Optimizing spreading dynamics in interconnected networks

Liming Pan,1 Wei Wang,2,1 Shimin Cai,1 and Tao Zhou1

1Web Sciences Center, School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu, 611731, China
2Cybersecurity Research Institute, Sichuan University, Chengdu 610065, China
(Dated: February 28, 2022)

Adding edges between layers of interconnected networks is an important way to optimize the spreading dynamics. While previous studies mostly focus on the case of adding a single edge, the theoretical optimal strategy for adding multiple edges still need to be studied. In this study, based on the susceptible-infected-susceptible (SIS) model, we investigate the problem of maximizing the stationary spreading prevalence in interconnected networks. For two isolated networks, we maximize the spreading prevalence near the critical point by choosing multiple interconnecting edges. We present a theoretical analysis based on the discrete-time Markov chain approach to derive the approximate optimal strategy. The optimal inter-layer structure predicted by the strategy maximizes the spreading prevalence, meanwhile minimizes the spreading outbreak threshold for the interconnected network simultaneously. Numerical simulations on synthetic and real-world networks show that near the critical point, the proposed strategy gives better performance than connecting large degree nodes and randomly connecting.

Spreading dynamics in interconnected networks relay on the inter-layer structure apart from the structure within each layer. For two given networks, how to design the interconnecting structure to optimize the spreading dynamics is a very appealing topic. Previous studies obtained the optimal strategy when considering adding a single edge in two-layer interconnected networks, while the optimal strategy of adding multiple edges lacks theoretical studies. Therefore in this study, a novel strategy is proposed to promote the spreading dynamics by adding multiple interconnecting edges for two isolated networks. Near the critical point, the spreading prevalence can be written in terms of the leading eigenvalue and the corresponding eigenvector of the adjacency matrix. Basing on an approximation scheme for the leading eigenvalue and eigenvector of the interconnected network, we optimize the spreading prevalence among all candidate edges. The optimal inter-layer structure is achieved by selecting edge candidates that are top-ranked by the product of eigenvector centrality of nodes in its two ends. Meanwhile, the optimal strategy also minimizes the outbreak threshold of the interconnected network. Numerical simulations on three pairs of synthetic networks and two pairs of real-world networks show that the strategy gives better performance than the heuristic strategies of connecting large degree nodes and of randomly connecting, especially near the critical point.

I. INTRODUCTION

In real-world social systems, individuals might communicate with others via multiple possible channels (such as Twitter, Facebook, LinkedIn). The communication relations with a specific channel can be represented by a network, where the nodes correspond to individuals, and the edges correspond to the communication relations. Therefore all the communication relations combined can be described by a multilayer network [1–4], where each layer corresponds to one of the communications channels. Multilayer networks and the dynamics on them have attracted attentions from diverse areas. It has been observed that multilayer networks display distinct collective behaviors compared to that of single-layer networks [5–13]. As an example, percolation processes on multilayer networks display first-order phase transitions [8, 14–17], which is intrinsically different from the second-order phase transitions on single-layer networks [18–20]. For evolutionary games, multilayer networks promote cooperation better than single-layer networks [12, 21, 22]. For synchronization processes, explosive synchronization and hysteresis loop are observed on multilayer networks [13].

Spreading dynamics on multilayer networks have attracted considerable attention in recent studies [2, 23]. Saumell-Mendiola, Serrano and Boguñá [31] have studied susceptible-infected-susceptible (SIS) model [32] on multilayer networks. They found that adding a small fraction of edges between layers can lead to the outbreak of epidemics while without these edges the epidemic extinct. The epidemic threshold depends on the structure of multilayer networks [9, 33] and it is possible to observe localization phenomena [34]. For susceptible-infected-recovered (SIR) model, Dickison, Havlin and Stanley [35] found that the system might exhibit a mixed phase, i.e., the epidemic outbreaks in one layer but not others. Refs. [36–38] studied a model with the spreading of information and epidemics simultaneously. The studies found that the diffusion of information can inhibit the spreading of epidemics.

The inter-layer structure in multilayer networks has significant impacts on the dynamics. Parshani et al. [39] found that a positive inter-layer degree correlation will inhibit large scale cascading failures. Ref. [31] found that those degree correlations will make the epidemic outbreak more easily, while Ref. [36] found that the positive inter-
We give some other heuristic strategies and finally in Sec. IV we investigate the problem of optimizing the spreading prevalence in two-layer networks by adding multiple inter-layer edges. For two isolated networks, we try to understand how to add a small number of edges to maximize the stationary spreading prevalence in the interconnected network. We mainly focus on the SIS model near the critical point. With a known formula given in [45], the epidemic prevalence near the critical point can be written in terms of the leading eigenvalue and eigenvector. We first develop a scheme for approximating the new leading eigenvalue and eigenvector for the interconnected network after adding those interconnecting edges. With this approximation scheme, we obtain a formula that predicts the stationary epidemic prevalence in the two-layer interconnected network. Then this approximated prevalence can be optimized among all possible inter-layer structures. The optimal inter-layer structure that maximizes the spreading prevalence will found to minimize the spreading outbreak threshold for the interconnected network simultaneously. Numerical simulations are performed to compare the strategy with some other heuristic strategies. The proposed strategy gives a better performance at least near the critical point when adding a small number of interconnecting edges.

