Synchronization-based scalability of complex clustered networks

Xiaojuan Ma,[1,2] Liang Huang,[1] Ying-Cheng Lai,[1,3] Yan Wang,[1,2] and Zhigang Zheng[2]
[1]*Department of Electrical Engineering, Arizona State University, Tempe, Arizona 85287, USA*
[2]*Department of Physics, Beijing Normal University, Beijing 100875, China*
[3]*Department of Physics, Arizona State University, Tempe, Arizona 85287, USA*

(Received 1 July 2008; accepted 24 September 2008; published online 6 November 2008)

Complex clustered networks arise in biological, social, physical, and technological systems, and the synchronous dynamics on such networks have attracted recent interests. Here we investigate system-size dependence of the synchronizability of these networks. Theoretical analysis and numerical computations reveal that, for a typical clustered network, as its size is increased, the synchronizability can be maintained or even enhanced but at the expense of deterioration of the clustered characteristics in the topology that distinguish this type of networks from other types of complex networks. An implication is that, for a large network in a realistic situation, if synchronization is important for its function, then most likely it will not have a clustered topology. © *2008 American Institute of Physics*. [DOI: 10.1063/1.3005782]

If a complex system of relatively small size possesses a certain dynamical function, can a similar system but of much larger size function the same way? This is the issue of scalability, which is particularly relevant to complex networks. For example, if a small network of certain topology is synchronizable, would a much larger network of the same topology still be synchronizable? Despite tremendous recent progress on complex networks, the issue of scalability has received little attention. It is, however, important to biology as networks of significantly different sizes can arise in a diverse array of contexts, and the occurrence of universal dynamics across the spectrum of networks is of fundamental interest. Scalability is also of paramount importance in technological fields, such as computer networks. Here we present a systematic study of the dynamics-based scalability in complex clustered networks, a general type of topology that has been found in various biological, social, physical, and technological networks. To be concrete and to be able to obtain analytic insights, we focus on the dynamics of synchronization. Our general finding is that synchronizability and the clustered topology cannot be maintained at the same time. In particular, as the network size is increased, its synchronizability can be maintained but at the expense of continuous deterioration of the clustered characteristics that are unique to this type of network. A practical implication is that, if synchronization-based scalability is required for network design, the clustered topology is undesirable.

I. INTRODUCTION

Recently, synchronization in complex networks has received considerable attention.[1–10] There are two main motivations: (1) Synchronization is fundamental to many phenomena in nature, especially in biology,[11] and (2) many natural and technological systems exhibit traits of complex networks.[12–16] Most existing studies have focused on the synchronizability, addressing the role played by different network topologies.[5–8] Various coupling schemes have been proposed to enhance the network synchronizability. The issue of scalability, i.e., the dependence of dynamical properties of the network on its size, has also begun to be considered.[17,18] The focus of this paper is on the scalability of synchronization of complex clustered networks, networks whose characteristics have been found in various biological, social, and technological systems.[19–24]

A clustered network consists of a number of subnetworks (clusters), where nodes within each cluster are densely connected but the linkage among the clusters is sparse. A clustered network can be complex in the sense that, not only the intercluster linkage can be random, but the connections within each individual cluster can also be random,[25] small-world,[12] or scale-free.[13] Recently synchronization in complex clustered networks has been studied,[9] but the issue of size-dependence has not been systematically explored. The main question addressed in this paper is then: If a clustered network of small size is synchronizable, under what conditions will networks of the same topology but of much larger size still be synchronizable? The answer to this question can reveal the interplay between synchronization and the clustered topology and help provide insights into whether *large* complex clustered networks can be pervasive in natural systems with respect to synchronization.

