Optimizing interconnections to maximize the spectral radius of interdependent networks

Huashan Chen, Xiuyan Zhao, Feng Liu, Shouhuai Xu, and Wenlian Lu

Phys. Rev. E **95**, 032308 — Published 7 March 2017

DOI: 10.1103/PhysRevE.95.032308

Optimizing inter-connections to maximize the spectral radius of interdependent networks

Huashan Chen,[1,2,3] Xiuyan Zhao,[2] Feng Liu,[1,*] Shouhuai Xu,[3,†] and Wenlian Lu[2,4,‡]

[1]*State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China*
[2]*School of Mathematical Sciences, Fudan University, Shanghai, P. R. China*
[3]*Department of Computer Science, University of Texas at San Antonio, USA*
[4]*Institute of Science and Technology for Brain-Inspired Intelligence, Fudan University, Shanghai, China*

(Dated: February 14, 2017)

Spectral radius (i.e., the largest eigenvalue) of the adjacency matrices of complex networks is an important quantity that governs the behavior of many dynamic processes on the networks, such as synchronization and epidemics. Studies in the literature focused on bounding this quantity. In this paper, we investigate how to maximize the spectral radius of interdependent networks by optimally linking k inter-network connections (or inter-connections for short). We derive formulas for the estimation of the spectral radius of interdependent networks and employ these results to develop a suite of algorithms that are applicable to different parameter regimes. In particular, a simple algorithm (Algorithm 2) is to link the k nodes with the largest k eigenvector centralities in one network to the node in the other network with a certain property related to both networks. We demonstrate the applicability of our algorithms via extensive simulations. We discuss the physical implications of the results, including how the optimal inter-connections can more effectively decrease the threshold of epidemic spreading in the Susceptible-Infected-Susceptible (SIS) model and the threshold of synchronization of coupled Kuramoto oscillators.

PACS numbers: 02.10.Ox, 89.75.Hc, 89.75.Fb

I. INTRODUCTION

In the last two decades, we witnessed a significant advance in understanding the structure and function of complex networks [1–4]. Many phenomena occurring on networks are now known to be affected by their structure, e.g., the absence of epidemic threshold in large scale-free networks [5–7]. The spectral radius ρ of the adjacency matrix of networks has emerged as a key quantity governing the properties of dynamical processes on networks, including the Susceptible-Infected-Susceptible (SIS) model [8–13], the Kuramoto type of synchronization of coupled oscillators [14], and percolation [15]. For example, the critical value of the coupling strength in a network of coupled oscillators is proportional to $1/\rho$ [14], and a network with a larger ρ reduces the critical value of the coupling strength more towards synchronization. For percolation in directed networks [15], the critical node removal probability is $1 - 1/\rho$, and a network with a larger ρ is more robust against random node removal. The importance of ρ has attracted a due amount of attention, but the focus was on bounding and approximating it [16–19]. Moreover, the issue of ρ has not been investigated in the context of *interdependent networks*, which aim to model the interactions between real-world complex networks (e.g., electricity and telecommunication infrastructures) [20–22].

In this paper, we study the spectral radius of interdependent networks from an optimization perspective: How can we maximize the spectral radius ρ of an interdependent network by optimally linking k inter-connections of weight α between two networks G_1 and G_2? The problem is interesting because ρ also governs the properties of dynamic processes occurring on interdependent networks, including the processes mentioned above. We use a parameter α to represent the coupling strength of the interconnection between G_1 and G_2 [23–25] (e.g., the infection rate associated to the inter-connections in epidemic model, or the coupling strength in the model of coupled oscillators). We stress that α can be arbitrary, and the case $\alpha = 1$ can be seen as the unweighted case. We use the Perturbation Theory to derive approximate expressions for the spectral radius of the resulting interdependent networks with respect to a small or large α, and use these approximations to identify the optimal inter-connections. The results are summarized as follows:

- In the case α is small (e.g., $\alpha \leq 1$), we identify a new algebraic property corresponding to a certain relation between networks G_1 and G_2, and then use this property to design an optimal algorithm for maximizing ρ. The algorithm is applicable when k is small (for example, when $k \leq 40$). For arbitrary k, we discover an interesting phenomenon that actually leads to a much faster algorithm.

- In the case α is large (e.g., $\alpha \geq 100$), we present a Genetic Algorithm (GA) that is applicable for arbitrary k. The algorithm leads to larger spectral radii than the other algorithms that use the random inter-connection strategy or various node-centrality based inter-connection strategies.

- The case that α is medium (in-between the small and large cases) can be seen as a (nonlinear) combination of the first two cases. We show that the algorithms that work for small or large α also work well in this case. In particular, we observe a threshold at $\alpha^* \approx 21$, above which the GA outperforms the algorithm that works for

* liufeng@iie.ac.cn
† shxu@cs.utsa.edu
‡ wenlian@fudan.edu.cn

smaller α and below which the GA is outperformed by the algorithm that works for smaller α.

It is worth mentioning that all the algorithms mentioned above are independent of the sizes of G_1 and G_2. Moreover, we use examples to show how the resulting ρ promotes the SIS spreading processes and accelerates the synchronization of coupled Kuramoto oscillators in interdependent networks.

II. OPTIMIZING SPECTRAL RADIUS OF INTERDEPENDENT NETWORKS

We use the following notations. I_n denotes the identity matrix of n dimensions. A network G with n nodes is represented by an $n \times n$ adjacency matrix $\mathsf{A} = [a_{i,j}]_{i,j=1}^n$, where $a_{i,j} = 1$ means nodes i and j are connected and $a_{i,j} = 0$ otherwise. We use $\rho(\mathsf{G})$ and $\rho(\mathsf{A})$ interchangeably for the spectral radius of adjacency matrix A or network G.

Let $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ be two connected (i.e., irreducible) undirected networks, where $V_1 = \{v_1, \ldots, v_m\}$ and $V_2 = \{u_1, \ldots, u_n\}$ at vertex sets, E_1 and E_2 are edge sets. Let E_* denote a set of k inter-connections between G_1 and G_2 with a uniform weight α, which can be an arbitrary positive real number. As mentioned above, α represents the strength of the couplings between G_1 and G_2, such as the infection rate associated to the inter-connections in epidemic model and the coupling strength in the model of coupled oscillators. Let $C = (c_{ij})_{m\times n}$ represent the inter-connections between G_1 and G_2 with

$$c_{ij} = \begin{cases} 1 & (i,j) \in E_* \\ 0 & \text{otherwise.} \end{cases}$$

Let $G = (V, E)$ denote the resulting interdependent network, where $V = V_1 \cup V_2$ and $E = E_1 \cup E_2 \cup E_*$. Let $A_1 = [a_{i,j}^{(1)}]_{i,j=1}^n$, $A_2 = [a_{i,j}^{(2)}]_{i,j=1}^n$ and $A = [a_{i,j}]_{i,j=1}^n$ respectively denote the adjacency matrix of G_1, G_2, and G. Assume further that both G_1 and G_2 are aperiodic (i.e., the largest common divisor of the lengths of all loops of each node is 1), implying that the algebraic dimensions of $\rho(A_1)$ and $\rho(A_2)$ are 1. Then,

$$A = \begin{bmatrix} A_1 & \alpha C \\ \alpha D & A_2 \end{bmatrix} = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} + \alpha \begin{bmatrix} 0 & C \\ D & 0 \end{bmatrix},$$

where $D = C^\top$ because G_1 and G_2 are undirected.

The optimization problem is formalized as: Given A_1, A_2 and α, find C with k nonzeros to maximize the spectral radius $\rho(A)$. In order to tackle this problem, we first derive analytic formulas for the spectral radius of interdependent networks via the perturbation approach, and then employ these formulas to design algorithms. In the present paper, we focus on the case $\rho(A_1) > \rho(A_2)$, while noting that the treatment of the case $\rho(A_1) < \rho(A_2)$ is equivalent. The investigation of the case $\rho(A_1) = \rho(A_2)$ is left as an open problem for future research, because it leads to a different optimization problem. Our investigation considers three cases of the inter-connection weight α: small, large, and medium.

A. The case of small inter-connection weight

Let $\psi = [\psi_1, \cdots, \psi_m]^\top$ be the right and left eigenvector of A_1 corresponding to $\rho(A_1)$ (noting that A_1 is symmetry and the algebraic dimension of $\rho(A_1)$ is assumed to be one), respectively. We permutate the indices of the nodes of G_1 such that $\psi_1 \geq \psi_2 \geq \cdots \geq \psi_m$ and re-scale them as $\sum_{i=1}^n \psi_i = 1$. Let $x^\top = \psi^\top C$ and $M = [\rho(A_1) I_n - A_2]^{-1} = [m_{ij}]_{n\times n}$. It can be proved that M has all elements positive because A_1 is irreducible (namely G_1 is connected) and $\rho(A_1) > \rho(A_2)$ (see [26]). For a sufficiently small inter-connection weight α, the perturbation approach leads to,

$$\rho(A) = \rho(A_1) + \alpha^2 x^\top M x + o(\alpha^2). \tag{1}$$

Please see Appendix A for the derivation of Eq.(1).

Since α is sufficiently small, the maximization of $\rho(A)$ implies the maximization of $\lambda_2 = x^\top M x$ in Eq.(1), which suggests us to investigate the numerical characteristics of M. Consider the relative difference between $\rho(A_1)$ and $\rho(A_2)$, denoted by

$$\mu = \frac{\rho(A_1) - \rho(A_2)}{\rho(A_2)},$$

where $\mu > 0$ because $\rho(A_1) > \rho(A_2)$. We observe that if $\rho(A_1) \gg \rho(A_2)$, then M is nearly diagonal. Define

$$\Gamma = \frac{\min\limits_{l=1,\ldots,n} m_{ll}}{\max\limits_{\substack{i\neq j \\ i,j=1,\ldots,n}} m_{ij}}.$$

Later we will numerically verify the fact that

$$\Gamma \sim b_* \mu, \tag{2}$$

where $b_* = \dfrac{\rho(A_2)}{\max_{i\neq j}(a_{i,j}^{(2)})}$. We also give a theoretical validation of Eq.(2) in Appendix B. This means that M is diagonally dominant because $\Gamma \gg 1$ when $\mu \gg 1$. The to-be-verified fact (2) leads to an approximation of λ_2 as follows. Let us rewrite λ_2 as

$$\lambda_2 = \Sigma_1 + \Sigma_2 \tag{3}$$

where $\Sigma_1 = \sum_{l=1}^n x_l x_l m_{ll}$ and $\Sigma_2 = \sum_{\substack{i,j=1,\cdots,n \\ i\neq j}} x_i x_j m_{ij}$. Then,

$$\Sigma_1/\Sigma_2 = \frac{\sum_{l=1}^n x_l x_l m_{ll}}{\sum_{\substack{i,j=1,\cdots,n \\ i\neq j}} x_i x_j m_{ij}} \geq \frac{\min(m_{ll}) \sum_{l=1}^n x_l^2}{\max(m_{ij}) \sum_{\substack{i,j=1,\cdots,n \\ i\neq j}} x_i x_j}.$$