The rest of the paper is organized as follows. In Sec. II, we introduce the basic setups of the model and some notations. In Sec. III, the theoretical derivations of the strategy is given. Then in Sec. IV the strategy is tested and compared to some other heuristic strategies and finally in Sec. V we give some conclusions and discussions.

II. MODEL

Starting with two isolated networks \( a \) and \( b \), we add a fixed number of edges to interconnect the two networks. The way of adding the edges will affect the dynamical behaviors on the interconnected network. We focus on the way of adding these edges that maximizing the spreading prevalence.

Let the adjacency matrices of the networks \( a, b \) be \( G_a \) and \( G_b \), respectively. The number of nodes in \( a(b) \) is \( N_a(N_b) \) and the number of edges is \( M_a(M_b) \). The total number of nodes is denoted by \( N = N_a + N_b \) and the total number of edges by \( M = M_a + M_b \). The adjacency matrix of the combined network then is

\[
G^0 = \begin{pmatrix}
G_a & 0 \\
0 & G_b
\end{pmatrix}.
\]

A set of edges with fixed cardinality \( \delta M \) will be added between the two networks. After the operation the new network combining \( a, b \) and interconnecting edges will have adjacency matrix \( G = G^0 + \delta G \), where

\[
\delta G = \begin{pmatrix}
0 & C \\
C^T & 0
\end{pmatrix}.
\]

Here \( C \) is an \( N_a \times N_b \) matrix which indicates how the inter-layer connections are added. Its elements take values \( C_{ij} \in \{0, 1\} \), where \( C_{ij} = 1 \) if and edge is added between the \( i \)-th node of network \( a \) and \( j \)-th node of network \( b \) and \( C_{ij} = 0 \) otherwise. The matrix satisfies the constraint on total number of added edges as

\[
\langle 1_{N_a}, C1_{N_a} \rangle = \delta M,
\]

where \( 1_{N_a} = [1, \cdots, 1]^T \) is all-one vector of length \( N_a \) and \( \langle \cdot \rangle \) is the inner product of two vectors.

By choosing among all possible assignment of \( C \), the epidemic prevalence can be maximized. In this paper, we consider the SIS model. For the SIS model, each node is in either the susceptible or infected state. In a discrete-time setting, at time step \( t \), infected nodes have a probability \( \lambda \) to infect their susceptible neighbors independently. Then the infected nodes (not including nodes get infected at the current time step) become susceptible again with probability \( \mu \). In large time limit, the density of infected nodes will converge to its stationary value. The target of the paper is to maximize the stationary spreading prevalence.

III. THEORY

We employ the discrete time Markov chain approach [46] to describe SIS model on general networks. For the discrete time Markov chain approach, the status of a node \( i \) is characterized by \( \rho_i(t) \), which is the probability that \( i \) is infected at time step \( t \). Then \( \rho_i(t) \) evolves according to the following discrete time equations

\[
\rho_i(t+1) = (1-\mu)\rho_i(t) + (1-\rho_i(t)) (1-q_i(t))
\] (4)
for $i \in \{1, \ldots, N\}$ and
\[
q_i(t) = \prod_{j=1}^{N} (1 - \lambda G_{ij} \rho_j(t)).
\] (5)

The first term on the r.h.s. of Eq. (4) corresponds to the probability that node $i$ is infected at $t$ and not recovered in $t+1$, and second term is the probability that $i$ is susceptible at $t$ and get infected by at least one infected neighbor. $q_i(t)$ in Eq. (5) is the probability that $i$ is not get infected by any of its neighbors at $t$.