We will use the standard approach of network spectral analysis, namely the master-stability-function (MSF) approach[26] to explore the size-dependence issue. In particular, previous works have established that the synchronizability of a network can be characterized by the spread of the eigenvalue spectrum of the underlying coupling matrix.[2–4] Given a clustered network, we shall obtain analytic estimates for both the smallest and the largest nontrivial eigenvalues as a function of the network size, based on which the range of the coupling parameter, say ε, for which synchronization is possible can be obtained. A network is regarded as scalable with respect to synchronization if there exists a finite range of ε in which synchronization can occur, insofar as the network size is finite. Likewise, a network is not scalable if the synchronizable parameter range becomes zero as the network size exceeds a critical value. To state our main result, it is necessary to define parameters to characterize a complex clustered network. In this regard the probabilities of intercluster and intracluster links, denoted by p_l and p_s, respectively, are most relevant. For the clustered topology to be distinct, it is required by definition that $p_l \ll p_s$. Our analysis indicates that, for fixed values of p_l and p_s, the network is scalable with respect to synchronization. However, as we will see, when p_l and p_s are fixed, the densities of the intercluster and intracluster linkages increase in different manners as the network size is increased. If the size is sufficiently large, the intercluster link density can surpass the intracluster link density (unless the number of clusters is small). When this occurs, the characteristics of the clustered topology are completely lost, reducing the network to one with the standard complex topology determined by the specific topology of the individual subnetwork. On the other hand, if the intercluster link density is fixed so that the clustered topology is maintained, the network's synchronizability is lost when its size becomes sufficiently large. The general phenomenon is then that complex clustered networks are not scalable with respect to synchronization. An implication is that, if synchronization is important to the functions of a large networked system, the complex clustered topology is not desirable. For large networks in biology, if synchronization is fundamental, they are most likely to be nonclustered. Our result also provides a dynamics-based explanation to the difficulty to achieve synchronization in many social networks that are typically large and clustered.

In Sec. II, we argue for the general applicability of the MSF formalism in scalability analysis. In Sec. III, we provide a spectral analysis for complex clustered networks under two coupling schemes. In Sec. IV, we apply the results in Sec. III to obtain analytic results concerning the scalability of such networks and provide numerical support. Conclusions are offered in Sec. V.

II. MASTER-STABILITY FUNCTION AND ITS GENERALITY

The aim of our study is to address network scalability by focusing on the network's ability to synchronize, not on actual synchronization. This approach would allow general conclusions to be drawn, in spite of the complexity of the problem. If actual synchronization were to be considered, general insights would be difficult to obtain as the synchronization would depend on many specific details, such as initial conditions. Thus, in this paper, when we say that certain networks are scalable with respect to synchronization, we mean only that the networks can be synchronized, regardless of its size, if the coupling parameter and initial conditions are chosen properly. In contrast, if a class of networks is not scalable, they absolutely cannot be synchronized if their sizes exceed a critical value, regardless of how the coupling parameter or initial conditions are adjusted. It is in this sense of scalability which makes the MSF formalism[26] a powerful theoretical tool. In what follows we shall briefly describe the MSF framework and argue for its applicability when different types of node dynamics are taken into account.

We consider the following network of N coupled oscillators:

$$\frac{d\mathbf{x}_i}{dt} = \mathbf{F}(\mathbf{x}_i) - \varepsilon \sum_{j=1}^{N} G_{ij}\mathbf{H}(\mathbf{x}_j), \tag{1}$$

where $i=1,\ldots,N$, $d\mathbf{x}/dt=\mathbf{F}(\mathbf{x})$ describes the dynamics of each individual oscillator, $\mathbf{H}(\mathbf{x})$ is the coupling function to each oscillator, $\mathbf{G}=(G_{ij})$ is the coupling matrix determined by the network topology, and ε is a coupling parameter. The matrix $\mathbf{G}$ satisfies the condition $\sum_{j=1}^{N} G_{ij}=0$ for any i, ensuring the existence of a synchronized state $\mathbf{x}_i(t)=\mathbf{s}(t)$, $\forall i$, where $d\mathbf{s}/dt=\mathbf{F}(\mathbf{s})$ is a solution to Eq. (1).