Noting that there are at most k nonzero components in x, we have $\sum_{\substack{i,j=1\cdots,n \\ i\neq j}} x_i x_j \leq (k-1) \sum_{l=1}^n x_l^2$. This implies

$$\frac{\sum_{l=1}^n x_l^2}{\sum_{\substack{i,j=1\cdots,n \\ i\neq j}} x_i x_j} \geq \frac{1}{k-1}.$$

Therefore, we have

$$\Sigma_1/\Sigma_2 \geq \frac{\Gamma}{k-1}.$$

The to-be-verified fact (2) implies $\Sigma_1 \gg \Sigma_2$ when $\mu \gg 1$, namely

$$\lambda_2 \approx \Sigma_1 = \sum_{l=1}^{n} x_l^2 m_{ll}. \tag{4}$$

Therefore, in the case of $\mu \gg 1$, we can use Eq.(4) to reformulate the problem of maximizing λ_2 approximately as:

$$\begin{cases} \max\limits_{C=(c_{ij})} & \Sigma_1 = \sum_{l=1}^{n} x_l^2 m_{ll} \\ \text{Subject to} & c_{ij} \in \{0,1\} \\ & C \text{ has } k \text{ nonzero elements.} \end{cases} \tag{5}$$

Since $x^\top = \psi^\top C$, C should be selected to keep the largest components of ψ (i.e., $\psi_1, \cdots, \psi_k$) and the largest diagonal elements in M as in the product term of Σ_1 as possible. We pick the k largest diagonal elements in M, denoted by $M_{j_1,j_1} \geq M_{j_2,j_2} \geq \cdots \geq M_{j_k,j_k}$. We take the $1st, \cdots, k\text{-}th$ rows and the $j_1th, \cdots, j_k\text{-}th$ columns of matrix C to comprise a submatrix of C, denoted by $\hat{C} = (\hat{c}_{pq}) \in R^{k,k}$ with $\hat{c}_{pq} = c_{p,j_q}$ for all p, q. As shown in the Appendix C, the following rule for C is necessary to maximize Σ_1.

Rule 1. *Matrix C should satisfy:*

- $c_{ij} = 0$ *for all* $i \notin I$ *or* $j \notin J$;
- $\hat{c}_{p,q} = 1$ *implies* $\hat{c}_{p',q'} = 1$ *for all* $p' \leq p$ *and* $q' \leq q$.

In fact, if there are some p, q and p', q' with $p' \geq p$ and/or $q' \geq q$ such that $\hat{c}_{p,q} = 1$ but $\hat{c}_{p',q'} = 0$, changing their values with $\hat{c}_{p',q'} = 1$ and $\hat{c}_{p,q} = 0$ (but keep the others intact) will lead to an increase in Σ_1. Please see Appendix C for an illustrative example of $k = 3$. Denote by $\mathcal{C}^{n,k}$ the set of all square matrices (of n dimensions) satisfying Rule 1. We observe from Rule 1 that the definition of $\mathcal{C}^{n,m,k}$ is self-contained and independent from the graph structures of G_1 and G_2. Hence, the cardinality of $\mathcal{C}^{n,m,k}$ is independent of the graph structures of G_1 and G_2. In addition, from Rule 1, the nonzero elements $\hat{c}_{p,q}$ of C should satisfy $p \leq k$ and $q \leq k$, which implies that these nonzero elements are located in the first k-rows and first k-columns. In the more interesting cases of $\min\{n, m\} > k$, $\mathcal{C}^{n,m,k}$ is independent of the sizes of G_1 and G_2, namely $\mathcal{C}^{k,k,k} = \mathcal{C}^{m,n,k}$ for all $m \geq k$ and $n \geq k$. In this case, we can denote $\mathcal{C}^{n,m,k}$ as $\mathcal{C}^k$ instead. For example, when $k = 3$, we have $\mathcal{C}^3$ consisting of

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \text{ and } \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

The preceding discussion leads to Algorithm 1, which selects the matrix C that maximizes λ_2 from all $C \in \mathcal{C}^k$.

Algorithm 1 Maximize $\rho(A)$ according to (5) via numerical features

Input: Adjacency matrices A_1, A_2; Link weight α; # of interconnections k
Output: $\rho(A)$, $\rho(A_1) + \alpha^2\lambda_2$

```
compute A_1's spectral radius ρ(A_1) and the corresponding left eigenvector ψ^⊤; compute matrix M = [ρ(A_1)I_n − A_2]^{−1}
initialize C to be a m * n zero matrix, and set k_edges ← ∅
compute the set P of all of the integer partitions of k
for each partition ξ ∈ P do
    let t be the maximum value in partition ξ
    find the largest t values in the diagonal elements of M
    return their indices as a set ζ
    for j ∈ ζ do
        let s be the number of nonzero elements in ξ
        find the largest s elements of eigenvector ψ^⊤
        return their indices as a set η
        for i ∈ η do
            C[i, j] ← 1
            add edge (i, j) to set k_edges
        end for
        for g ∈ ξ do
            if g > 0 then
                g ← g − 1
            end if
        end for
    end for
    compute ρ(A), ρ(A_1) + α^2 λ_2
end for
return the maximum ρ(A), ρ(A_1) + α^2 λ_2 among all partitions
```

The computational complexity of Algorithm 1 can be measured by the cardinality of $\mathcal{C}^k$ when $\min\{m, n\} > k$. The cardinality of $\mathcal{C}^k$ is equivalent to the integer partitions of k. A partition of a positive integer k is defined to be a sequence of positive integers whose sum is k. We let the function $f(k)$ denote the number of partitions of the integer k. As an example, $f(4) = 5$, and here are all 5 of the partitions of the integer $k = 4$:

$$\begin{aligned} 4 &= 4 \\ 4 &= 3 + 1 \\ 4 &= 2 + 2 \\ 4 &= 2 + 1 + 1 \\ 4 &= 1 + 1 + 1 + 1 \end{aligned}$$

We can use the following function $f(k, p)$ to count the partitions of k:

$$f(k,p) = \begin{cases} 1 & (k = 1) \\ 1 + f(k, k-1) & (k = p) \\ f(k-p, p) + f(k, p-1), & (k > p) \end{cases} \tag{6}$$

Hence, the cardinality of $\mathcal{C}^k$ can be obtained as $|\mathcal{C}^k| = f(k) = f(k, k)$, which was given by Hardy and Ramanujan [27] and independently by Uspensky[28, 29] as:

$$f(k) \approx \frac{e^{c*\sqrt{k}}}{4 * k * \sqrt{3}},$$

where c is given as:

$$c = \pi * (2/3)^{1/2}$$

This is, however, just an approximation to $f(k)$.

In order to get a concrete value k with which Algorithm 1 is practical, let us compare the computational complexity $\mathcal{C}^k$ with the polynomial k^3 as a concrete example. In this example, FIG.1 shows that when $k \leq 40$, Algorithm 1 is faster than polynomial k^3 and the exponential complexity e^k that is incurred by the brute-force method. That is, when $k \leq 40$, Algorithm 1 is practical; when $k > 40$, we need to design more efficient algorithms. We should stress that the number 40 is specific to the upper-bounding computational complexity that is considered practical, which is k^3 in this example. In other words, this number will vary according to the upper-bounding computational complexity, which reflects the available computer resource.

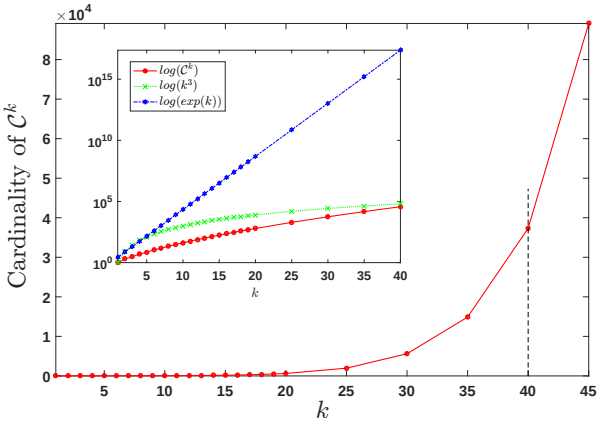


FIG. 1. Computational complexity of Algorithm 1.

In order to verify the results of Algorithm 1, we consider four kinds of complex network models as follows.

- REG(n, m): denoting regular a network with n vertices of degree m [30];
- ER(n, p): denoting an ER random network with n nodes and edge probability p [31];
- WS(n, m, p): denoting a WS small-world network with n nodes with each node having m nearest neighbors and rewiring probability p [32];
- BA(n, m): denoting a BA scale-free network with n nodes and m new edges being added at each time step [33].

Since the to-be-verified fact (2) serves as a foundation for Algorithm 1, let us first confirm it, namely that $\mu \to \infty$ implies $\Gamma \sim b_*\mu$, we consider a network generated by WS$(n, 0.2n, 0.2)$. FIG.2 shows that as μ increases, Γ/μ converges to the horizontal line $\rho(A_2)/\max(a^{(2)}_{i,j}), i \neq j$. The same simulation results hold for the other three network models (data are not shown here).

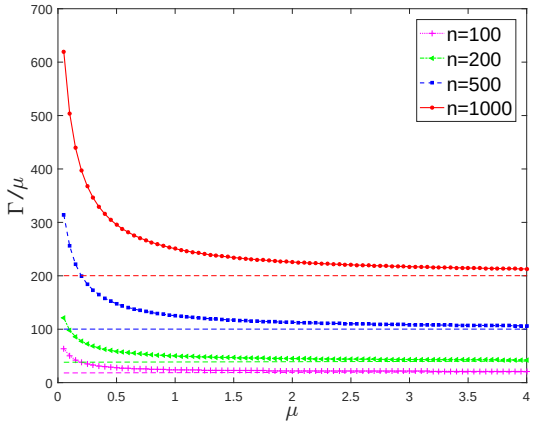


FIG. 2. Confirming the fact (2) via network WS$(n, 0.2n, 0.2)$ with $n = 100, 200, 500, 1000$.

Now we verify the results of Algorithm 1. First, we compare the spectral radius resulting from Algorithm 1 and the spectral radius resulting from the theoretical approximation detailed in Eq.(A7) of Appendix A. We plot the simulation result in FIG.3 with $k = 40$, G_1=ER$(100, 0.8)$ and G_2=ER$(100, 0.4)$, while noting that the same effect is observed in the other scenarios of G_1 and G_2. We observe that the spectral radius resulting from the theoretical approximation Eq.(A7) is in an excellent agreement with the spectral radius resulting from Algorithm 1 for $\alpha \leq 10^0$ (i.e., small α). However, this agreement disappears for large α (e.g., $\alpha \geq 10^1$).