The stationary solution is given by the limit in $t$
\[
\rho_i = \lim_{t \to \infty} \rho_i(t).
\] (6)

Let $\rho$ and $q$ be the vector with elements $\rho_i$ and $q_i$ for $i \in \{1, \ldots, N\}$. Near the critical point $\rho$ is expected to be small and the equation can be linearized as $q \approx 1 - \lambda G \rho$, the stationary equation reads
\[
G \rho = (\mu/\lambda) \rho.
\] (7)

The spreading outbreaks only when $\lambda/\mu > 1/\omega_1$ where $\omega_1$ is the leading eigenvalue of epidemic. It has been shown in [45] the stationary prevalence of epidemic is approximately
\[
\langle \rho \rangle \approx (\lambda^* \omega_1 - 1) \frac{\sum_{i=1}^{N} u_i}{N \sum_{i=1}^{N} u_i^2}
\] (8)

where $\langle \rho \rangle = (1/N) \sum_{i=1}^{N} \rho_i$, $\lambda^* = \lambda/\mu$ is the effective infection probability and $u$ is the eigenvector corresponding to the eigenvalue $\omega_1$. It can be seen that $\langle \rho \rangle$ is determined by the leading eigenvalue and eigenvector approximately near the critical point. Since $G$ is obtained by adding a small number of edges to $G^0$, its spectra should be closely related to that of $G^0$. Next we develop a scheme to approximate the spectra of $G$ from that of $G^0$.

First consider the spectra of matrix $G^0$. Let the $\omega_k^a$ for $k \in \{1, \ldots, N_a\}$ be the $k$-th eigenvalue of $G_a$ and $v_k^a$ the corresponding eigenvector. Similarly $\omega_l^b$ and $v_l^b$ with $l \in \{1, \ldots, N_b\}$ are eigenvalues and eigenvectors for $G_b$. The adjacency matrix of networks $a$ and $b$ combined $G^0$ is a diagonal block matrix by putting $G_a$ and $G_b$ in the diagonal, thus with eigenvalues
\[
\{\omega_k^a : k = 1, \ldots, N_a\} \cup \{\omega_l^b : l = 1, \ldots, N_b\}.
\] (9)

Clearly the leading eigenvalue of $G^0$ is $\max\{\omega_1^a, \omega_1^b\}$. Without losing of generality we assume $\omega_1^a \geq \omega_1^b$. The corresponding eigenvector for an eigenvalue $\omega_k^a$ is
\[
\hat{v}_k^a = (v_k^a, 0)^T,
\] (10)

which is by combining $v_k^a$ and all-zero vector of length $N_b$. Similarly, for eigenvalue $\omega_l^b$, the corresponding eigenvector of $G^0$ is
\[
\hat{v}_l^b = (0, v_l^b)^T
\] (11)

with zero vector of length $N_a$.

Now we consider adding a small number of interconnecting edges. By assuming these edges won’t shift the spectra of the two isolated networks too much, a first approximation of $u$ would be proportional to $v_1^a$ in the first $N_a$ elements, and to $v_1^b$ in the last $N_b$ elements, thus written as
\[
\begin{align*}
u^0 &= \beta^a v_1^a + \beta^b v_1^b
\end{align*}
\] (12)

where $\beta^a, \beta^b \in \mathbb{R}$ are coefficients to be determined. The eigenvector $u$ is given by the following limit
\[
\begin{align*}
u &= c \lim_{n \to \infty} (\omega_1^n)^{-n} G^n u^0
\end{align*}
\] (13)

for some constant $c$ and $G^n$ denotes self matrix multiplication of $G$ for $n$ times. Since the number of interconnecting edges is small, it can be assumed that $u^0$ is already close to $u$ by choosing $\beta^a$ and $\beta^b$ properly. The limit thus approximated by setting $n = 1$ and this gives
\[
\begin{align*}
u &\approx \frac{c}{\omega_1} \left( \beta^a \omega_1^a \hat{v}_1^a + \beta^b \omega_1^b \hat{v}_1^b + \beta^a \delta G \hat{v}_1^a + \beta^b \delta G \hat{v}_1^b \right).
\end{align*}
\] (14)

Rescale the parameters by
\[
\begin{align*}
\beta^a &\rightarrow \beta^a \frac{c}{\omega_1} \beta^b \omega_1^b \rightarrow \beta^b,
\end{align*}
\] (15)

the eigenvector $u$ is approximated by the form
\[
\begin{align*}
u &\approx u^0 + \delta u,
\end{align*}
\] (16)

where
\[
\begin{align*}
\delta u &= \beta^a \omega_1^a \delta G \hat{v}_1^a + \beta^b \omega_1^b \delta G \hat{v}_1^b.
\end{align*}
\] (17)

For scale-free networks which we mainly consider in the paper, it has been shown that the leading eigenvalue diverges in the thermodynamic limit [47]. Thus $\delta u$ can be ignored for large enough networks and $u$ is approximated as $u \approx u^0$.