Linearizing Eq. (1) about the synchronized state yields

$$\frac{d\delta\mathbf{x}_i}{dt} = \mathbf{DF}(\mathbf{s}) \cdot \delta\mathbf{x}_i - \varepsilon \sum_{j=1}^{N} G_{ij}\mathbf{DH}(\mathbf{s}) \cdot \delta\mathbf{x}_j, \tag{2}$$

where $\mathbf{DF}(\mathbf{s})$ is the Jacobian matrix evaluated with respect to the synchronous state $\mathbf{s}(t)$. For linear coupling function $\mathbf{H}(\mathbf{x})$, $\mathbf{DH}(\mathbf{s})$ is a constant matrix independent of $\mathbf{s}(t)$. Let $0=\lambda_1 < \lambda_2 \leq \cdots \leq \lambda_N$ be the eigenvalue spectrum of the coupling matrix $\mathbf{G}$. Diagonalizing $\mathbf{G}$ in Eq. (2) leads to the following generalized variational equation:

$$\frac{d\delta\mathbf{y}}{dt} = [\mathbf{DF}(\mathbf{s}) - K\mathbf{DH}(\mathbf{s})] \cdot \delta\mathbf{y}, \tag{3}$$

where $\delta\mathbf{y}$ is an infinitesimal variation from the synchronous solution $\mathbf{s}(t)$, and $K=\varepsilon\lambda_i (i=2,\ldots,N)$ is a generalized coupling parameter whose spread for given value of ε is determined by the eigenvalue spectrum of the coupling matrix. The largest Lyapunov exponent $\Psi(K)$ of this variational system is the MSF (Ref. 26) of the coupled system (1), which determines whether the synchronized state is physically realizable. The state is stable if $\Psi(\varepsilon\lambda_i) < 0$ for all $i=2,\ldots,N$ and unstable otherwise. Note that the MSF being negative is a necessary but not sufficient condition for synchronization to actually occur. In particular, for random initial conditions synchronization is likely if the MSF is negative. However, if the MSF is positive, synchronization will not occur, regardless of the choice of the initial condition.

For different types of node dynamics, the MSF shows some different behaviors. What has often been assumed in the network-synchronization literature[2–4] is that the MSF is negative in a single, finite interval. However, to encompass all possible situations, we shall also address the cases where the interval tends to infinity and where the MSF has several distinct stable regions.

Class-I node dynamics: To be concrete, we assume chaotic dynamics on any single node so that $\Psi(0) > 0$. For synchronization to be possible, $\Psi(K)$ must be negative in some region of K. There is thus a cross point of $\Psi(K)$ with the K-axis at which $\Psi(K)$ becomes negative, say K_1. As K is increased, $\Psi(K)$ becomes positive again at K_2 and remains positive thereafter. In this case, $\Psi(K)$ is negative in a finite interval (K_1,K_2), and the stability condition for synchronization becomes

$$\lambda_2 > \frac{K_1}{\varepsilon} \quad \text{and} \quad \lambda_N < \frac{K_2}{\varepsilon},$$

or

$$\frac{K_1}{\lambda_2} < \varepsilon < \frac{K_2}{\lambda_N}.$$

Let $\varepsilon_1 = K_1/\lambda_2$ and $\varepsilon_2 = K_2/\lambda_N$. For a given network ($\lambda_2$ and λ_N fixed) and given node dynamics (K_1 and K_2 fixed), the interval of coupling strength that permits a stable synchronization of the system is

$$\Delta\varepsilon = \varepsilon_2 - \varepsilon_1 = \frac{K_2}{\lambda_N} - \frac{K_1}{\lambda_2}. \tag{4}$$

If $\Delta\varepsilon > 0$, the system can be made synchronizable for proper choices of the coupling parameter ε and of the initial conditions. For given K_1 and K_2, this leads to the condition for synchronization,

$$Q \equiv \frac{\lambda_N}{\lambda_2} < \frac{K_2}{K_1}, \tag{5}$$

where Q is the eigenratio. If $\Delta\varepsilon < 0$ or $Q > K_2/K_1$, synchronization will not occur no matter how the coupling parameter ε may be adjusted.[27] The eigenratio Q can thus be used as an indicator of the synchronizability of the network;[2] the smaller the value of Q, the higher the probability that the system can synchronize. For a network with given node dynamics, K_2/K_1 is constant. For the network to be scalable with respect to synchronization, $\Delta\varepsilon$ should be positive for any size N of the network. Or equivalently, the eigenratio Q should not exceed K_2/K_1 as the size of the network is increased.