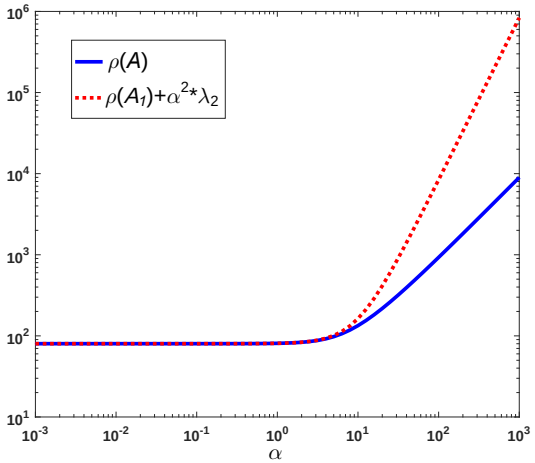
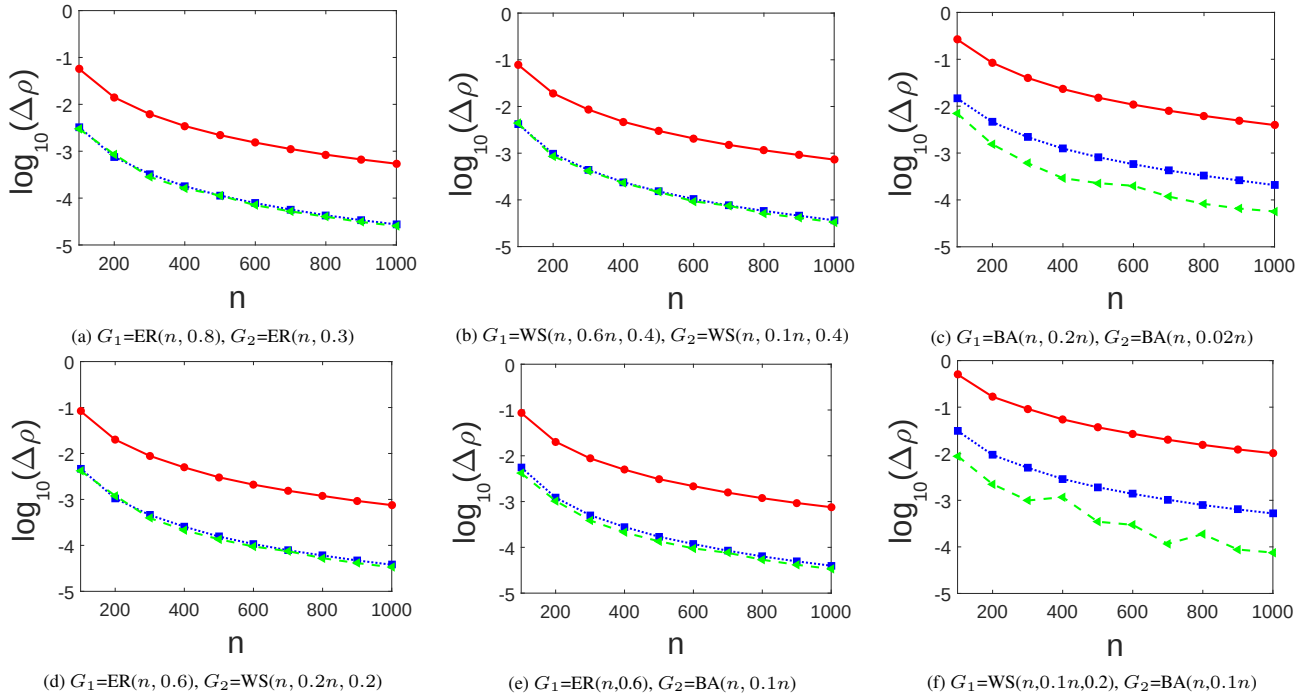


FIG. 3. Comparison between the spectral radius resulting from Algorithm 1 (the solid blue curve) and theoretical approximation of Eq.(A7) (the dashed red curve) with respect to α, where G_1 =ER(100,0.8), G_2 =ER(100,0.4), and $k = 40$.

Second, in order to show that Algorithm 1 maximizes $\rho(A)$, we consider two alternate algorithms: the Random algorithm that picks k inter-connections uniformly at random; the Degree centrality algorithm that connects the k highest-degree nodes in G_1 respectively to the k highest-degree nodes in G_2.



(a) G_1=ER$(n, 0.8)$, G_2=ER$(n, 0.3)$ (b) G_1=WS$(n, 0.6n, 0.4)$, G_2=WS$(n, 0.1n, 0.4)$ (c) G_1=BA$(n, 0.2n)$, G_2=BA$(n, 0.02n)$

(d) G_1=ER$(n, 0.6)$, G_2=WS$(n, 0.2n, 0.2)$ (e) G_1=ER$(n,0.6)$, G_2=BA$(n, 0.1n)$ (f) G_1=WS$(n,0.1n,0.2)$, G_2=BA$(n,0.1n)$

FIG. 4. Comparing Algorithm 1 to the two alternate algorithms via the difference of the spectral radius of the resulting interdependent network and the maximum of the two networks G_1 and G_2 (indicating the increase in spectral radius because of the inter-connections), namely $\Delta\rho = \rho(A)$-max$(\rho(A_1), \rho(A_2))$ with respect to n for small α. The solid red curve represents Algorithm 1, the dotted blue curve represents the Degree centrality algorithm, and the dashed green curve represents the Random algorithm, where $k = 20$, $\alpha = 1$, G_1 and G_2 are indicated in the captions.

Since G_1 and G_2 are generated by the ER, WS and BA network models, there are six combinations for the resulting interdependent networks.

The network parameters are set to meet the condition that $\mu >> 0$. We illustrate the result via the case of $m = n$, while noting that the general case ($m \neq n$) can be treated similarly. FIG.4 plots the simulation results with $\alpha = 1$, averaged over 100 simulation runs. We observe that Algorithm 1 always leads to a much larger $\rho(A)$ than the other algorithms, while noting that the Random algorithm leads to the smallest $\rho(A)$.

Dealing with the case of arbitrary k. Although Algorithm 1 leads to optimal results when k is small (e.g., $k < 40$ in the example mentioned above, it becomes infeasible when k is large (e.g., $k \geq 40$ in the example). Fortunately, we observed an interesting phenomenon in the simulation study: The '1' elements in the adjacency matrix C generated by Algorithm 1 always stay on the same column, coinciding with the index of the largest diagonal element of matrix M. In other words, the matrix C generated by Algorithm 1 exhibits a star structure. Moreover, we observed that the larger the networks, the more obvious the phenomenon.

The above phenomenon can be explained as follows. FIG.5 shows that the relative difference between the k largest components of ψ, $\psi_1, \ldots, \psi_k$ decreases with the network size n, meaning that for $n \gg k$,

$$\psi_1 \approx \psi_2 \approx \ldots \approx \psi_k. \tag{7}$$

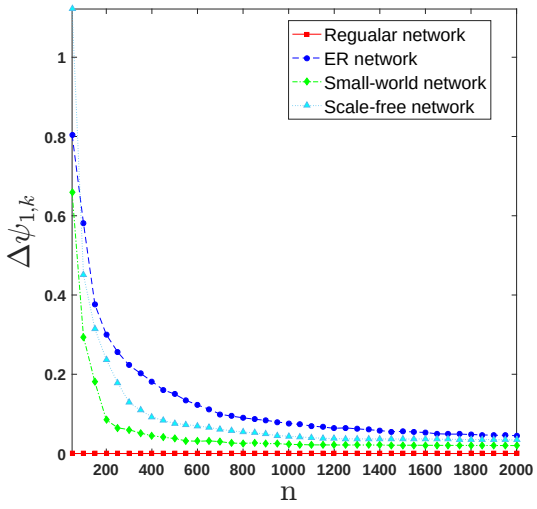


FIG. 5. Confirming assumption (7) via the relative difference between ψ_1 and ψ_k in four network models: REG$(n, 0.4n)$, ER$(n, 0.1)$, WS$(n, 0.2n, 0.2)$, and BA$(n, 0.2n)$, where $\Delta\psi_{1,k} = (\psi_1^2 - \psi_k^2)/\psi_k^2$, $k = 10$.

This observation implies that in order to maximize Σ_1, all of the nonzero components of x should contribute to the largest diagonal element of M, namely, M_{j_1,j_1}. In other words, the case $\hat{C}_{1,j_q} = 1$ for all $q = j_1, \cdots, j_k$ and $\hat{C}_{p,q} = 0$ otherwise can maximize Σ_1 approximately. That is, when n is large

and k is arbitrary (especially for a large k, such as $k \geq 40$ in the example mentioned above, that cannot be handled by Algorithm 1), we can find (or approximate) the optimal inter-connections by connecting the nodes in G_1 with the indices associated to the largest k components of ψ to the single node in G_2 with the index associated to the largest diagonal element of M. This insight leads to Algorithm 2, which is much faster than Algorithm 1.

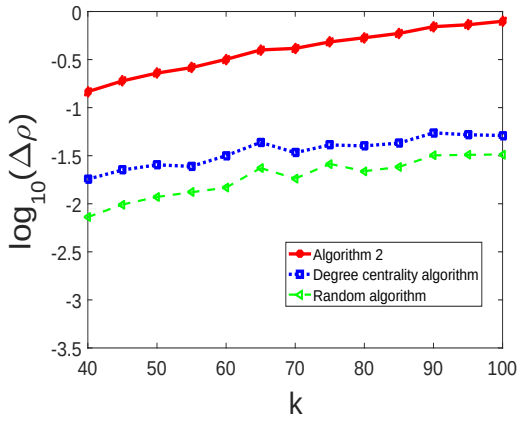
Algorithm 2 Maximize $\rho(A)$ for arbitrary k

Input: Adjacency matrices A_1, A_2; # of inter-connections k;
Output: matrix C, $\rho(A)$

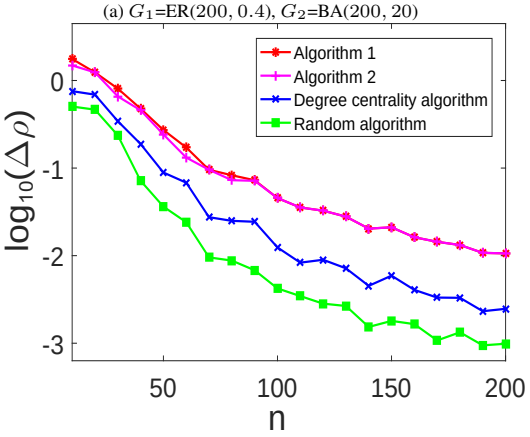
1: compute $\rho(A_1)$ and the corresponding left eigenvector $\psi^\top$; compute matrix $M = [\rho(A_1)I_n - A_2]^{-1}$
2: initialize C as a $m * n$ zero matrix, $k\_edges \leftarrow \emptyset$
3: find the largest diagonal element of M, return its index as j
4: find the largest k absolute values in the eigenvector ψ, return their indices as set η
5: **for** $i \in \eta$ **do**
6: $\quad C[i, j] \leftarrow 1$
7: $\quad$ add (i, j) to k_edges
8: **end for**
9: compute $\rho(A)$ corresponding to matrix C
10: return matrix C , $\rho(A)$

For a large k (such as $k \geq 40$ in the example mentioned above, in which case Algorithm 1 is not feasible), we compare Algorithm 2 with the Random and Degree centrality algorithms mentioned above. FIG.6(a) plots the simulation result with G_1=ER(200,0.4) and G_2=BA(200,20), while noting that the same effect is observed in the other scenarios of network models. We observe that Algorithm 2 always leads to the largest spectral radius in the example case of $k \geq 40$. For the case of small k, we consider $k = 10$ and compare Algorithm 2, Algorithm 1, the Random algorithm, and the Degree centrality algorithm. FIG.6(b) shows that Algorithm 1 always leads to the largest $\rho(A)$, and that Algorithm 2 perfectly coincides with Algorithm 1 except for some $n \in (10, 100)$ in which case the difference is still very small, where G_1=ER(n,0.4) and G_2=BA(n,0.2n). It is clear that both Algorithm 1 and Algorithm 2 perform much better than the Degree centrality algorithm and the Random algorithm.