Denote the gap by $g = \omega_1^a - \omega_1^b$. The leading eigenvalue $\omega_1$ of $G$ can be written as $\omega_1 = \omega_1^a + \delta \omega_1$, which is $\omega_1^a$ plus a correction term $\delta \omega_1$. With these approximations we arrive at the following eigenvalue equation
\[
(G^0 + \delta G) u^0 = (\omega_1^a + \delta \omega_1) u^0.
\] (18)

By definition
\[
G^0 u^0 = \beta^a \omega_1^a \hat{v}_1^a + \beta^b \omega_1^b \hat{v}_1^b,
\] (19)

and after some algebra Eq. (18) becomes
\[
(\delta G - \delta \omega_1 \mathbb{I}) u^0 = \beta^b \delta g \hat{v}_1^b,
\] (20)

where $\mathbb{I}$ is the $N$ by $N$ identity matrix. Multiplying $\hat{v}_1^a$ and $\hat{v}_1^b$ from the left separately gives the following equations
\[
\begin{align*}
\beta^b \langle v_1^a, Cv_1^b \rangle &= \delta \omega_1 \beta^a,
\beta^a \langle v_1^b, Cv_1^b \rangle - \beta^b &= \delta \omega_1 \beta^b.
\end{align*}
\] (21)
The equations can be written in the form of an eigenvalue equation
\[
\begin{pmatrix}
0 & E(C) \\
E(C) & -g
\end{pmatrix}
\begin{pmatrix}
\beta^a \\
\beta^b
\end{pmatrix}
= \delta \omega_1
\begin{pmatrix}
\beta^a \\
\beta^b
\end{pmatrix}
\] (22)
where we have denoted by \( E(C) = \langle v^a_1, C v^b_1 \rangle \). The equation gives two eigenvalues and pick the larger one which is
\[
\delta \omega_1 = \frac{1}{2} \left( (\sqrt{4E^2(C) + g^2} - g) \right).
\] (23)
The corresponding eigenvectors gives
\[
\beta^a = \sqrt{1 + \left( \frac{g}{2E(C)} \right)^2 + \frac{g}{2E(C)}},
\] (24)
\[
\beta^b = 1.
\] In summary, the leading eigenvalue and eigenvector of the interconnected network are approximated as
\[
\omega_1 = \omega_1^a + \delta \omega_1,
\]
\[
u = \beta^a v^a_1 + \beta^b v^b_1,
\] (25)
where \( \delta \omega_1, \beta^a \) and \( \beta^b \) are given in Eq. (23) and Eq. (24). Clearly when \( g/E(C) \to 0, \beta^a \to 1 \) and the two networks play equal role. Meanwhile if \( g/E(C) \gg 0 \), network \( a \) dominates \( b \). Normalizing \( u \) to unity and substituting back into (25) gives
\[
\langle \rho \rangle = \left( \lambda^* \omega_1^a + \lambda^* \delta \omega_1 - 1 \right)
\times \frac{\left( (\beta^a)^2 + 1 \right) \theta^a_1 + \left( (\beta^a)^2 + 1 \right) \theta^b_1}{N \left( (\beta^a)^2 \theta^a_3 + \theta^b_3 \right)}
\] (26)
where
\[
\theta^a_1 = \langle 1_{N_a}, v^a_1 \rangle, \quad \theta^b_1 = \langle 1_{N_a}, v^a_1 \cdot v^b_1 \rangle
\] (27)
and similarly for \( \theta^b_1, \theta^b_3 \).

With this approximation scheme, we can optimize \( \langle \rho \rangle \) over \( C \). Since \( \langle \rho \rangle \) depends on the interconnecting matrix \( C \) only in the form of \( E(C) \), it is sufficient optimize \( \langle \rho \rangle \) as a function of \( E(C) \). After determining the optimal \( E(C) \), we choose the matrix \( C \) such that \( E(C) \) is closest to its optimal value. Empirically, for all the networks considered in this study (see Sec. IV), \( \langle \rho \rangle \) is an increasing function of \( E(C) \) for small \( E(C) \). Thus it is sufficient to choose \( C \) such that \( E(C) \) is maximized, or in other words, to choose the top \( \delta M \) edges ranked by \( v^a_1(i) \times v^b_1(j) \) for \( i \in \{1, \ldots, N_a\}, j \in \{1, \ldots, N_b\} \). Here \( v^a_1(i) \) denotes the \( i \)th element of \( v^a_1 \) and similarly for \( v^b_1(j) \). Note that a node can be connected to multiple nodes in the other layer. After determine \( C \) in this way, we perform the check to ensure that \( \langle \rho \rangle \) is non-decreasing in the region \((0, E(C)]\). From Eq. (23), \( \delta \omega_1 \) is also an increasing function of \( E(C) \). Therefore, the optimal strategy also maximizes \( \delta \omega_1 \) as well as \( \langle \rho \rangle \). Since the outbreak threshold for the interconnected network is given by \( 1/(\omega_1^a + \delta \omega_1) \), as a consequence, the optimal inter-layer structure also minimizes the outbreak threshold. In the rest of the paper, the proposed strategy is called large eigenvector connecting (LEC), since it is by connecting inter-layer node pairs with larger product of the two nodes’ eigenvector centrality.

