0 \leq n_a^r \leq k_a^r, 0 \leq n_b^r \leq k_b^r$ and $0 \leq n_o \leq k_o$.

We can see that Eq. (3) should be closed by establishing the expression of the joint probability $\mathbb{P}[n_a^r, n_b^r, n_o]$, which is actually the conditional probability $\mathbb{P}\left[n_a^r, n_b^r, n_o|\vec{k^r}\right]$. In the present work, the topological connectivity for each layer is *unclustered*, which means that one neighbor's state is independent of another. Moreover, the aggregated network structure is also required to be unclustered. This holds approximately for a multiplex network with a small number of cycles of length 3. In that case, the joint probability can be approximately decomposed into the product of three terms, i.e., $\mathbb{P}\left[n_a^r, n_b^r, n_o|\vec{k^r}\right] = \mathbb{P}\left[n_a^r|\vec{k^r}\right]\mathbb{P}\left[n_b^r|\vec{k^r}\right]\mathbb{P}\left[n_o|\vec{k^r}\right]$.

For further analysis, we let $\theta_{\vec{k^r}, x}(t)$ be the probability that a randomly-chosen node among $x = k_a^r, k_b^r, k_o$ neighbors of a susceptible node with vector degree $\vec{k^r}$ is infected at time t. If $x = 0$, then it is natural to define $\theta_{\vec{k^r}, x}(t) = 0$. In what follows, we focus on the nonzero case. Generally, $\theta_{\vec{k^r}, x}(t)$ is

related to $\vec{k^r}$ class [36, 37] and can be computed by a conditional or joint probability

$$\theta_{\vec{k^r},x}(t) = \frac{\sum_{\vec{k^r}'} x' \mathbb{P}[\vec{k^r}, \vec{k^r}'] \rho_{\vec{k^r}'}(t)}{\sum_{\vec{k^r}'} x' \mathbb{P}[\vec{k^r}, \vec{k^r}']}, \qquad (4)$$

Here x' is a component of $\vec{k^r}'$ and represents k_a', k_b' or k_o'. By using these quantities, we can get that

$$\mathbb{P}\left[n_a^r, n_b^r, n_o\right]$$
$$= \mathbb{P}\left[n_a^r | \vec{k^r}\right] \mathbb{P}\left[n_b^r | \vec{k^r}\right] \mathbb{P}\left[n_o | \vec{k^r}\right]$$
$$= \mathbf{B}_{k_a^r, n_a^r}(\theta_{\vec{k^r}, k_a^r}(t)) \mathbf{B}_{k_b^r, n_b^r}(\theta_{\vec{k^r}, k_b^r}(t)) \mathbf{B}_{k_o, n_o}(\theta_{\vec{k^r}, k_o}(t)),$$

where

$$\mathbf{B}_{k,i}(x) = \binom{k}{i} x^i (1-x)^{k-i},$$

denotes the binomial factor [38]. As a result, we can write the dynamical model as follows

$$\frac{d}{dt}\rho_{\vec{k^r}}(t) = -\gamma \rho_{\vec{k^r}}(t) + \beta \left[1 - \rho_{\vec{k^r}}(t)\right] \sum_{n_a^r, n_b^r, n_o} \{\psi(\vec{k^r}, \vec{n^r}) n_b$$
$$\times\; \mathbf{B}_{k_a^r, n_a^r}(\theta_{\vec{k^r}, k_a^r}(t)) \mathbf{B}_{k_b^r, n_b^r}(\theta_{\vec{k^r}, k_b^r}(t)) \mathbf{B}_{k_o, n_o}(\theta_{\vec{k^r}, k_o}(t))\}. \qquad (5)$$