We conclude that when n is large, for any k (including the case of a small k or $k \leq 40$ in the example mentioned above), Algorithm 2 is as good as Algorithm 1 but incurs a much smaller computational complexity. It is also worth mentioning that we considered other centrality notions (e.g., the closeness centrality, the eigenvector centrality), and found they lead to the same results. Moreover, the effect of the Betweenness centrality algorithm coincides with the Degree Centrality algorithm. Finally, it is not surprising to see that the Random algorithm gives the worst result.



(a) G_1=ER(200, 0.4), G_2=BA(200, 20)



(b) G_1=ER(n, 0.4), G_2=BA(n, 0.1n)

FIG. 6. (a) Comparing the spectral radius $\Delta\rho = \rho(A)$-max$(\rho(A_1), \rho(A_2))$ of Algorithm 2 and the Random and Degree centrality inter-connection algorithms mentioned above, where $k \geq 40$, G_1=ER(200, 0.4), G_2=BA(200, 0.1n), and $\alpha = 1$. (b) Comparing the spectral radius $\Delta\rho = \rho(A)$-max$(\rho(A_1), \rho(A_2))$ of Algorithms 1-2 and the Random and Degree-centrality inter-connection algorithms, where $k = 10$, G_1=ER(n, 0.4), G_2=BA(n, 0.1n), and $\alpha = 1$.

B. The case of large inter-connection weight

In the case of sufficiently large α, we can rewrite matrix A as

$$A = \alpha \left(\epsilon \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} + \begin{bmatrix} 0 & C \\ C^\top & 0 \end{bmatrix} \right) = \alpha \tilde{A} \tag{8}$$

where $\epsilon = 1/\alpha$. We want to estimate $\rho(\tilde{A})$ as $\varepsilon \to 0$. As detailed in Appendix D, the spectral radius of the interdependent networks can be expressed as

$$\rho(A) = \alpha\sqrt{\mu_0} + \mu_1 + o(1/\alpha), \tag{9}$$

where μ_0 is the spectral radius of matrix $C^\top C$, and

$$\mu_1 = \frac{a^\top A_1 a + b^\top A_2 b}{a^\top a + b^\top b} \tag{10}$$

with $\begin{bmatrix} a \\ b \end{bmatrix}$ being the right eigenvector of matrix $Z = \begin{bmatrix} 0 & C \\ C^\top & 0 \end{bmatrix}$ corresponding to the largest eigenvalue $\sqrt{\mu_0}$. As a matter of fact, a is the right eigenvector of $CC^\top$ and b is the right eigenvector of $C^\top C$ corresponding to the largest eigenvalue.

In order to maximize $\rho(A)$, we need to consider two optimization problems instead. First, we need to maximize μ_0 because it has the highest order term:

$$\begin{cases} \max\limits_{C} & \rho(C^\top C) \\ \text{Subject to :} & c_{ij} \in \{0,1\} \\ & C \text{ has } k \text{ nonzero elements.} \end{cases} \tag{11}$$

This problem clearly has multiple solutions. Second, we need to maximize μ_1 for all the solutions of (11), which can be written as:

$$\begin{cases} \max\limits_{C} & \mu_1 \\ \text{Subject to:} & C \in \text{solution(11)} \end{cases}$$

where "solution(11)" represents the set of solutions to (11).

It can be proved that $\mu_0 = \rho(C^\top C) = k$ if and only if either C has a star structure centered at some node k_0 with links to some $Q = \{q_1, \ldots, q_k\}$, or C has an inverse star structure centered at some mode k_0 with links to Q.

In the case C has a star structure with $Q = \{1, \cdots, k\}$ and $k_0 = 1$ (without loss of generality), $CC^\top$ and $C^\top C$ have the following structures:

$$CC = \begin{bmatrix} ones(k,k) & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, C^\top C = \begin{bmatrix} k & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},$$

where "$ones(p,q)$" stands for a matrix $\in \mathbb{R}^{p,q}$ whose elements are all equal to 1. Thus, we have $b = (b_1, \cdots, b_n)^\top$ with $b_i = 1$ when $i = k_0$ and $b_i = 0$ otherwise; we also have $a = (a_1, \cdots, a_m)^\top$ with $a_i = 1$ if $i \in Q$ and $a_i = 0$ otherwise. Now, Eq.(10) becomes

$$\mu_1 = \frac{\sum_{i \in Q, j \in Q} a^{(1)}_{i,j} + a^{(2)}_{k_0 k_0}}{k+1}. \tag{12}$$

In the case C has an inverse star structure, a similar reasoning makes Eq.(10) become

$$\mu_1 = \frac{\sum_{i \in Q, j \in Q} a^{(2)}_{i,j} + a^{(1)}_{k_0 k_0}}{k+1}. \tag{13}$$

Therefore, the maximum value of μ_1 is

$$\max_{k_0, |Q|=k} \left(\frac{\sum_{i \in Q, j \in Q} a^{(1)}_{i,j} + a^{(2)}_{k_0 k_0}}{k+1}, \frac{\sum_{i \in Q, j \in Q} a^{(2)}_{i,j} + a^{(1)}_{k_0 k_0}}{k+1} \right), \tag{14}$$

and the optimal inter-connection links formulate a star-like or inverse star-like structure according to the discussion above. In the case that G_1 and G_2 do not have self-links, meaning $a^{(1)}_{kk} = a^{(2)}_{kk} = 0$ for all k, the maximal values of μ_1 is independent of the selection of k_0. Thus, the optimal solution corresponding to the subset-sum maximization problem over Q.

We now present a Generic Algorithm (GA) for identifying the k inter-connection links according to (14). The algorithm searches over the set J as follows: (i) the fitness is defined by Eq.(14); (ii) Crossover is realized by picking one half elements in Q from one parent and replacing them with those of the other parents; and (iii) mutation is realized by randomly picking elements with probability 0.1 and replacing them with the other elements that are randomly chosen.

Algorithm 3 Generic Algorithm (GA)

Input: Adjacency matrices A_1, A_2; # of inter-connections k;
Output: matrix C, $\rho(A)$

```
initialize a number T of iterations, a size S of population, a rate P_c of crossover, a rate P_m of mutation
initialize Q as a m * 1 (and (n * 1)) vector by randomly picking k components as 1 and the others zero.
initPop()
for i = 1 : T do
    newpop=selection(pop)
    newpop=crossover(newpop, P_c, S)
    newpop=mutation(newpop, P_m, S)
    pop=newpop
end for
result=best(pop)
return matrix C (a star or inverse star structure), ρ(A)
```

To verify Algorithm 3, we consider star or inverse star structure for C via:

- Random algorithm for large α: Randomly pick k_0 and J with equal probability;
- Degree centrality algorithm for large α: Pick k_0 and J according to the maximum weight degree (sum).

All picking operations are done without replacement. We then take the maximum value of μ_1 from the star and inverse star structures.

FIG.7 compares the spectral radius resulting from the optimal choices of inter-connection links according to Algorithm 3 and the spectral radius resulting from the theoretical approximation of Eq.(9). We observe that Eq.(9) gives an excellent approximation of the spectral radius $\rho(A)$ for a large α ($\alpha \geq 10^2$).

FIG.8 shows that Algorithm 3 corresponding to Eq.(14) always leads to the largest spectral radius when compared with the other two algorithms mentioned above. The simulation results are averaged over 100 simulation runs.

C. The case of medium inter-connection weight

For a small α ($\alpha \leq 10^0$), Algorithm 1 works well for a small k (e.g., $k \leq 40$ in the example mentioned above) and Algorithm 2 work well for arbitrary k. For a large α ($\alpha \geq 10^2$), Algorithm 3 works well. In what follows, we consider the case of medium α, namely $10^0 \leq \alpha \leq 10^2$.

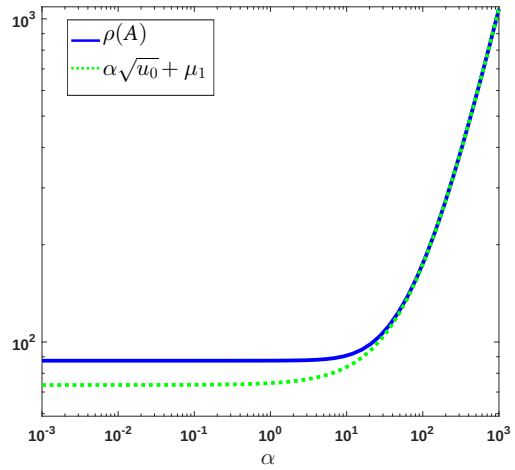


FIG. 7. Comparison between the spectral radius resulting from the Genetic algorithm (solid blue curve) and the spectral radius resulting from the theoretical approximation of Eq.(9) (dashed green curve), where G_1 =ER(100,0.6), G_2 =WS(100,40,0.4), and $k = 80$.

For $\alpha \in [10^{-3}, 10^0]$, Fig.9(a) shows that Algorithms 2 and 3 always perform better than their respective alternate algorithms. For $10^0 \leq \alpha \leq 10^2$, Fig.9(b) shows that Algorithm 3 is as good as Algorithm 2, both of which perform better than their two alternate algorithms (Random and Degree centrality algorithms), respectively. However, there is a critical value of α around $\alpha^* = 21$, below which (i.e., $\alpha < \alpha^*$) Algorithm 2 is slightly better than Algorithm 3, and above which (i.e., $\alpha > \alpha^*$) Algorithm 3 is slightly better than Algorithm 2. However, the differences are very small as shown in Fig.9(c). We conclude that for medium α, namely $10^0 \leq \alpha \leq 10^2$, it is reasonable to use Algorithm 2 instead of Algorithm 3 because the former is computationally more efficient.

The results are summarized as follows:

- For $\alpha \leq 1$, Algorithm 2 is almost as good as Algorithm 1 when k is small (e.g. $k \leq 40$ in the example mentioned above). Since Algorithm 1 is not efficient when k is large (e.g., $k \geq 40$ in the example mentioned above) and Algorithm 2 is always fast regardless of k, Algorithm 2 should be used when $\alpha \leq 1$.

- For $\alpha \geq 100$, Algorithm 3 should be used.

- For $0 \leq \alpha \leq 100$, we observe a critical value $\alpha = 21$, below which Algorithm 2 is slightly better than Algorithm 3 and above which Algorithm 3 is slightly better than Algorithm 2. Since Algorithm 2 is fast and the difference between the results of Algorithm 2 and Algorithm 3 is small, Algorithm 2 should be used.

III. PHYSICAL IMPLICATIONS

In this section we discuss the physical implications of the results mentioned above, while noting that maximizing the spectral radius can enhance the network robustness against failures, blackouts, jamming, and attacks [34]. In what follows we elaborate this effect in two concrete scenarios: epidemic spreading and synchronization.