Class-II node dynamics: In this case, $K_2 \to \infty$, i.e., the MSF $\Psi(K)$ is negative for $K \in (K_1, \infty)$. With respect to scalability, this is a special case of *class-I node dynamics*, in the following sense: a synchronizable or scalable system for class-I dynamics is also synchronizable or scalable for class-II dynamics.

Class-III node dynamics: In this case, $\Psi(K)$ is negative in several distinct regions, say (K_{a1},K_{b1}), $(K_{a2},K_{b2}),\ldots,(K_{af},K_{bf})$, where $K_{a1} < K_{b1} < K_{a2} < K_{b2} < \cdots < K_{af} < K_{bf}$ and K_{bf} can be either finite or infinite. When K_{bf} is finite, if each interval (K_{ai},K_{bi}) is regarded as the synchronizable interval (K_1,K_2) for class-I node dynamics, results in the scalability for class-I dynamics can be applied. Note that it is possible that K_i may reside in different stable intervals where the system is still synchronizable. For $K_{bf} \to \infty$, if a particular finite interval is of interest, results from class-I node dynamics are pertinent; otherwise results for class-II dynamics are applicable.

The above discussion suggests that, in order to address the issue of scalability, focusing on class-I node dynamics suffices.

III. SPECTRAL ANALYSIS OF COMPLEX CLUSTERED NETWORKS

We consider the following general clustered network model:[16,28] there are N nodes in a network which are divided into M clusters, and each cluster contains $n = N/M$ nodes. Nodes in the same cluster are connected with probability p_s, and the probability for two nodes, each belonging to a different cluster, to be linked is p_l. The clustered topology requires $p_l \ll p_s$. For typical clustered networks arising in different situations, the topology of the subnetworks in individual clusters is mostly random,[19–24] which we shall assume for our analysis in this paper.

For a given node dynamics, the values of the general coupling parameter that define the stable synchronization regime, K_1 and K_2, are fixed. The synchronizability of the oscillator network is then determined by its topology as characterized by the smallest and the largest nontrivial eigenvalues of the coupling matrix, λ_2 and λ_N, respectively. In the following, we shall consider two different coupling schemes and derive analytic formulas for λ_2 and λ_N.

A. Type-I coupling

In this case, the coupling matrix is defined as: for any $i(1 \le i \le N)$, $G_{ii} = k_i$, where k_i is the degree (the number of links) of node i, $G_{ij} = -1$ $(i \ne j)$ if there is a link between node i and j, and $G_{ij} = 0$ otherwise. This matrix is in fact the generalized Laplacian matrix.

To obtain an analytic estimate for λ_N, we make use of the relation between λ_N and the maximum degree of the network as derived in Ref. 17,

$$\lambda_N \approx k_{\max} + 1. \tag{6}$$

Our goal is thus to obtain an expression for $k_{\max}$ for random clustered networks.

In a single random network with connection probability p, the degree k_i of a node i follows a binomial distribution $B(N-1,p)$: $P(k_i = k) = C_{N-1}^k p^k (1-p)^{N-1-k}$, where $C_{N-1}^k = (N-1)!/[k!(N-1-k)!]$ is the binomial coefficient. When N is large, a straightforward application of the law of large numbers yields the following standard approximation:

$$P(k) \approx \frac{1}{\sqrt{Np(1-p)}}\phi\left[\frac{k-Np}{\sqrt{Np(1-p)}}\right],$$

where $\phi(x) = (1/\sqrt{2\pi})e^{-(1/2)x^2}$. For a clustered network, node i connects to the remaining $n-1$ nodes in the same cluster with probability p_s, and connects to the $N-n$ nodes in different clusters with probability p_l. Therefore, the degree distribution of the network consists of two parts: $B(n-1,p_s)$ for intracluster links and $B(N-n,p_l)$ for intercluster links. Using the approximation of normal distribution, we have

$$P_s(k) \approx \frac{1}{\sqrt{np_s(1-p_s)}}\phi\left[\frac{k-np_s}{\sqrt{np_s(1-p_s)}}\right],$$

$$P_l(k) \approx \frac{1}{\sqrt{(N-n)p_l(1-p_l)}}\phi\left[\frac{k-(N-n)p_l}{\sqrt{(N-n)p_l(1-p_l)}}\right].$$