For mathematical analysis, we assume that the topological connectivity for each layer is *uncorrelated*, which indicates that the degree k' of a randomly-chosen neighbor of a susceptible node is uncorrelated to the degree k of the susceptible node; then the probability of a randomly-chosen neighbor of a susceptible node being k' is proportional to $k'p(k')$. Since the uncorrelated assumption is made for every layer, $\mathbb{P}[k_a, k_a'] = \mathbb{P}[k_a]\mathbb{P}[k_a']$ and $\mathbb{P}[k_b, k_b'] = \mathbb{P}[k_b]\mathbb{P}[k_b']$. This is however insufficient for modelling the multiple degree correlations in multiplex networks. It is thus further assumed that $\mathbb{P}[\vec{k^r}, \vec{k^r}'] = \mathbb{P}[\vec{k^r}]\mathbb{P}[\vec{k^r}']$, which holds approximately when the link overlap is randomly generated. Under this assumption, $\theta_{\vec{k^r},x}(t)$ is not related to $\vec{k^r}$ and obeys

$$\theta_{\vec{k^r},x}(t) \simeq \frac{\sum_{\vec{k^r}'} x' q(\vec{k^r}') \rho_{\vec{k^r}'}(t)}{\sum_{\vec{k^r}'} x' q(\vec{k^r}')} := \theta_x(t). \qquad (6)$$

After substituting (6) into (5), we obtain a simple dynamical system

$$\frac{d}{dt}\rho_{\vec{k^r}}(t) = -\gamma \rho_{\vec{k^r}}(t) + \beta[1 - \rho_{\vec{k^r}}(t)] \sum_{n_a^r, n_b^r, n_o} \{\psi(\vec{k^r}, \vec{n^r}) n_b$$
$$\times\; \mathbf{B}_{k_a^r, n_a^r}(\theta_{k_a^r}(t)) \mathbf{B}_{k_b^r, n_b^r}(\theta_{k_b^r}(t)) \mathbf{B}_{k_o, n_o}(\theta_{k_o}(t))\}. \qquad (7)$$

Remark 1: For each layer of the network, we actually allocate two colors to all the edges. That is, every overlapped edge has color 1 and every remaining edge has color 2. This

framework can be generalized to the SIS epidemic model with general colored degree in single-layer networks, where l colors are allocated to all the edges and the joint probability is denoted by $q(\vec{k}) = q(k_1, k_2, \cdots, k_l)$ with $1 \le l \le L$. Here, L is the total number of edges in the single-layer network. In this sense, the proposed HMF model is referred to as the *(edge-)colored heterogenous mean-field* (cHMF) model.

Remark 2: In Eq. (7), we cannot simplify the model formulation. However, when we consider a linear form of awareness function, i.e., $\psi(\vec{k^r}, \vec{n^r}) = 1 - \alpha n_a / k_a$, the summation of the system can be computed by making use of the properties of binomial distribution.

Remark 3: Note that the configuration network has a small clustering coefficient and is degree-uncorrelated since it is randomly generated by the vector degree distribution. The mean-field rate equations (7) thus hold naturally for the configuration multiplex network randomly generated by $q(\vec{k^r})$.

C. Epidemic thresholds

We shall work on determining the condition of epidemic outbreak, by which we can study the impact of individual awareness on the epidemic threshold. We firstly perform the linear stability analysis on Eq. (7). The linear system takes the form

$$\frac{d}{dt}\rho_{\vec{k^r}}(t)$$
$$= -\gamma \rho_{\vec{k^r}}(t) + \beta \psi(\vec{k^r}, 0, 0, 1) k_o \theta_o(t) + \beta k_b^r \theta_{b^r}(t). \qquad (8)$$

Thus the Jacobian matrix of the above model at point disease-free equilibrium (denoted by J) is

$$J = \gamma \mathcal{I} + \beta H,$$

where

$$H_{\vec{k^r}, \vec{k^r}'}$$
$$= \left[\frac{\psi(\vec{k^r}, 0, 0, 1) k_o k_o'}{\langle k_o \rangle_q} + \frac{k_b^r k_b^{r'}}{\langle k_b^r \rangle}\right] q(\vec{k^r}').$$