IV. RESULTS

In this section we test the strategy proposed in Sec. III both for synthetic and real-world networks. For synthetic networks, we consider scale-free networks generated by uncorrelated configuration model (UCM). Denote by ‘SF \( \alpha \)’ for scale-free network with degree distribution \( p(k) \sim k^{-\alpha} \) and degree exponent \( \alpha \) we consider the following three pairs of networks: (i) SF 3.0-SF 3.0, (ii) SF 3.0-SF 2.3, (iii) SF 3.0-SF 4.0 and two pairs of real world networks: (iv) Advogato-Facebook, (v) HepPh-HepTh. Here Advogato and Facebook are two online social networks, and HepPh and HepTh are citation networks of papers from the high energy physics-theory and high energy physics-phenomenology sections of the e-print arXiv. The networks are downloaded from [51]. TABLE I shows some basic statistics of the four real-world networks. When performing Monte-Carlo simulations on these networks, we first choose \( \lfloor 0.02N \rfloor \) nodes as infected seeds. Then the spreading runs at discrete time steps. At each time step, all the infected nodes infect each of their susceptible neighbors with probability \( \lambda \) independently and then returns to susceptible state with probability \( \mu \). The steps are then repeated until the density of infected nodes in the network reaches the stationary value.

TABLE I. Some basic statistics of four real-world networks. The statistics includes the number of nodes (\( N \)), the number of edges (\( M \)), the maximal degree (\( k_{\max} \)), the first moment of the degree distribution (\( \langle k \rangle \)), the clustering coefficient (\( c \)), the theoretical epidemic threshold predicted by \( \lambda^*_c = 1/\omega_1 \) and the modularity (\( Q \)).

| Network   | \( N \)  | \( M \)  | \( k_{\max} \) | \( \langle k \rangle \) | \( c \) | \( \lambda^*_c \) | \( Q \) |
|-----------|---------|---------|----------------|----------------|------|----------------|------|
| Advogato  | 5042    | 39227   | 803            | 15.56          | 0.092| 0.014          | 0.337|
| Facebook  | 2988    | 769     | 20644          | 0.0004         | 0.036| 0.809          |      |
| HepPh     | 420784  | 24463   | 0.013          | 0.408          |      |                |      |
| HepTh     | 352021  | 2468    | 0.269          | 0.328          |      |                |      |

First, we verify the accuracy of epidemic prevalence predicted by Eqs. (4). Define the ratio \( p = \delta M/M \). For the five network pairs considered, \( \lfloor pM \rfloor \) of edges are added uniformly random between the two networks. Starting with random initial conditions, we iterate Eqs. (4) until reaching a tolerated error. We perform Monte-Carlo simulations on the networks generated to compare with the theoretical predictions. Fig. 1 shows how \( \langle \rho \rangle \) changes as a function of \( \lambda \). The recovery rate is fixed to \( \mu = 0.5 \) and as well in the rest of the paper. The density of interconnecting edges is \( p = 0.01 \). In the insets of Fig. 1 the theoretical epidemic threshold \( \lambda^*_c = 1/\omega_1 \) is compared to the simulation results. The variability measure \( \Delta \) [53, 54] is applied to determine the
spreading threshold in simulations. In particular
\[ \Delta = \sqrt{\langle p \rangle^2 / \langle p \rangle^2} - 1, \]  
(28)
where the overline \( \overline{\cdot} \) denotes the average over independent runs of the simulations. As a convention \( \Delta = 0 \) when \( \langle p \rangle = 0 \). The simulated outbreak threshold is given by where the variability measure reaches its maximal. We mark the two thresholds by vertical lines both in Fig. 1 and the insets. From Fig. 1 for all the network pairs considered expect Advogato-Facebook (Fig. 1d), \( \langle p \rangle \) and the epidemic threshold predicted by Eqs. (4) agree with simulations in good accuracy. Meanwhile, for the network pair Advogato-Facebook, predictions of Eqs. (4) are less accurate. Since mean-field theory ignores dynamical correlations \[55\], i.e., ignores dependencies among the states of each node and its neighbors, it has been shown that mean-field theory could be less accurate for real-world networks with high clustering coefficient \[56\], high modularity \[57\] and low average degree \[58\]. As shown in TABLE I the network Facebook has high modularity and low average degree, therefore, this could be the possible origins that Eqs. (4) are less accurate for Advogato-Facebook. In general, quenched mean-field theory still predicts well the simulations, therefore can provide a necessary guarantee for the accuracy of the theory developed in Sec. III.