A. Spreading Processes

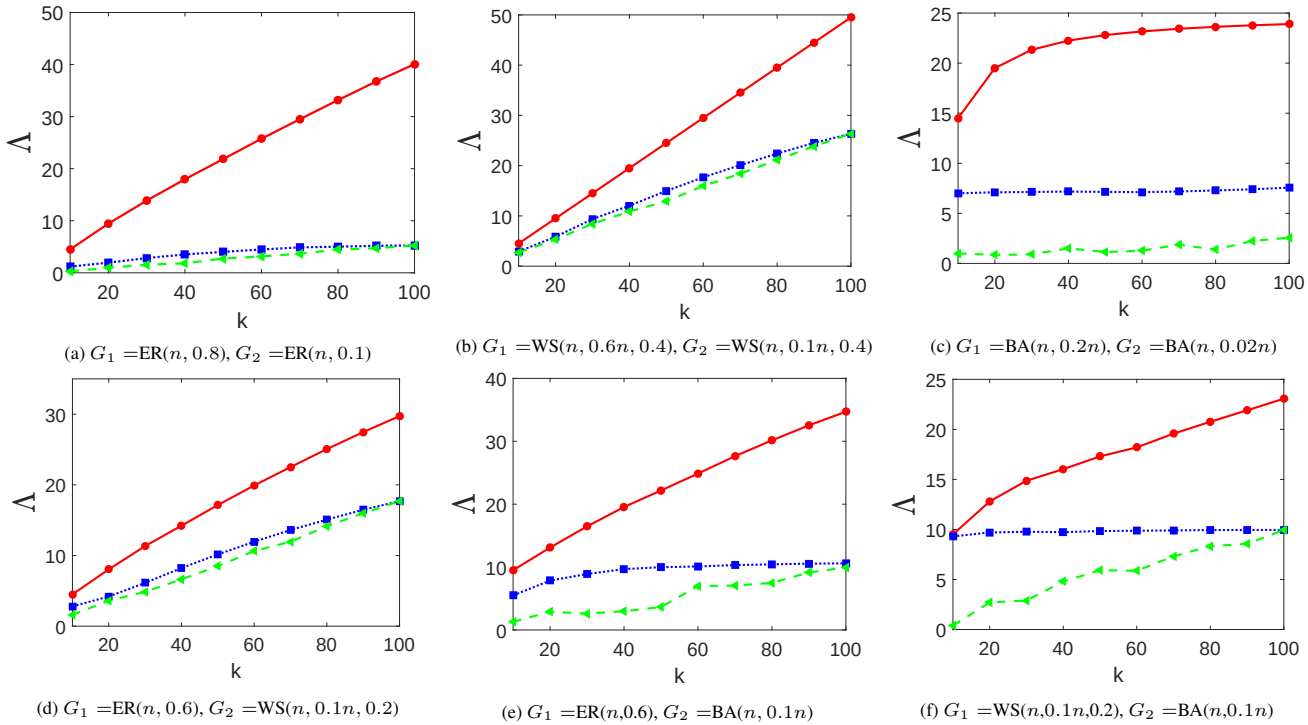
Let us first consider the Susceptible-Infected-Susceptible (SIS) epidemic model in networks [8–13], which has been used to model a family of spreading processes, ranging from the viral propagation in social and technological networks to the dissemination of information such as rumors and data [35]. In this model, each node may be in one of two states: *susceptible* or *infected*. Each node in the network represents an individual, and each link represents a connection along which the infection can propagate. A susceptible node becomes infected with probability γ_{vu} over an edge connecting to an infected node. An infected node returns to the susceptible state with probability β. The SIS model considers a finite network graph $G = (V, E, W)$, where $V = \{1, 2, \ldots, n\}$ is the set of vertices, E is the set of edges, and W is the weight matrix denoting the weight at each link and can be regarded as the weight adjacent matrix of graph G.

For concreteness, consider a discrete-time model with time $t = 0, 1, 2, \ldots$. Denote by $s_v(t)$ the probability that $v \in V$ is susceptible at time t, and $i_v(t)$ the probability that $v \in V$ is infected at time t, where $s_v(t) + i_v(t) = 1$. The master equation of the nonlinear dynamical system is [8–13]

$$i_v(t+1) = \big(1 - \prod_{(u,v)\in E} [1 - \gamma_{vu} i_u(t)]\big)(1 - i_v(t)) + (1-\beta) i_v(t). \tag{15}$$

A very recent result [13] shows that the dynamics always converge to a unique equilibrium $(i_1^*, \ldots, i_n^*)^\top$, namely that $\lim_{t\to 0} i_v(t) = i_v^*$ and a remarkable property of the SIS epidemic model is the appearance of a phase transition when $\rho(W)$ approaches the spreading threshold τ with $W = (\gamma_{vu})$. If $\rho(W) \leq \tau$, the dynamics converge to the origin equilibrium (i.e., $i_v^* = 0$ for all $v \in V$, which implies that the spreading dies out); if $\rho(W) > \tau$, the dynamics converge to a unique nonzero equilibrium (i.e., $i_v^* > 0$ for all $v \in V$) [13]. We note that [13] considered the case that the elements of W take values in $\{0, 1\}$, but the extension to the weighted case is straightforward. Hence, the spectral radius $\rho(W)$ is a measure of the network spreading power: the larger the spectral radius, the more powerful the spreading [13, 35].

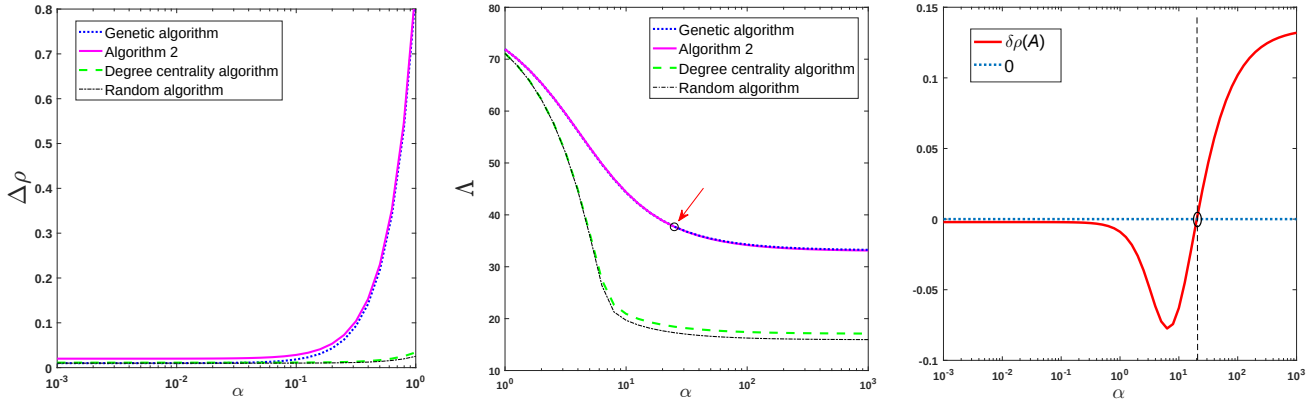
In order to quantify the effect of optimal inter-connections to maximize the spectral radius of the resulting interdependent networks, and therefore the effect on the equilibrium state, we consider an interdependent network G resulting from two networks $G_1 = (V_1, E_1, W_1)$ and $G_2 = (V_2, E_2, W_2)$ with adjacent matrix A_1 and A_2 respectively. Suppose the infection spreading rate within networks G_1 and G_2 as a uniform γ and the inter-network transmission rate as $[1 - (1 - \gamma)^r]$ for some constant $r > 0$ [36]. Let $\alpha(r) = [1 - (1 - \gamma)^r]/r$, which can be regarded as the weight of edges linking G_1 and G_2. The



(a) G_1 =ER$(n, 0.8)$, G_2 =ER$(n, 0.1)$ (b) G_1 =WS$(n, 0.6n, 0.4)$, G_2 =WS$(n, 0.1n, 0.4)$ (c) G_1 =BA$(n, 0.2n)$, G_2 =BA$(n, 0.02n)$

(d) G_1 =ER$(n, 0.6)$, G_2 =WS$(n, 0.1n, 0.2)$ (e) G_1 =ER$(n,0.6)$, G_2 =BA$(n, 0.1n)$ (f) G_1 =WS$(n,0.1n,0.2)$, G_2 =BA$(n,0.1n)$

FIG. 8. Comparison of $\Lambda = \rho(A) - \alpha\sqrt{u_0}$ with respect to k and large α: the solid red curve corresponds to Algorithm 3, the dotted blue curve corresponds to the Degree centrality algorithm for large α, and the dashed green curve corresponds to the Random algorithm for large α, where $n = 100$ and $\alpha = 10^3$.



(a) (b) (c)

FIG. 9. (a) Comparison of $\Delta\rho = \rho(A) - \max(\rho(A_1), \rho(A_2))$ derived from Algorithm 2, Algorithm 3 (i.e., the genetic algorithm), and the two alternate algorithms with respect to $10^{-3} \le \alpha \le 10^0$. (b) Comparison of $\Lambda = \rho(A) - \alpha\sqrt{u_0}$ derived from Algorithm 2, Algorithm 3 (i.e., the genetic algorithm), and the two alternate algorithms with respect to $10^0 \le \alpha \le 10^3$. The solid pink curve represents Algorithm 2, the dotted blue curve represents Algorithm 3, the dashed green curve represents Degree centrality algorithm, and the black chain curve represents Random algorithm. (c) Detailed difference between Algorithm 3 (i.e., the genetic algorithm) and Algorithm 2 in the range of $10^{-3} \le \alpha \le 10^3$, where $\delta\rho(A)$ denotes the spectral radius resulting from the Genetic algorithm subtracting the spectral radius resulting from Algorithm 2. The parameters are G_1 =ER(100,0.8), G_2 =ER(100,0.4), and $k = 80$.

SIS model in the interdependent networks becomes

$$i_v(t+1) = \Big(1 - \prod_{(u,v)\in E_j} [1 - \gamma i_u(t)] - \prod_{(u,v)\in E_*} [1 - \gamma\alpha(r) i_u(t)]\Big) * (1 - i_v(t)) + (1-\beta) i_v(t), \quad v \in G_1 \cup G_2. \tag{16}$$

Thus, the critical value of spreading dynamics should satisfy $\gamma\rho(A) = \beta$ with

$$A = \begin{pmatrix} A_1 & \alpha(r)C \\ \alpha(r)C^\top & A_2 \end{pmatrix}$$

Denote by $\tau_c = \gamma/\beta$.

For simulation, we consider G_1=WS(100,8,0.4) and G_2=BA(100,2) with linking weight $\alpha = 1$ (equivalently

$r = 1$) and k inter-connections. We simulate Eq.(16) with $(\gamma, \beta) = (0.1, 0.9)$ with $\tau_c = 1/9$ and the initial infection of 5% of nodes that are randomly picked. We calculate the quantity $\frac{1}{|G|}\sum_{v\in V_1\cup V_2} i_v^*$ at the $500th$ step to represent the equilibrium infection rate. Fig.10 (a) shows that the spectral radius $\rho(A)$ of interdependent network by connecting G_1 and G_2 according to Algorithm 2 is larger than the Random algorithm and the Degree centrality algorithm. This results in $\rho(A)$ exceeds the critical value $1/\tau_c = \beta/\gamma = 9$ at around $k = 10$, earlier than the two alternate algorithms. This means that Algorithm 2 leads to a much earlier outbreak of the information or epidemics when increasing the number of inter-connections, as shown in FIG.10 (b). This is in a good agreement with our theoretical analysis.