From this, one can observe that the epidemic threshold τ_c is inversely proportional to the leading eigenvalue of matrix H, i.e.,

$$\tau_c = \frac{1}{\lambda_{\max}(H)}. \qquad (9)$$

This indicates that the impact of epidemic information on the epidemic threshold is related to the overlapping structure. Specifically, the structure of network overlapping can be divided into five separate cases: (i) $E_a \cap E_b = \emptyset$; (ii) $E_a = E_b$; (iii) $E_a \supset E_b$; (iv) $E_a \subset E_b$; and (v) $E_a \cap E_b \ne \emptyset$, $E_a \backslash E_b \ne \emptyset$, $E_b \backslash E_a \ne \emptyset$. It is easy to compute their respective overlapping ratios. For case (i), $\bar{\sigma} = 0$; for cases (ii)(iii)(iv), $\bar{\sigma} = 1$; and for case (v), $0 < \bar{\sigma} < 1$.

FIG. 2: The clustering coefficient as a function of mean overlapping ratio for a multiplex network overlapped by two same ER graphs with mean degree $z = 5$. All results are averaged by 20 independent network realizations.

FIG. 3: Comparison between the mean-field models and the stochastic simulations on an ER network with the mean degree $z = 5$. Results by the colored mean-field model (i.e., $\bar{\sigma} = 0.5$) are shown by the solid line and results by the standard HMF model (i.e., $\bar{\sigma} = 1.0$) are shown by the dashed line. The GA stochastic simulation results are shown in open square symbols.

By Eq. (9), we can obtain the results of epidemic thresholds for different cases. These results are summarized in Table I, where $\tau_c^1 = \frac{\langle k_b^r \rangle}{\langle (k_b^r)^2 \rangle}$, $\tau_c^2 = \frac{\langle k_o \rangle_q}{\langle \psi(\vec{k^r},0,0,1)k_o^2 \rangle}$, and τ_c^3 is computed by Eq. (9).

IV. SIMULATION RESULTS AND DISCUSSIONS

In this section, we perform continuous-time stochastic simulations by using the Gillespie algorithm (GA) method, which is frequently used in the network epidemiology [11, 39]. The multiplex network is overlapped by two Erdős-Rényi (ER) random graphs [40] with a size $N = 1000$, the maximal degree $M_a = M_b = M = 10$ and a mean degree $z = 5$ [11]. Unless otherwise specified, the stochastic simulations start with 10 initially infectious seeds and the recovery rate $\gamma = 1.0$. Additionally, the infection density in the metastable state ρ is averaged over 100 epidemic dynamics and the error bars show the standard deviation.

TABLE I: Linear approximation of the joint probability.

	(i)	(ii)	(iii)	(iv)	(v)
θ_{a^r}	+	0	+	0	+
θ_{b^r}	+	0	0	+	+
θ_{o^r}	0	+	+	+	+
$\bar{\sigma}$	0	1	1	1	$(0,1)$
τ_c	τ_c^1	τ_c^2	τ_c^2	τ_c^3	τ_c^3

In fact, τ_c^3 can also be explicitly given. Note that Eq. (8) is similar to the formulation of Eq. (1) in ref. [21]. Using almost the same technique, the explicit expression of epidemic threshold τ_c^3 can be derived as

$$\tau_c^3 = \frac{2}{c_2 + c_3 + \sqrt{(c_2 - c_3)^2 + 4c_1}}.$$

Here, $c_1 = \frac{\langle \psi(\vec{k^r},0,0,1)k_o^2 \rangle \langle k_b^r k_o \rangle}{\langle k_o \rangle_q \langle k_b^r \rangle}$, $c_2 = \frac{\langle \psi(\vec{k^r},0,0,1)k_o^2 \rangle}{\langle k_o \rangle_q}$, and $c_3 = \frac{\langle k_b^{r2} \rangle}{\langle k_b^r \rangle}$. This formula clearly shows that the epidemic threshold is closely related to the moment of vector degree distribution $q(\vec{k^r})$. Meanwhile, we also can see that the inter-layer degree correlation has no direct impact on the epidemic threshold since k_a^r does not present in this expression except for the awareness term. This is different from the SIS dynamics on multiplex networks [21], where the inter-layer correlation has a strong impact on the outbreak threshold.