Assuming that intracluster and intercluster links are independent of each other, we can sum the two distributions to obtain a new normal distribution for the degree distribution,

$$P(k) \approx \frac{1}{\sigma}\phi\left(\frac{k-\langle k\rangle}{\sigma}\right), \tag{7}$$

where the mean and the variance are given by

$$\langle k\rangle = np_s + (N-n)p_l,$$
$$\sigma^2 = np_s(1-p_s) + (N-n)p_l(1-p_l). \tag{8}$$

The maximum degree $k_{\max}$ of the network can be calculated by following the condition that the probability of a node to have a degree larger than or equal to $k_{\max}$ is $1/N$, i.e.,

$$\int_{k_{\max}}^{\infty} P(k)dk = 1/N.$$

Using Eq. (7), we obtain

$$k_{\max} = \mathrm{erf}^{-1}(1-2/N)\cdot\sqrt{2}\sigma + \langle k\rangle, \tag{9}$$

where $\mathrm{erf}^{-1}(x)$ is the inverse of the error function $\mathrm{erf}(x) = 2/\sqrt{\pi}\int_0^x e^{-t^2}dt$. The largest eigenvalue λ_N can then be approximated as

$$\lambda_N \approx k_{\max} + 1 = \mathrm{erf}^{-1}(1-2/N)\cdot\sqrt{2}\sigma + \langle k\rangle + 1. \tag{10}$$

For λ_2, we have $\lambda_2 = \mathbf{e}_2^T\cdot\mathbf{G}\cdot\mathbf{e}_2 = \Sigma_{i,j=1}^N e_{2i}G_{ij}e_{2j}$, where $\mathbf{e}_2$ is the eigenvector associated with λ_2 and e_{2i} is the ith component of $\mathbf{e}_2$. A recent work[18] has revealed that for a clustered network, the components of the eigenvector $\mathbf{e}_2$ have approximately the same value within a cluster. Thus the eigenvector $\mathbf{e}_2$ can be written as $\mathbf{e}_2 \approx [\tilde{e}_1,\ldots,\tilde{e}_1,\tilde{e}_2,\ldots,\tilde{e}_2,\ldots,\tilde{e}_M,\ldots,\tilde{e}_M]^T$, and for each index I, $1\leq I\leq M$, there are $n\tilde{e}_I$'s in $\mathbf{e}_2$. We have

$$\lambda_2 \approx \sum_{i=1}^N e_{2i}\{G_{i1}\tilde{e}_1 + G_{i2}\tilde{e}_1 + \cdots + G_{in}\tilde{e}_1 + G_{i(n+1)}\tilde{e}_2 + \cdots + G_{iN}\tilde{e}_M\}. \tag{11}$$

For type-I coupling, the matrix elements are: (1) $G_{ii}=k_i$, (2) $G_{ii}=-1$ with probability p_s and $G_{ii}=0$ with probability $1-p_s$ if nodes i and j belong to the same cluster, and (3) $G_{ii}=-1$ with probability p_l and $G_{ii}=0$ with probability $1-p_l$ if nodes i and j belong to different clusters. Substituting these matrix elements into Eq. (11), we have

$$\lambda_2 \approx \sum_{i=1}^N e_{2i}\{-np_l\tilde{e}_1 - np_l\tilde{e}_1 + \cdots + k_i\tilde{e}_I - np_s\tilde{e}_I + \cdots - np_l\tilde{e}_M\},$$

where $\tilde{e}_I$ is the eigenvector component associated with the cluster that contains nodes i. For a random subnetwork, the degree distribution has a narrow peak centered at $k=np_s+(N-n)p_l$, which leads to $k_i\approx k$. We can thus write λ_2 as