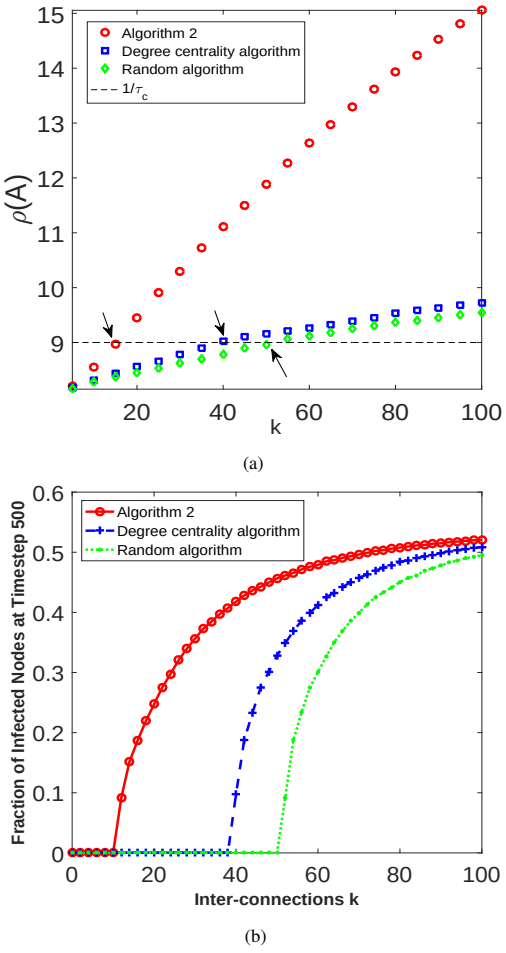


FIG. 10. (a) The spectral radius of the interdependent complex networks $\rho(A) = \rho(G)$ with k inter-connections resulting from different inter-connection algorithms. (b) Comparision of equilibria resulting from different inter-connection algorithms with respect to the number of inter-connections k. All values are averaged over 1,000 simulation runs with $(\gamma, \beta) = (0.1, 0.9)$, G_1=WS(100,8,0.4), G_2=BA(100,2), and α=1.

B. Synchronization of coupled oscillators

The problem of synchronization in complex networks, where each node is a Kuramoto oscillator [37], was first reported for WS networks [38, 39] and BA networks [40]. These studies were mainly numerical explorations of the onset of synchronization, with the main goal of characterizing the critical coupling beyond which groups of nodes beating coherently first appear. Exact analytical results to determine the transition to synchronization on general complex networks were presented in [14, 34, 41, 42].

In the extended Kuramoto model of coupled oscillators [14], the dynamics of oscillators can be approximated by an equation for the phases θ_i of the form

$$\dot{\theta}_i = \omega_i + s\sum_{j=1}^{N} z_{ij}\sin(\theta_j - \theta_i), \tag{17}$$

where ω_i is the natural frequency of oscillator i, N is the total number of oscillators, and s represents the overall coupling strength. For each i, the corresponding ω_i is independently chosen from a known oscillation frequency probability distribution $g(\omega)$. In order to incorporate the presence of a heterogeneous network, let z_{ij} denote the elements of a $n \times n$ adjacency matrix Z.

An important characteristic of the collective dynamics of the ensemble is the global complex-valued order parameter:

$$r = \frac{\sum_{i=1}^{N} r_i}{\sum_{i=1}^{N} d_i}, \tag{18}$$

where d_i is the degree of node i defined as $d_i = \sum_{j=1}^{N} z_{ij}$ and r_i is defined by

$$r_i e^{i\psi_i} = \sum_{j=1}^{N} z_{ij}\langle e^{i\theta_j}\rangle_t. \tag{19}$$

Note that r measures the extent of coherence of the system, ψ is the average phase of all of the oscillators, $r = 1$ corresponds to the complete in-phase synchronization, and $r = 0$ corresponds to the absence of an in-phase synchronization. Studies have showed that the onset of synchronization occurs at a critical coupling strength that is inversely proportional to the spectral radius of the adjacency matrix of the coupling network. The critical transition value, denoted by s_c, is

$$s_c = \frac{s_0}{\rho(Z)}, \tag{20}$$

where $s_0 \equiv 2/(\pi g(0))$ and $\rho(Z)$ is the largest eigenvalue of the adjacency matrix Z.

Consider interdependent network G that is created from two networks G_1 and G_2, whose adjacency matrices are respectively denoted by $A_1 = [a_{i,j}^{(1)}]_{i,j=1}^{m}$ and $A_2 = [a_{i,j}^{(2)}]_{i,j=1}^{n}$. The adjacency matrix of G is

$$A = \begin{bmatrix} A_1 & \alpha C \\ \alpha C^\top & A_2 \end{bmatrix}$$

Thus, for node i in network G_1, Eq.(17) becomes

$$\dot{\theta}_i = \omega_i + s\left[\sum_{j\in G_1} a_{i,j}^{(1)}\sin(\theta_j-\theta_i) + \alpha\sum_{j\in G_2} c_{ij}\sin(\theta_j-\theta_i)\right],$$

and for node i in network G_2, Eq.(17) becomes

$$\dot{\theta}_i = \omega_i + s\left[\sum_{j\in G_2} a_{i,j}^{(2)}\sin(\theta_j-\theta_i) + \alpha\sum_{j\in G_1} c_{ij}^{\top}\sin(\theta_j-\theta_i)\right],$$

The critical value of the coupling strength towards synchronization on interdependent complex networks becomes

$$s_c = \frac{s_0}{\rho(A)}. \tag{21}$$

This suggests that a larger spectral radius of the interdependent network G will reduce the critical value of the coupling strength more towards coherence and thus will advance the emergence of the synchronization phenomenon. Accelerating synchronization can be physically useful in, for example, biological processes, power grids, and transportation networks.

In order to quantify the effect, we choose a distribution of natural frequencies given by $g(\omega) = (3/4)(1-\omega^2)$ for $-1<\omega<1$ and $g(\omega) = 0$ otherwise. We consider the interdependent network G resulting from G_1=WS(100,10,0.4) and G_2=BA(100,2) with $\alpha = 1$ and $k = 80$. The spectral radius $\rho(A)$ of the interdependent network with k inter-connections resulting from Algorithm 2, the Degree centrality algorithm, and the Random algorithm are 15.6993, 11.6679, and 10.4733, respectively. This suggests that Algorithm 2 can accelerate the synchronization in interdependent networks. We set Algorithm 2 as the baseline by taking $s_c = s_0/\rho(A)$, where $\rho(A)$ is the spectral radius resulting from Algorithm 2. Fig.11 shows that when the coupling strength s exceeds the critical value s_c (i.e., $s/s_c = 1$), the onset of synchronization appears in the interdependent network with inter-connections resulting from Algorithm 2. However, the onsets of synchronization in the interdependent networks resulting from the two alternate algorithms are triggered when $s/s_c > 1$. Furthermore, one can see that r has a larger value in the interdependent networks resulting from Algorithm 2 than in the interdependent networks resulting from the two alternate algorithms. This is in a good agreement with our theoretical analysis.

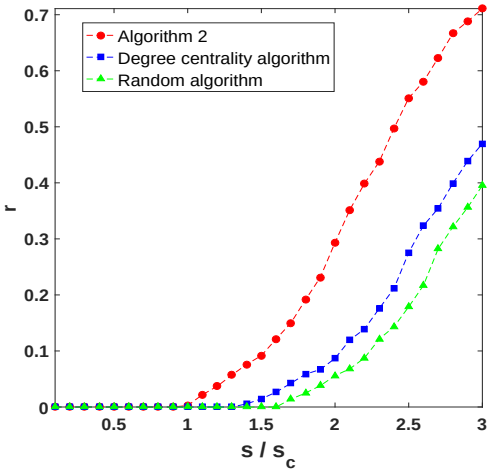


FIG. 11. The order parameter r obtained from the numerical solution of Eq.(21) and Eq.(21) as a function of s/s_c for interdependent network G, which is obtained from G_1 generated by WS(100,10,0.4) and G_2 generated by BA(100,2) with $\alpha = 1$ and $k = 80$. All values are averaged from 1,000 simulation runs.

IV. DISCUSSIONS AND CONCLUSIONS

We have studied the problem of how to select k inter-connections between two networks to maximize the spectral radius of the resulting interdependent network. We have proposed algorithms that are applicable in different scenarios. For the case of small and medium inter-connection weight α, a fast algorithm based on the numerical characteristic of the adjacent matrices of G_1 and G_2 performs well, better than the alternate methods of random inter-connections or node-centrality based inter-connections. We have found that the other notions of node-centrality, including betweenness, eigenvector, and closeness centralities, perform similarly with the node-centrality. The research has both theoretical significance and practical value, as shown in the context of the SIS model and the onset of synchronization in the coupled oscillators of Kuramoto model.

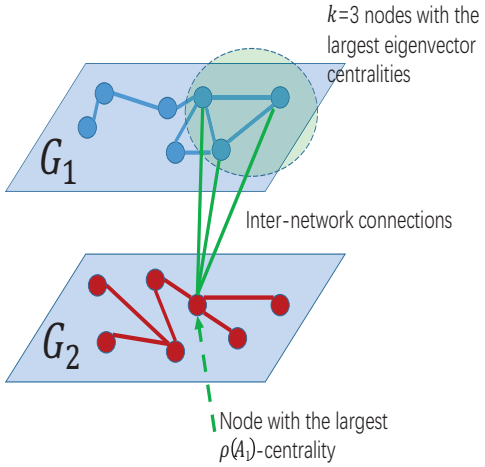


FIG. 12. Illustration of Algorithm 2.

It can be seen that Algorithm 2 and Algorithm 3 do not depend on the parameter α, but their performance is dependent upon α. Specifically, Algorithm 2 generally provided an efficient and fast method of connecting two networks to maximize the spectral radius when the inter-connection weight α is not very large. As illustrated in FIG.12, Algorithm 2 leads to a star-like subgraph of inter-connections, which is composed of the k nodes in network G_1 (which has a larger spectral radius) with the highest eigenvector centralities, and the node in G_2 that corresponds to the largest diagonal element in $(\rho(A_1)I - A_2)^{-1}$. This inspires us to introduce a new notion of node-centrality, dubbed ζ-centrality, in a graph G with adjacent matrix F for each $\zeta > \rho(F)$ as follows: let $M(\zeta) = (\zeta I - F)^{-1}$ with elements $m_{ij}(\zeta)$; then, the

ζ-centrality of node i is $m_{ii}(\zeta)$. It can be seen that Algorithm 2 selects the node in G_2 with the largest ζ-centrality with $\zeta = \rho(A_1)$. As shown in Fig.13, this centrality is different from the popular notions of centralities, for example, degree centrality, betweenness centrality, closeness centrality and eigenvector centrality.