To analyze the effects of overlap on the epidemic spreading [23, 24], we construct a set of multiplex networks by edge rewiring. Specifically, we firstly generate an ER random graph as both information layer and virtual layer. Then in the virtual layer we successively choose one link which lies in both layers and another link without any common endpoint, and we swap them with a rate ω by changing their endpoints while avoiding having new overlapping edges or multi-edge in the virtual layer. Finally, when no other link can be chosen, the final virtual layer is regarded as the contact layer. When $\omega = 0$, the two ER networks completely overlap and $\bar{\sigma} = 1$. When $\omega = 1$, there is not any link overlap in the network and $\bar{\sigma} = 0$. In Fig. 2, we can see how the link overlap affects the clustering of the duplex network. With $\bar{\sigma}$ increasing, C linearly declines. Since the clustering coefficient is very small, our model can be used to the whole span of parameter $\bar{\sigma}$.

We first consider $\psi(\vec{k^r}, \vec{n^r}) \equiv 1$ as a reference. In this case, the spreading process reduces to the SIS epidemic spreading in a single network and the corresponding epidemic model can

be changed to

$$\frac{d}{dt}\rho_{k_b^r,k_o}(t) = -\gamma\rho_{k_b^r,k_o}(t) + \beta[1 - \rho_{k_b^r,k_o}(t)] \times$$

$$\left[\frac{k_b^r \sum k_b^r q(k_b^r,k_o)\rho_{k_b^r,k_o}(t)}{\sum k_b^r q(k_b^r,k_o)} + \frac{k_o \sum k_o q(k_b^r,k_o)\rho_{k_b^r,k_o}(t)}{\sum k_o q(k_b^r,k_o)}\right].$$
(10)

Moreover, when $\bar{\sigma} = 0$ or 1, Eq. (10) becomes the standard HMF model [26]. In Fig. 3, we compare Eq. (10) and the standard HMF model with the GA simulations for the basic SIS epidemic model in ER networks with a mean degree $z = 5$. The distribution $q(k_b^r, k_o)$ is generated by the multiplex network with two overlapping ER networks with the same topology. When $\bar{\sigma} = 0$, this corresponds the standard HMF model prediction. While when $\bar{\sigma} \simeq 0.5$ (see Fig. 2), this is a typical colored HMF model. From this figure, one can observe that the results of cHMF model are very close to those HMF model (both thresholds $\lambda_c \simeq 1/(z+1) = 0.1667$). Since the cHMF model (10) does not consider the vector degree correlation, its prediction is not better than HMF model. Mainly because that all these mean-field like models have not taken the dynamical correlation between two connected nodes into consideration, the resulting theoretical predictions are slightly lower than the corresponding simulation results [37, 41].

Next, we consider $\psi(\vec{k^r}, \vec{n^r}) = 1 - \alpha n_a/k_a$, which is linearly dependent on the fraction of infected neighbors in the information layer. This awareness function reflects that individual behavior has a weak effect on the epidemic risk [10]. In Fig. 4, we compare the GA simulations and the model prediction (7) for two extreme cases: $\bar{\sigma} = 1$ and $\bar{\sigma} = 0$. For the former case (Fig. 4(a)), one can see that the local awareness has a remarkable effect on the epidemic threshold when the two subnetworks completely overlap. It is noticed that this case corresponds to the disease awareness model in a single network [3]. For the latter case, as shown in Fig. 4(b), we cannot find any change of epidemic threshold since the overlapping ratio is zero. Hence, one can see that the link overlap has a strong impact on the outbreak threshold. This also shows that only contact-based epidemic information can affect the epidemic threshold and other epidemic information (e.g., n_a^r) has no effect. However, other epidemic information indeed can decrease the final epidemic size, as can be clearly observed in Fig. 4(b). We also can see a small shift between the simulation and model, which is mainly due to the dynamical correlation between two connected nodes [29, 37, 42]. Compared to the discrepancy for the standard HMF model in Fig. 3, the results predicted by the colored HMF model are acceptable.