$$\begin{aligned}\lambda_2 &\approx \sum_{i=1}^N e_{2i}\left\{(N-n)p_l\tilde{e}_I - np_l\sum_{J\neq I}^M \tilde{e}_J\right\}\\ &= \sum_{i=1}^N e_{2i}\left\{Np_l\tilde{e}_I - np_l\sum_{J=1}^M \tilde{e}_J\right\}\\ &\approx \sum_{I=1}^M n\tilde{e}_I\left\{Np_l\tilde{e}_I - np_l\sum_{J=1}^M \tilde{e}_J\right\}\\ &= Np_l\sum_{I=1}^M n\tilde{e}_I^2 - \left(n\sum_{J=1}^M \tilde{e}_J\right)^2 p_l.\end{aligned}$$

Note that $\Sigma_{I=1}^M n\tilde{e}_I^2 \approx \Sigma_{i=1}^N e_{2i}^2 = 1$, and $n\Sigma_{J=1}^M\tilde{e}_J = \Sigma_{i=1}^N e_{2i} = 0$ ($\mathbf{G}$ is symmetric for this type of coupling). We obtain, finally,

$$\lambda_2 \approx Np_l \tag{12}$$

for $p_l \ll p_s$ so that the clustered structure of the network is maintained.

B. Type-II coupling

Type-II coupling is defined by the following normalized Laplacian matrix: For any $i(1\leq i\leq N)$, $G_{ii}=1$, $G_{ij}=-1/k_i$ ($i\neq j$) if there is a link between node i and j, and $G_{ij}=0$ otherwise. For such a matrix, if $N\geq 2$ and the network is connected, then $0<\lambda_2\leq N/(N-1)$ and $N/(N-1)\leq\lambda_N\leq 2$.[5,29] λ_2 is more crucial in determining network synchronizability than λ_N is, because a slight change in λ_2 could lead to drastic change in the eigenratio Q, while the change of λ_N will not. Therefore, in the following, we estimate λ_N in one way, and estimate λ_2 in another more accurate way.

For λ_N, note that $\mathbf{G}=\mathbf{I}-\mathbf{D}^{-1}\mathbf{A}$, where $\mathbf{I}$ is the unit matrix, $\mathbf{D}=\mathrm{diag}\{k_1,\ldots,k_N\}$, and $\mathbf{A}$ is the adjacency matrix. For a random network, its spectrum follows the Wigner semicircle law.[30] The minimum eigenvalue of $\mathbf{A}$ is thus given by

$$\begin{aligned}\lambda_{\min}^A &= -2\sqrt{np_s(1-p_s)+(N-n)p_l(1-p_l)}\\ &\approx -2\sqrt{np_s+(N-n)p_l}\\ &= -2\sqrt{\langle k\rangle}.\end{aligned}$$

Because of the narrow degree distribution, we have $k_i\approx\langle k\rangle$, which leads to[5,31]

$$\lambda_N \approx 1 - \lambda_{\min}^A/\langle k\rangle \approx 1 + \frac{2}{\sqrt{\langle k\rangle}}. \tag{13}$$

For the smallest nontrivial eigenvalue, λ_2 can be obtained in a more precise manner from Eq. (11). In particular, recall that for type-II coupling, $G_{ii}=1$, and if i and j belong to the same cluster, G_{ij} equals $-1/k_i$ with probability p_s and 0 with probability $1-p_s$, while if they belong to different clusters, G_{ij} equals $-1/k_i$ with probability p_l and 0 with probability $1-p_l$. Using $1-np_s/k_i=(N-n)p_l/k_i$ and performing a similar analysis as for the case of type-I coupling, we obtain

$$\lambda_2 \approx \frac{Np_l}{np_s + (N-n)p_l} = \frac{Np_l}{\langle k \rangle}. \quad (14)$$

Numerical results show, indeed, that Eq. (14) predicts more accurately λ_2 than the random-matrix prediction $\lambda_2 \approx 1 - 2/\sqrt{\langle k \rangle}$.