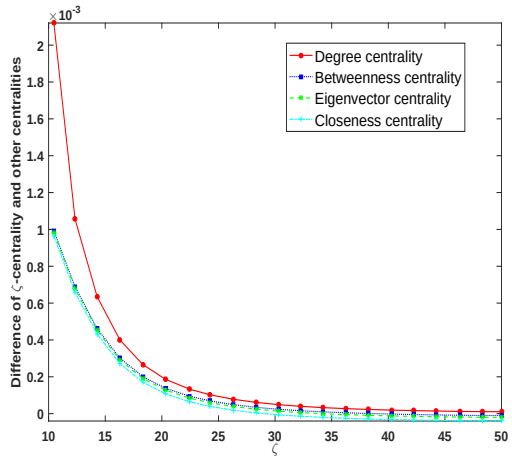


FIG. 13. Comparison between the ζ-centrality and the other four centralities, where G_2=WS(100,8,0.4), and "$\Delta\zeta$-centrality" is the largest ζ-centrality (Algorithm 2) subtracting the ζ-centrality of the node with the largest centrality of another kind (i.e., the degree, betweenness, eigenvector and closeness centrality).

There are several interesting problems for future research. For example, the case of $\rho(A_1) = \rho(A_2)$ leads to a different optimization problem, which would require another algorithm. The case of $\rho(A_1) \to \rho(A_2)$ or $\rho(A_1) \approx \rho(A_2)$ also would require a separate treatment. This is because these conditions render Algorithms 1 and 2, or more specifically the calculation of the auxiliary matrix M, numerically unreliable and inefficient, due to the fact that $\rho(A_1)I_n - A_2$ is ill-conditioned and therefore Approximation (2) cannot be used. Besides, the optimization problem appears to be NP-hard. However, a formal proof appears to be hard because known NP-hard problems are often about combinatorial properties rather than algebraic properties of matrices. Another problem is to inter-connect two networks with k inter-connections to minimize the spectral radius of the resulting interdependent network.

ACKNOWLEDGEMENTS

We thank the reviewers for their constructive comments that helped improve the paper significantly. W. L. Lu is jointly supported by the National Natural Sciences Foundation of China under Grant No. 61673119, the Key Program of the National Science Foundation of China No. 91630314, the Laboratory of Mathematics for Nonlinear Science, Fudan University and the Shanghai Key Laboratory for Contemporary Applied Mathematics, Fudan University. F. Liu is supported by the Strategic Priority Research Program of the Chinese Academy of Sciences under Grant No. XDA06010701, the National Key Research and Development Program of China with No. 2016YFB0800100. S. Xu is supported in part by ARO Grant #W911NF-13-1-0141 and NSF Grant #1111925. The views and opinions reported in the paper are those of the author(s) and do not reflect those of the funding agencies.

APPENDIX A: Derivation of Eq.(1)

Our goal is to approximate the spectral radius of interdependent networks with a small inter-connection weight α as:

$$\rho(A) = \rho(A_1) + \alpha\lambda_1 + \alpha^2\lambda_2 + o(\alpha^2), \tag{A1}$$

and the right eigenvector of A corresponding to $\rho(A)$ is

$$\xi = \begin{bmatrix} \phi + \alpha u_1 + \alpha^2 u_2 \\ \alpha v_1 + \alpha^2 v_2 \end{bmatrix} + o(\alpha^2), \tag{A2}$$

as $\alpha \to 0$. According to the property of nonnegative matrices, all of the quantities mentioned above are real, where $u_1, u_2, v_1, v_2 \in \mathbb{R}$ and $\lambda_1, \lambda_2 \in \mathbb{R}$ are to be determined below. According to $A\xi = \rho(A)\xi$, we have

$$\begin{aligned} A\xi &= \begin{bmatrix} A_1 & \alpha C \\ \alpha D & A_2 \end{bmatrix} \begin{bmatrix} \phi + \alpha u_1 + \alpha^2 u_2 \\ \alpha v_1 + \alpha^2 v_2 \end{bmatrix} + o(\alpha^2) \\ &= \begin{bmatrix} A_1\phi + \alpha A_1 u_1 + \alpha^2 A_1 u_2 + \alpha^2 C v_1 \\ \alpha D\phi + \alpha^2 D u_1 + \alpha A_2 v_1 + \alpha^2 A_2 v_2 \end{bmatrix}, \end{aligned}$$

$$\begin{aligned} &\rho(A)\xi \\ &= \begin{bmatrix} \rho(A_1)\phi + \alpha\rho(A_1)u_1 + \alpha^2\rho(A_1)u_2 + \alpha\lambda_1\phi + \alpha^2\lambda_1 u_1 + \alpha^2\lambda_2\phi \\ \alpha\rho(A_1)v_1 + \alpha^2\rho(A_1)v_2 + \alpha\lambda_1 v_1, \end{bmatrix} \end{aligned}$$

which lead to:

$$A_1 u_1 = \rho(A_1)u_1 + \lambda_1\phi, \tag{A3}$$

$$A_1 u_2 + C v_1 = \rho(A_1)u_2 + \lambda_1 u_1 + \lambda_2\phi, \tag{A4}$$

$$A_2 v_1 + D\phi = \rho(A_1)v_1, \tag{A5}$$

$$A_2 v_2 + D u_1 = \rho(A_1)v_2 + \lambda_1 v_1. \tag{A6}$$

Multiplying $\psi^\top$ to Eq.(A4), we have

$$\psi^\top A_1 u_1 = \rho(A_1)\psi^\top u_1 + \lambda_1\psi^\top\phi.$$

Since $\psi^\top\phi \neq 0$ (both are positive because A_1 is irreducible, according to the Perron-Frobenius theorem [26]), we have

$$\lambda_1 = 0.$$

From Eq.(A6), we can obtain

$$v_1 = [\rho(A_1)I_n - A_2]^{-1} D\phi.$$

Multiplying Eq.(A5) by $\psi^\top$, we have

$$\lambda_2 = \frac{\psi^\top C[\rho(A_1)I_n - A_2]^{-1} D\phi}{\psi^\top\phi}.$$

Hence, we have the following approximation:

$$\begin{aligned}\rho(A) &= \rho(A_1) + \alpha^2\lambda_2 + o(\alpha^2)\\ &= \rho(A_1) + \alpha^2 \frac{\psi^\top C[\rho(A_1)I_n - A_2]^{-1} D\phi}{\psi^\top \phi} + o(\alpha^2).\end{aligned}$$

For undirected graphs, we have $C = D^\top$. This leads to:

$$\rho(A) = \rho(A_1) + \alpha^2 \frac{\psi^\top C[\rho(A_1)I_n - A_2]^{-1} C^\top \phi}{\psi^\top \phi} + o(\alpha^2). \tag{A7}$$

Since A is symmetric, we have ϕ=ψ. By the normalization mentioned above, we have $\psi^\top\psi = 1$. Due to $x^\top = \psi^\top C$ and $M = [\rho(A_1)I_n - A_2]^{-1} = [m_{ij}]_{n\times n}$, Eq.(A7) becomes (1).

APPENDIX B: Derivation of Eq.(2)

Denote by M^* the adjoint matrix of M. Since $M^{-1} = \frac{1}{det(M)}M^*$, we have

$$\begin{aligned}M &= (\rho(A_1)I - A_2)^{-1}\\ &= \frac{1}{det(\rho(A_1)I - A_2)}(\rho(A_1)I - A_2)^*\\ &\sim \frac{1}{\rho(A_1)^n}(\rho(A_1)I - A_2)^*.\end{aligned} \tag{B1}$$

From Eq.(B1), we can obtain m_{ll} and $m_{ij}, i \neq j$. In what follows we give the derivations of the main diagonal element m_{11} and the non-diagonal element m_{1n}, while noting that the other elements elements can be derived in a similar fashion.

$$\begin{aligned}m_{1n} &\sim \frac{1}{\rho(A_1)^n}(\rho(A_1)I - A_2)^*_{1n}\\ &= \frac{-1}{\rho(A_1)^n}\begin{vmatrix} -a^{(2)}_{21} & \rho(A_1)-a^{(2)}_{22} & -a^{(2)}_{13} & \dots & -a^{(2)}_{2,n-1}\\ -a^{(2)}_{31} & -a^{(2)}_{32} & \rho(A_1)-a^{(2)}_{33} & \dots & -a^{(2)}_{3,n-1}\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ -a^{(2)}_{n1} & -a^{(2)}_{n2} & -a^{(2)}_{n3} & \dots & -a^{(2)}_{n,n-1}\end{vmatrix}\\ &= \frac{-1}{\rho(A_1)^n}\cdot((-a^{(2)}_{n1})\prod_{l=2}^{n-1}(\rho(A_1) - a^{(2)}_{ll}) + O(\rho(A_1)^{n-3}))\\ &= \frac{1}{\rho(A_1)^n}\cdot a^{(2)}_{n1}\cdot\left(\rho(A_1)^{n-2} - \left(\sum_{l=2}^{n-1} a^{(2)}_{ll} + f_{1n}(A_2)\right)\right.\\ &\left.\cdot\,\rho(A_1)^{n-3} + O(\rho(A_1)^{n-4})\right).\end{aligned} \tag{B2}$$

Let $T_1 = \sum_{l=2}^{n-1} a^{(2)}_{ll} + f_{1n}(A_2)$, where $f_{1n}(A_2)$ is the function of the cofactors of $A_2(1,n)$. Since $\forall i,j, a^{(2)}_{ij}$ and $\sum_j a^{(2)}_{ij}$ are both finite, hence T_1 is finite. When $\mu \gg 1$, namely $\rho(A_1)\to\infty$, we have

$$\begin{aligned}m_{1n} &\sim \frac{1}{\rho(A_1)^n}\cdot a^{(2)}_{n1}\cdot\rho(A_1)^{n-3}\left(\rho(A_1) - \sum_{l=2}^{n-1} a^{(2)}_{ll} + f_{1n}(A_2)\right)\\ &\sim \frac{1}{\rho(A_1)^n}\cdot a^{(2)}_{n1}\cdot\rho(A_1)^{n-2}\\ &= \frac{1}{\rho(A_1)^2}\cdot a^{(2)}_{n1}.\end{aligned} \tag{B3}$$

On the other hand, we have

$$\begin{aligned}m_{11} &\sim \frac{1}{\rho(A_1)^n}(-1)^{1+1}\begin{vmatrix}\rho(A_1)-a^{(2)}_{22} & -a^{(2)}_{23} & -a^{(2)}_{24} & \dots & -a^{(2)}_{2,n}\\ -a^{(2)}_{32} & \rho(A_1)-a^{(2)}_{33} & -a^{(2)}_{34} & \dots & -a^{(2)}_{3,n}\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ -a^{(2)}_{n1} & -a^{(2)}_{n2} & -a^{(2)}_{n3} & \dots & \rho(A_1)-a^{(2)}_{n,n}\end{vmatrix}\\ &= \frac{1}{\rho(A_1)^n}\cdot\left(\rho(A_1)^{n-1} - \left(\sum_{l=2}^{n} a^{(2)}_{ll}\right)\rho(A_1)^{n-2} + O(\rho(A_1)^{n-3})\right)\end{aligned} \tag{B4}$$

$$\sim \frac{1}{\rho(A_1)^n}\cdot\rho(A_1)^{n-1} = \frac{1}{\rho(A_1)}. \tag{B5}$$