To further investigate the influence of other epidemic information on ρ, we also consider $\psi(\vec{k^r}, \vec{n^r}) = 1 - \alpha n_a^r/k_a$, i.e., the risk perception is only related to the other epidemic information. In this case, the epidemic threshold is unchanged for any α. In Fig. 5, we assume that $\beta = 1.0$ and focus on the impact of α on ρ. One can see that other epidemic information can obviously suppress the epidemic spreading by reducing the final epidemic size.

FIG. 4: The fraction of infected nodes ρ in the metastable state as a function of β for two types of duplex networks: (a) $\langle op \rangle = 1$ and (b) $\langle op \rangle = 0$. In each plane, the model prediction is presented as the dashed ($\alpha = 0.1$) and solid ($\alpha = 0.9$) lines and the GA stochastic simulation results are represented by open square ($\alpha = 0.1$) and circle ($\alpha = 0.9$).

V. CONCLUSIONS

In conclusion, we analyzed the impact of local awareness on the epidemic spreading in multiplex networks overlapped by an information layer and a contact layer. Each susceptible node may receive risk information from its neighborhood on the information layer and hence reduces its susceptibility to the infection along the edges on the contact layer. The information layer thus helps prevent epidemic spreading. As such a model cannot be easily analyzed using the traditional mean-field approach, we developed a novel mean-field model considering a colored degree of each node. Given the joint probability of three colored degrees in duplex networks, we derived the explicit expression of epidemic threshold and found that the only the risk information from the overlapped neighborhood can affect the epidemic threshold, regardless of the inter-layer correlation and link overlap. As corroborated by the previous work [3, 5, 10], the epidemic threshold increases only with risk perception of contact-based epidemic spreading, suppressing thereby the disease transmission over the entire network.

The theoretical work reported in this paper may be ex-

tended to analyze spreading dynamics on multi-relation networks [43] and multiplayer networks [15, 29, 44], etc. The observation that only the epidemic information being available in network nodes overlapping neighborhood helps increase epidemic threshold implicates the importance of promoting epidemic information spreading through those social connections potentially allowing epidemic spreading as well. While extensive research efforts have been made to analyzing epidemic dynamics in complex netowrks using various mean-field like models [2, 37, 41, 42, 45–47], there is always a need to keep a good balance between the accuracy and the tractability of the analysis [48], which shall be of our future research interest.

FIG. 5: The fraction of infected nodes ρ in the metastable state as a function of α when $\beta = 1.0$. In this figure, the model predictions are given by the solid ($\bar{\sigma} = 0.5$) and dashed ($\bar{\sigma} = 0$) lines and the GA stochastic simulation results are represented by open circle ($\bar{\sigma} = 0.5$) and square ($\bar{\sigma} = 0$) symbols.

[1] R. Pastor-Satorras, C. Castellano, P. Van Mieghem and A. Vespignani, Epidemic processes in complex networks, Rev, Mod. Phys. 87 (2015) 925.

[2] S. Funk, E. Gilad and V.A. Jansen, Endemic disease, awareness, and local behavioural response, J. Theor. Biol. 264(2) (2010) 501-509.

[3] Q.C. Wu, X.C. Fu, M. Small and X.J. Xu, The impact of awareness on epidemic spreading in networks, Chaos 22(1) (2012) 013101.

[4] Z. Wang, M.A. Andrews, Z.X. Wu, L. Wang and C.T. Bauch, Coupled disease-behavior dynamics on complex networks: A review, Physics of life reviews 15 (2015) 1-29.

[5] Z.J. Yan, H. Huang, Y.H. Chen and Y.H. Pan, Identifying the direct risk source to contain epidemics more effectively, Phys. Rev. E 93 (2016) 012308.