IV. SCALABILITY OF CLUSTERED NETWORKS: THEORY AND NUMERICAL SUPPORT

The synchronization-based scalability of a random clustered network can be analyzed by exploring how the key eigenvalues λ_N and λ_2 of the coupling matrix vary as the size of the network is increased. There are two ways by which the network size $N=nm$ can be increased: either n or m is increased. In addition, for a clustered network with fixed intracluster connecting probability p_s, there are two distinct situations. First, the intercluster connection probability p_l is fixed. In this case, the average number of intercluster links per node μ increases with the network size $N=nm$. Second, μ is fixed. In this case, when N is increased, the probability p_l needs to be decreased accordingly. With the two types of coupling schemes treated here, there are *eight* distinct combinatorial cases of interest. In the following, we will analyze each case and provide numerical support. Our approach will be as follows. Recall that, insofar as $Q=\lambda_N/\lambda_2 < K_2/K_1$, there is a finite parameter interval $(\varepsilon_1, \varepsilon_2)$, where $\varepsilon_1 = K_1/\lambda_2$ and $\varepsilon_2 = K_2/\lambda_N$, within which the oscillator network is synchronizable. We shall then focus on λ_N and λ_2, investigate when the condition $Q < K_2/K_1$ is satisfied, and plot ε_1 and ε_2 as functions of N to reveal the synchronizable (scalable) region in the two-dimensional parameter space (N, ε).

For numerical exploration, we shall use the chaotic Rössler oscillators for node dynamics, which is given by $\mathbf{F}(\mathbf{x}) = [-(y+z), x+0.2y, 0.2+z(x-9)]^T$. Parameters adopted here permit a funnel attractor in the phase space and the system is in the chaotic state. The coupling function is chosen to be $\mathbf{H}(\mathbf{x}) = x$. We obtain $K_1 \approx 0.2$, $K_2 \approx 4.62$, and the synchronization boundaries of the system are given by $\varepsilon_1 = 0.2/\lambda_2$ and $\varepsilon_2 = 4.62/\lambda_N$.

A. Scalability for fixed intercluster connecting probability

For each case below, we fix $p_s = 0.3$ and $p_l = 0.01 \ll p_s$ in numerical computations so as to ensure the clustered topology of the network.

1. Type-I coupling

Case 1: Fixing n and varying m. In this case, the size of individual clusters is fixed while the number of clusters is varied. Theoretical results for λ_N and λ_2 can be obtained from Eqs. (10) and (12), as shown by the solid curves in Figs. 1(a) and 1(b). The data points are from numerical computations. There is a reasonable agreement between theory and numerics. In particular, as m is increased, both λ_N and λ_2 increase, but Q decreases, as shown in Fig. 1(c). This means that, insofar as $Q < K_2/K_1$ is satisfied, larger networks are more synchronizable. The synchronization region in the (m, ε) parameter plane can be determined by Eq. (4), as shown in Fig. 2 as the region between the top and the bottom curves (open triangles are numerical results). It can be seen that as the number of clusters is increased, there exists a finite interval $\Delta\varepsilon$ within which the oscillator system can be synchronized. We thus see that for type-I coupling, random clustered networks with fixed cluster size are scalable with respect to synchronization.