Next we test the accuracy of predictions for the leading eigenvalue \( \omega_1 \) and eigenvector \( u \) of the interconnected network. We consecutively add interconnecting edges between the two networks according to the LEC strategy and compare the eigenvalue and eigenvector predicted by Eqs. (25) to the true values \( \omega_1^* \) and \( u^* \). Start with two isolated networks, we consecutively add the edges and do the comparison after each adding. To quantify the accuracy, for the leading eigenvalue we consider the relative error \( \left( \omega_1^* - \omega_1 \right) / \omega_1^* \), where \( \omega_1^* \) is the true eigenvalue and \( \omega_1 \) is that predicted by Eqs. (25). For leading eigenvector the accuracy is measured by the cosine similarity between the true one \( u^* \) and predicted one \( u \). Let \( \theta \) be the angle between \( u^* \) and \( u \),
\[ \cos \theta = \langle u^*, u \rangle. \]  
(29)
Note that we have assumed both \( u \) and \( u^* \) are normalized to unity.

Still denote the fraction of edges added by \( p = \delta M / M \). \( \left( \omega_1^* - \omega_1 \right) / \omega_1^* \) and \( \cos \theta \) versus \( p \) are shown in Figs. 2(a) and (c) respectively, while \( \omega_1^* \) and \( \omega_1 \) versus \( p \) are shown in Fig. 2(b). From Figs. 2 the theoretical predictions gives close approximations to it’s true value. As can observed in Figs. 2(a) and (b), errors in the eigenvalue \( \omega_1^* - \omega_1 \) have a small magnitude compared to the value of \( \omega_1^* \), at least for small values of \( p \). Also \( \cos \theta \) is close to 1 which indicates \( u \) and \( u^* \) are closely aligned. After tested the accuracy of Eqs. (4) and Eqs. (25), the interconnecting strategy proposed in Sec. III is expected to be reliable.

Next we test the performance of the strategy of optimizing spreading. For comparison, the two following heuristic strategies are considered. (i) Large degree connecting (LDC). This is by choosing top candidate edges that ranked by the total degree of its two ends \( k_i + k_j \). (ii) Random connecting (RC). The edges are chosen uniformly random among all node pairs. Remind that the strategy proposed in Sec. III is called largest eigenvector connecting (LEC).

As the model depends on the parameter \( p \), we first consider comparing the three strategies with fixed \( p = 0.005 \). The results for different small values of \( p \) are similar as will be shown in the next. After choosing \( C \) according to the strategies, Eqs. (4) are iterated to stationary. For the convenience of visualization, the performance is evaluated by \( \delta(p) \), which is the difference of stationary \( \langle p \rangle \) after and before adding the interconnecting edges in \( C \). \( \delta(p) \) versus the infect rate \( \lambda \) are shown in Fig. 3 for the five pairs of networks. Still, the recovery rate is fixed to \( \mu = 0.5 \). The threshold value of \( \lambda_c \) for the isolated network \( G^0 \) is denoted by a black vertical line. For \( \lambda \) near the critical point, LEC strategy gives highest \( \delta(p) \) for all the five pairs of networks. For the three model network pairs in Figs. 3(a)-(c), LDC gives a very close performance to LEC. This is because that LEC and LDC predict similar \( C \) for these three network pairs, as the degree centrality and eigenvector centrality are strongly correlated in ranking for the UCM \[59\]. For real-world networks, the structure becomes complex; therefore, degree and eigenvector centrality are less correlated. Thus a better performance of LEC is observed near the critical point for the two real-world network pairs.

As \( \lambda \) becomes large and deviates the critical value, those nodes with large eigenvalue centrality have a high probability of to be infected. Thus interconnecting these nodes becomes unnecessary. Meanwhile, high degree nodes still play a central role in maintaining the epidemic. Thus gradually, better performance of the LDC strategy is observed. Eventually, when \( \lambda \) become very large, the infected state is prevalent; therefore, all central nodes have a high probability to be infected. Meanwhile, the RC strategy more likely picks low degree nodes due to the power-law degree distribution. Thus in this region of \( \lambda \), RC is expected to give the best performance among the three.