Similarly, we can ignore the part $T_2 = \left(\sum_{l=2}^{n} a^{(2)}_{ll}\right)$ when deriving Eq.(B5) from Eq.(B4). Hence, we derive the following approximation:

$$\Gamma = \frac{\min\limits_{l=1,\dots,n}(m_{ll})}{\max\limits_{\substack{i\neq j\\ i,j=1,\dots,n}}(m_{ij})} \sim \frac{\rho(A_1)}{\max\limits_{\substack{i\neq j\\ i,j=1,\dots,n}} a^{(2)}_{ij}} \sim \frac{\mu\cdot\rho(A_2)}{\max\limits_{\substack{i\neq j\\ i,j=1,\dots,n}} a^{(2)}_{ij}} \sim b_*\mu. \tag{B6}$$

APPENDIX C: Validation of Rule 1

Let

$$C^{3,1} = \begin{pmatrix}1&1&1\\0&0&0\\0&0&0\end{pmatrix}, C^{3,2} = \begin{pmatrix}1&1&0\\1&0&0\\0&0&0\end{pmatrix}, C^{3,3} = \begin{pmatrix}1&0&0\\1&0&0\\1&0&0\end{pmatrix}.$$

Denote by $\lambda_{3,*}$ the value of $\Sigma_1 = \sum_{l=1}^n x_l^2 m_{ll}$ after substituting $C^{3,*}$ into it, we have

$$\lambda_{3,1} = \psi_1^2 m_{11} + \psi_1^2 m_{22} + \psi_1^2 m_{33} \tag{C1}$$

$$\lambda_{3,2} = (\psi_1+\psi_2)^2 m_{11} + \psi_1^2 m_{22} \tag{C2}$$

$$\lambda_{3,3} = (\psi_1+\psi_2+\psi_3)^2 m_{11} \tag{C3}$$

We cannot compare $\lambda_{3,1}, \lambda_{3,2}, \lambda_{3,3}$ based on their algebraic expressions in general, but can compare them in specific scenarios. Suppose the matrix C that leads to the maximum Σ_1 is not any of $C^{3,1}, C^{3,2}, C^{3,3}$. For example, suppose C has the form

$$C^{3,4} = \begin{pmatrix}1&0&1\\0&1&0\\0&0&0\end{pmatrix}.$$

Thus, we have $x^\top_{3,4} = [\psi_1, \psi_2, \psi_1] \leq x^\top_{3,1} = [\psi_1, \psi_1, \psi_1]$. Obviously,

$$\lambda_{3,4} = x^\top_{3,4} M x_{3,4} < \lambda_{3,1} = x^\top_{3,1} M x_{3,1}.$$

This means that when C takes the form of $C^{3,4}$, $\lambda_{3,4}$ must be smaller than the $\lambda_{3,1}$ resulting from $\mathcal{C}^{3,1}$, which contradicts the aforementioned assumption. Similarly, given an arbitrary adjacency matrix $C^{3,s} \in \mathbb{R}^{3\times 3}$, if $s > 3$, one can always find a $C^{3,i}, i = 1,2,3$ that makes $\lambda_{3,s} < \lambda_{3,i}$.

APPENDIX D: Derivation of Eq.(9)

Denote by λ and $\begin{bmatrix} a \\ b \end{bmatrix}$ the eigenvalue and eigenvector of matrix $Z = \begin{bmatrix} 0 & C \\ D & 0 \end{bmatrix}$ with $D = C^\top$ in our case, respectively.

$$\begin{bmatrix} 0 & C \\ D & 0 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \lambda \begin{bmatrix} a \\ b. \end{bmatrix}$$

This implies

$$CDa = \lambda^2 a, DCb = \lambda^2 b.$$

So, the largest eigenvalue of Z, denoted by μ_0, is the square root of CD (equivalently DC), and the right eigenvector a with respect to DC and the right eigenvector b with respect to CD compose the eigenvector of Z.

Suppose the algebraic dimension of the largest eigenvalues of CD and DC is one. The right-eigenvector of A corresponding to the largest eigenvalue is a perturbation from $[a^\top, b^\top]^\top$. Let

$$\rho(\tilde{A}) = \sqrt{\mu_0} + \epsilon\mu_1 + o(\epsilon),$$

and its right eigenvector be

$$\pi(\epsilon) = \begin{bmatrix} a + \epsilon r_1 \\ b + \epsilon s_1 \end{bmatrix} + o(\epsilon),$$

where $r_1, s_1 \in \mathbb{R}^m$ and $\mu_1 \in \mathbb{R}$ are to be determined later. Similar to the arguments above, we have

$$A_1 a + C s_1 = \sqrt{\mu_0} r_1 + \mu_1 a. \quad \text{(D1)}$$
$$A_2 b + D r_1 = \sqrt{\mu_0} s_1 + \mu_1 b. \quad \text{(D2)}$$

Let $\left[c^\top, d^\top\right]$ be the left eigenvector of Z corresponding to the largest eigenvalue $\sqrt{\mu_0}$, meaning

$$c^\top C = \sqrt{\mu_0} d^\top, d^\top D = \sqrt{\mu_0} c^\top.$$

Multiplying (D1) by $c^\top$ and (D2) by $d^\top$, we have

$$c^\top A_1 a + \sqrt{\mu_0} d^\top s_1 = \sqrt{\mu_0} c^\top r_1 + \mu_1 c^\top a,$$
$$d^\top A_2 b + \sqrt{\mu_0} c^\top r_1 = \sqrt{\mu_0} d^\top s_1 + \mu_1 d^\top b.$$

Summing them up gives

$$\mu_1 = \frac{c^\top A_1 a + d^\top A_2 b}{c^\top a + d^\top b}. \quad \text{(D3)}$$

So, the approximation can be given by

$$\rho(A) = \alpha\sqrt{\mu_0} + \frac{c^\top A_1 a + d^\top A_2 b}{c^\top a + d^\top b} + o(1/\alpha). \quad \text{(D4)}$$

[1] A.-L. Barabasi and R. Albert, Science **286**, 509 (1999).
[2] D. J. Watts and S. H. Strogatz, Nature **393**, 440 (1998).
[3] J. G. Restrepo, E. Ott, and B. R. Hunt, Physical review letters **97**, 094102 (2006).
[4] J. Aguirre, D. Papo, and J. M. Buldú, Nature Physics **9**, 230 (2013).
[5] M. Barthélemy, A. Barrat, R. Pastor-Satorras, and A. Vespignani, Phys. Rev. Lett. **92**, 178701 (2004).
[6] M. Boguñá, R. Pastor-Satorras, and A. Vespignani, Phys. Rev. Lett. **90**, 028701 (2003).
[7] R. M. May and A. L. Lloyd, Phys. Rev. E **64**, 066112 (2001).
[8] Y. Wang, D. Chakrabarti, C. Wang, and C. Faloutsos, in *Proc. of the 22nd IEEE Symposium on Reliable Distributed Systems (SRDS'03)* (2003), pp. 25–34.
[9] A. Ganesh, L. Massoulie, and D. Towsley, in *Proceedings of IEEE Infocom 2005* (2005), pp. 1455–1466.
[10] P. V. Mieghem, J. Omic, and R. Kooij, IEEE/ACM Trans. Netw. **17**, 1 (2009).
[11] S. Xu, W. Lu, and L. Xu, ACM Transactions on Autonomous and Adaptive Systems (ACM TAAS) **7**, 32 (2012).
[12] S. Xu, W. Lu, and Z. Zhan, IEEE Trans. Dependable Sec. Comput. **9**, 30 (2012).
[13] R. Zheng, W. Lu, and S. Xu, CoRR **abs/1602.06807** (2016).
[14] J. G. Restrepo, E. Ott, and B. R. Hunt, Phys. Rev. E **71**, 036151 (2005).
[15] J. G. Restrepo, E. Ott, and B. R. Hunt, Phys. Rev. Lett. **100**, 058701 (2008).
[16] K. C. Das and P. Kumar, Discrete Mathematics **281**, 149 (2004).
[17] V. Nikiforov, Linear Algebra and its Applications **418**, 257 (2006).
[18] Y. Hong, J.-L. Shu, and K. Fang, Journal of Combinatorial Theory, Series B **81**, 177 (2001).
[19] A. Milanese, J. Sun, and T. Nishikawa, Phys. Rev. E **81**, 046112 (2010).
[20] S. V. Buldyrev, R. Parshani, G. Paul, and et al., Nature **464**, 1025 (2010).
[21] S. W. Son, P. Grassberger, and M. Paczuski, Phys. Rev. Lett. **107**, 195702 (2011).
[22] G. J. Baxter, S. N. Dorogovtsev, A. V. Goltsev, and et al., Phys. Rev. Lett. **109**, 248701 (2012).
[23] J. Martin-Hernandez, H. Wang, P. Van Mieghem, and G. DAgostino, Physica A: Statistical Mechanics and its Applications **404**, 92 (2014).
[24] S. Tauch, W. Liu, and R. Pears, in *2015 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)* (IEEE, 2015), pp. 683–688.
[25] J. Martin-Hernandez, H. Wang, P. Van Mieghem, and G. D'Agostino, arXiv preprint arXiv:1304.4731 (2013).
[26] A. Berman and N. Shaked-Monderer, *Completely positive matrices* (World Scientific, 2003).
[27] G. H. Hardy and S. Ramanujan, Proceedings of the London Mathematical Society **2**, 75 (1918).
[28] J. V. Uspensky and M. A. Heaslet, *Elementary number theory* (McGraw-Hill book Company, Incorporated, 1939).
[29] D. M. Burton, *Elementary number theory* (Tata McGraw-Hill

Education, 2006).
[30] S. H. Strogatz, Nature **410**, 268 (2001).
[31] P. Erdös and A. Rényi, Publicationes Mathematicae (Debrecen) **6**, 290 (1959).
[32] D. J. Watts and S. H. Strogatz, nature **393**, 440 (1998).
[33] A.-L. Barabási and R. Albert, science **286**, 509 (1999).
[34] D. Taylor and J. G. Restrepo, Physical Review E **83**, 066112 (2011).
[35] D. Taylor and D. B. Larremore, Physical Review E **86**, 031140 (2012).
[36] W. Wang, M. Tang, H.-F. Zhang, H. Gao, Y. Do, and Z.-H. Liu, Physical Review E **90**, 042803 (2014).
[37] S. H. Strogatz, Physica D: Nonlinear Phenomena **143**, 1 (2000).
[38] D. J. Watts, *Small worlds: the dynamics of networks between order and randomness* (Princeton university press, 1999).
[39] H. Hong, M.-Y. Choi, and B. J. Kim, Physical Review E **65**, 026139 (2002).
[40] Y. Moreno and A. F. Pacheco, EPL (Europhysics Letters) **68**, 603 (2004).
[41] J. G. Restrepo, E. Ott, and B. R. Hunt, Physical review letters **96**, 254103 (2006).
[42] P. Van Mieghem, EPL (Europhysics Letters) **97**, 48004 (2012).