[6] F. Bagnoli, P. Liò and L. Sguanci, Risk perception in epidemic modeling, Phys. Rev. E. 76(6) (2007) 61904.

[7] Y.L. Shang, Modeling epidemic spread with awareness and heterogeneous transmission rates in networks, J. Biol. Phys. 39(3) (2013) 489-500.

[8] H.F. Zhang, J.R. Xie, M. Tang, et al., Suppression of epidemic spreading in complex networks by local information based behavioral responses, Chaos 24(4) (2014) 043106.

[9] Q.C. Wu and W.F. Zhu, Toward a generalized theory of epidemic awareness in social networks, Int. J. of Mod. Phys. C 28(05) (2017) 1750070.

[10] Q.C. Wu and S.F. Chen, Susceptible-infected-recovered epidemics in random networks with population awareness, Chaos 27(10) (2017) 103107.

[11] J. Lindquist, J.L. Ma, P. van den Driessche, et al., Effective degree network disease models, J. Math. Biol. 62 (2011) 143-164.

[12] J.P. Gleeson, High-accuracy approximation of binary-state dynamics on networks, Phys. Rev. Lett. 107(6) (2011) 068701.

[13] X.G. Zhang, C.P. Zhang and Z. Jin, Structure of growing complex networks coupling with the friendship and contact relations, Chaos Solit. Fract. 104 (2017) 758-765.

[14] Y. Min, J. Hu, W. Wang, et al., Diversity of multilayer networks and its impact on collaborating epidemics, Phys. Rev. E 90(6) (2014) 062803.

[15] S. Boccaletti, G. Bianconi, R. Criado, et al., The structure and dynamics of multilayer networks, Phys. Rep. 544(1) (2014) 1-122.

[16] Y.H. Pan and Z.J. Yan, The impact of individual heterogeneity on the coupled awareness-epidemic dynamics in multiplex networks, Chaos, 28 (2018) 063123.

[17] C. Granell, S. Gómez, A. Arenas, Dynamical interplay between awareness and epidemic spreading in multiplex networks, Phys. Rev. Lett. 111 (2013) 128701.

[18] E. Massaro, F. Bagnoli, Epidemic spreading and risk perception in multiplex networks: a self-organized percolation method, Phys. Rev. E. 90(5) (2014) 052817.

[19] F. Battiston, V. Nicosia and V. Latora, Structural measures for multiplex networks, Phys. Rev. E 89(3) (2014) 032804.

[20] V. Nicosia and V. Latora, Measuring and modeling correlations in multiplex networks, Phys. Rev. E 92(3) (2015) 032805.

[21] Q.C. Wu, Y.J. Lou and W.F. Zhu, Epidemic outbreak for an SIS model in multiplex networks with immunization, Math. Biosci. 277 (2016) 38-46.

[22] G. Bianconi, Statistical mechanics of multiplex networks: Entropy and overlap, Phys. Rev. E 87(6) (2013) 062806.

[23] H. Yang, Impact of network overlapping on dynamical interplay between information and epidemics[C]//Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), 2016 12th International Conference on. IEEE, 2016: 316-320.

[24] X.L. Chen, R.J. Wang, M. Tang, et al., Suppressing epidemic spreading in multiplex networks with social-support, New J. Phys. 20(1) (2018) 013007.

[25] Q.C. Wu and T. Hadzibeganovic, Pair quenched mean-field approach to epidemic spreading in multiplex networks, Appl.

Math. Model. 60 (2018) 244-254.

[26] R. Pastor-Satorras and A. Vespignani, Epidemic spreading in scale-free networks, Phys. Rev. Lett. 86 (2001) 3200.

[27] N. Masuda and N. Konno, Multi-sate epidemic processes on complex networks. J. Theor. Biol. 243 (2006) 64-75.

[28] X.L. Peng, X.J. Xu, X.C. Fu and T. Zhou, Vaccination intervention on epidemic dynamics in networks, Phys. Rev. E 87 (2) (2013) 022813.