To get a more comprehensive picture, we study how the performance of the three strategies depend on \( p \) and \( \lambda \). For each point in the parameter space of \( p \) and \( \lambda \), we compute \( \langle p \rangle \) for the three strategies and find the differences of LEC minus LDC and LEC minus RC. In other words, a positive difference indicates that LEC performs better than the other one compared. The results are shown in Fig. 4. The first row in Fig. 4 i.e. Fig. 4(a1)-(e1), corresponds to LEC minus LDC for the five network pairs, and the second row (Figs. 4(a2)-(e2)) correspond to LEC minus RC. The contours of \( \epsilon = 1e-6 \) are shown in the plots to indicate the region when LEC performs better. Here \( \epsilon \) is not set to be \( \epsilon = 0 \) to eliminate possible numeric errors. In Fig. 4 the position of \( \lambda_c = \mu / \omega_1^0 \) for the unconnected network is marked by a vertical black dashed line. For all the cases in Fig. 4 there is a vertical band in the parameter space where LEC gives better performance. The vertical band locates in where \( \lambda \) is close to the critical value. Thus we can conclude that LEC is better than the other two near the critical point when the number of added edges is small.
FIG. 1. (Color online) The stationary epidemic size $\langle \rho \rangle$ as a function of the infection rate $\lambda$ with fixed recovery rate $\mu = 0.5$ and inserted edge density $p = 0.01$. $\lambda$ is changed from 0.01 to 1 at steps of 0.01. The network pairs considered are (a) SF 3.0-SF 3.0, (b) SF 3.0-SF 2.3, (c) SF 3.0-SF 4.0, (d) Advogato-Facebook and (e) HepPh-HepTh. The orange circles correspond to the Monte Carlo simulations of $\langle \rho \rangle$ and the purple vertical lines denote the $\lambda_c$ estimated from simulations. The blue line is by iterating the Eqs. (4), and the yellow dashed vertical line is the theoretical threshold. The gray line in the insets shows how the variability measure changes with $\lambda$ near the threshold. The variability measure is obtained by 500 independent runs of Monte Carlo simulations. Note that in the main plot of (d) and (e), $\lambda_c$ and $\mu/\omega_1$ might be too close to distinguish.

V. DISCUSSION

In the study, we investigated the optimal strategy for interconnecting two isolated networks in order to maximize the spreading prevalence of SIS model. We develop a scheme to approximate the leading eigenvalue and eigenvector for the interconnected network. This approximation gives a formula that predicts the epidemic prevalence for the interconnected network. By maximizing the spreading prevalence over the interconnecting matrix, we obtain the optimal inter-layer structure. By numerically iterating the discrete-time Markov equations, we find the strategy gives better performance than some other heuristic strategies.

By maximizing the spreading prevalence $\langle \rho \rangle$ among all interconnecting matrices, it turns out that the strategy is equivalent to select edge candidates that are top-ranked by the product of eigenvector centrality of nodes in its two ends. Actually, for the networks considered in this study, $\langle \rho \rangle$ is not always an increasing function of $E(C)$ for large $E(C)$ and large $\lambda$. Nevertheless, in this region, the number of edges added must be very large, and the infection probability is far from critical; as a consequence, the eigenvalue approximation scheme becomes inaccurate. What is a better strategy when adding a large number of edges cannot be answered by the current approach.

Besides, the current method mainly focuses on the region near the epidemic threshold. In summary, we mainly work in regions where linear approximations are reliable. For larger effective spreading rates and larger numbers of added edges, the problem might become nonlinear and more complex. If it is possible to get a more comprehensive theory in all parameter regions requires further studies.

ACKNOWLEDGMENTS

This work was partially supported by National Natural Science Foundation of China (Nos. 61673150, 11622538)
FIG. 2. (Color online) The prediction accuracy of the leading eigenvalue and eigenvector of the interconnected network. The interconnected network is obtained by applying LEC to add edges between the two layers. (a) The relative error \((\omega_1^* - \omega_1)/\omega_1^*\) between the true eigenvalue \(\omega_1^*\) and the theoretical prediction \(\omega_1\) versus \(p\). (b) \(\omega_1^*\) and \(\omega_1\) versus \(p\). (c) The cosine similarity between the true leading eigenvector and the theoretical prediction \(u\). Different lines correspond to different network pairs. The fraction of edges added \(p = \delta M/M\) starts from 0.001 to 0.01 at steps of 0.001. In (b) the solid lines correspond to the theoretical predictions \(\omega_1\) while the circles correspond to the true values \(\omega_1^*\). In (b), the data points for SF 3.0-SF 3.0 and SF 3.0-SF 4.0 are highly overlapped.

[1] S. Boccaletti, G. Bianconi, R. Criado, C. I. Del Genio, J. Gómez-Gardeñes, M. Romance, I. Sendiña-Nadal, Z. Wang, and M. Zanin, Physics Reports 544, 1 (2014).
[2] J. Gao, S. V. Buldyrev, H. E. Stanley, and S. Havlin, Nature Physics 8, 40 (2012).
[3] M. Kivelä, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter, Journal of Complex Networks 2, 203 (2014).
[4] M. De Domenico, C. Granell, M. A. Porter, and A. Arenas, Nature Physics 12, 901 (2016).
[5] M. Cai, W. Wang, Y. Cui, and H. E. Stanley, Physica A 490, 1 (2018).
[6] A. Solé-Ribalta, S. Gómez, and A. Arenas, Physical Review Letters 116, 108701 (2016).
[7] C. Buono and L. A. Braunstein, Europhysics Letters 109, 26001 (2015).
[8] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, Nature 464, 1025 (2010).
[9] E. Cozzo, R. A. Banos, S. Meloni, and Y. Moreno, Physical Review E 88, 050801 (2013).
[10] J. Y. Kim and K.-I. Goh, Physical Review Letters 111, 058702 (2013).
[11] K.-M. Lee, J. Y. Kim, W.-k. Cho, K.-I. Goh, and I. Kim, New Journal of Physics 14, 033027 (2012).
[12] Z. Wang, L. Wang, A. Szolnoki, and M. Perc, European Physical Journal B 88 (2015), [1504.04359]
[13] X. Zhang, S. Boccaletti, S. Guan, and Z. Liu, Physical Review Letters 114, 038701 (2015).
[14] L. D. Valdez, P. A. Macri, H. E. Stanley, and L. A. Braunstein, Physical Review E 88, 050803 (2013).
[15] M. Di Muro, C. La Rocca, H. Stanley, S. Havlin, and L. Braunstein, Scientific Reports 6, 22384 (2016).
[16] M. Di Muro, S. Buldyrev, H. Stanley, and L. Braunstein, Physical Review E 94, 042304 (2016).
[17] L. Valdez, M. Di Muro, and L. Braunstein, Journal of Statistical Mechanics: Theory and Experiment 2016, 093402 (2016).
[18] R. Cohen, D. Ben-Avraham, and S. Havlin, Physical Review E 66, 036113 (2002).
[19] M. Á. Serrano and M. Boguná, Physical Review Letters 97, 088701 (2006).
[20] A. Goltsvev, S. Dorogovtsev, and J. Mendes, Physical Review E 78, 051105 (2008).
[21] M. Perc, J. J. Jordan, D. G. Rand, Z. Wang, S. Boccaletti, and A. Szolnoki, Physics Reports 687, 17 (2017).
[22] Z. Wang, A. Szolnoki, and M. Perc, Scientific Reports 3...
The values of $\lambda$ are chosen near the critical point thus depending on the specific network pairs considered. For the three synthetic network pairs, $\lambda$ is chosen from 0.002 to 0.2 with steps of 0.002. For the real world networks $\lambda$ starts from 0.0002 and ends at 0.02 with steps of 0.0002. The blue solid lines correspond to LEC, the orange dashed lines to LDC and the yellow dash-dot line to RC. The black vertical lines mark the critical value of $\lambda_c = \mu/\omega^a$ of the unconnected network $G^o$.
FIG. 4. (Color online) The difference of $\langle \rho \rangle$ for LEC−LDC in the first row ((a1)-(e1)) and LEC−RC in second row ((a2)-(e2)). Each one of the five columns corresponds to one of the the network pairs, which are (a1-2) SF 3.0-SF 3.0, (b1-2) SF 3.0-SF 2.3, (c1-2) SF 3.0-SF 4.0, (d1-2) Advogato-Facebook and (e1-2) HepPh-HepTh. The fraction of added edges $p$ is changed from 0.0001 to 0.01 at steps of 0.0001. The values of $\lambda$ are chosen close to the critical point and depends on specific network pairs. For the three synthetic network pairs, $\lambda$ is chosen from 0.002 to 0.2 with steps of 0.002. For the real-world network pairs $\lambda$ starts from 0.0002 and ends to 0.02 at steps of 0.0002. The while lines correspond to the contour of $\epsilon = 1e^{-6}$, which bounds the region where LEC performs better than the other by at least $\epsilon$. The vertical black dashed lines indicate the position of $\lambda_c$ for the unconnected network.

[51] http://konect.uni-koblenz.de/networks/
[52] M. E. Newman, Proceedings of the National Academy of Sciences 103, 8577 (2006).
[53] P. Shu, W. Wang, M. Tang, and Y. Do, Chaos 25, 063104 (2015).
[54] P. Shu, W. Wang, M. Tang, and P. Zhao, Chaos 26, 063108 (2016).
[55] W. Wang, M. Tang, H. E. Stanley, and L. A. Braunstein, Reports on Progress in Physics 80, 036603 (2017).
[56] F. Radicchi and C. Castellano, Physical Review E 93, 030302 (2016).
[57] W. Wang, Q.-H. Liu, L.-F. Zhong, M. Tang, H. Gao, and H. E. Stanley, Scientific Reports 6, 24676 (2016).
[58] J. P. Gleeson, S. Melnik, J. A. Ward, M. A. Porter, and P. J. Mucha, Physical Review E 85, 026106 (2012).
[59] L. Lü, D. Chen, X.-L. Ren, Q.-M. Zhang, Y.-C. Zhang, and T. Zhou, Physics Reports 650, 1 (2016).