[29] A. Saumell-Mendiola, M. Serrano and M. Boguna, Epidemic spreading on interconnected networks, Phys. Rev. E 86 (2012) 026106.

[30] J. Sanz, C.Y. Xia, S. Meloni, et al., Dynamics of interacting diseases, Phys. Rev. X 4(4) (2014) 041005.

[31] M.E.J. Newman, S.H. Strogatz, and D.J. Watts, Random graphs with arbitrary degree distributions and their applications, Phys. Rev. E 64 (2001) 026118.

[32] G.J. Baxter, D.Cellai, S.N. Dorogovtsev and J.F.F. Mendes, Cycles and clustering in multiplex networks, Phys. Rev. E 94 (2016)062308.

[33] N. C. Grassly, C. Fraser, and G. P. Garnett, Host immunity and synchronized epidemics of syphilis across the United States, Nature 433 (2005) 417-421.

[34] L. Böttcher, J. Nagler and H.J. Herrmann, Critical behaviors in contagion dynamics, Phys. Rev. Lett. 118 (2017) 088301.

[35] J.P. Gleeson, S. Melnik, et al., Accuracy of mean-field theory for dynamics on real-world networks, Phys. Rev. E. 85 (2012) 026106.

[36] X.C. Fu, M. Small, D.M. Walker and H.F. Zhang, Epidemic dynamics on scale-free networks with piecewise linear infectivity and immunization, Phys. Rev. E 77 (2008) 036113.

[37] C.R. Cai, Z.X. Wu, M.Z.Q. Chen, P. Holme and J.Y. Guan, Solving the dynamic correlation problem of the susceptible-infected-susceptible model on networks, Phys. Rev. Lett. 116(25) (2016) 258301.

[38] V. Nagy, Mean-field theory of a recurrent epidemiological model, Phys. Rev. E 79 (2009) 066105.

[39] P.G. Fennell, S. Melnik and J.P. Gleeson, Limitations of discrete-time approaches to continuous-time contagion dynamics, Phys. Rev. E 94(5) (2016) 052125.

[40] P. Erdös and A. Rényi, On the evolution of random graphs, Publ. Math. Inst. Hung. Acad. Sci 5 (1960) 17-61.

[41] A.S. Mata, R.S Ferreira and S.C. Ferreira, Heterogeneous pair-approximation for the contact process on complex networks, New J. Phys. 16 (2014) 053006.

[42] S.C. Ferreira, C. Castellano and R. Pastor-Satorras, Epidemic thresholds of the susceptible-infected-susceptible model on networks: a comparison of numerical and theoretical results, Phys. Rev. E 86(4) (2012) 041125.

[43] R.Q. Li, B.M. Xu and M. Tang, Epidemic spreading on multi-relation networks, Acta Phys. Sin. 62 (2013) 168903.

[44] F.D. Sahneh and C. Scoglio, Competitive epidemic spreading over arbitrary multilayer networks, Phys. Rev. E 89(6) (2014) 062817.

[45] G. St-Onge, J.G. Young, E. Laurence, C. Murphy and L.J. Dubé, Phase transition of the susceptible-infected-susceptible dynamics on time-varying configuration model networks, Phys. Rev. E 97 (2018) 022305.

[46] M.J. Keeling, T. House, A.J. Cooper and L. Pellis, Systematic approximations to susceptible-infectious-susceptible dynamics on networks. PLoS Comput. Biol. 12(12) (2016) e1005296.

[47] W. Wang, M. Tang, H.E. Stanley and L.A. Braunstein, Unification of theoretical approaches for epidemic spreading on complex networks, Rep. Prog. Phys. 80 (2017) 036603 (16pp).

[48] C. Kyriakopoulos, G. Grossmann, V. Wolf, et al., Lumping of degree-based mean-field and pair-approximation equations for multistate contact processes, Phys. Rev. E 97(1) (2018) 012301.