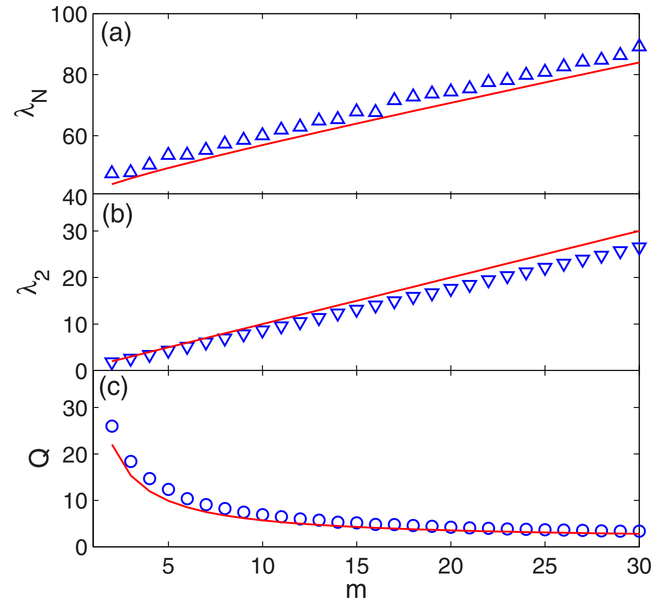

FIG. 1. (Color online) For type-I coupling, fixed cluster size [(a)–(c)] λ_N, λ_2, and Q vs m, the number of clusters, respectively. Simulation parameters are $p_s=0.3$, $p_l=0.01$, and $n=100$. The curves represent theoretical results and data points are numerical results averaged over 10 random network realizations.

Case 2: Fixing m and varying n. In this case, the number of clusters is fixed and the size of the network is controlled by n, the size of each individual cluster. Theoretical and numerical results show that the behaviors of Q and of the critical values of the coupling parameter are similar to those

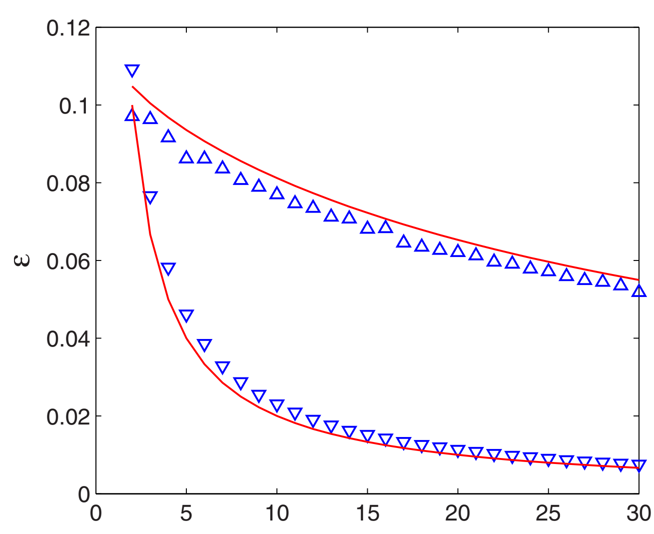

FIG. 2. (Color online) For the same setting as in Fig. 1, synchronizable region in the two-dimensional parameter plane (m, ε) as enclosed by the two curves. Data points are numerical results.

for case 1. We conclude that, for type-I coupling and fixed intercluster connecting probability, a clustered network is scalable with respect to synchronization.

2. Type-I coupling: Scaling theory

We now provide analytic insights into the behaviors of ε for type-I coupling. For this type of coupling, the relevant eigenvalues λ_N and λ_2 are given by Eqs. (10) and (12). Based on these formulas, we can write down the eigenratio Q as

$$Q=\lambda_N/\lambda_2 \approx \frac{\mathrm{erf}^{-1}(1-2/N)\sqrt{2}\sigma+\langle k\rangle+1}{Np_l}.$$

We proceed by making use of the following series expansion for the inverse error function:[32]

$$[\mathrm{erf}^{-1}(x)]^2 \sim \eta-\tfrac{1}{2}\ln\eta+\eta^{-1}\left(\tfrac{1}{4}\ln\eta-\tfrac{1}{2}\right)+\ \ldots, x\to 1,$$

where $\eta=-\ln[\sqrt{\pi}(1-x)]$. This expansion is valid for $x\to 1$, which holds in our problem as $1-2/N\to 1$ for large N. Keeping only the first-order term, we have

$$\mathrm{erf}^{-1}(1-2/N)\approx\sqrt{-\ln(\sqrt{\pi}\cdot 2/N)}=\sqrt{\ln(N/2\sqrt{\pi})}.$$

Substituting this into the expression for Q and omitting irrelevant constants, we have

$$Q=\frac{\sqrt{2\ln(N/2\sqrt{\pi})}\sigma+\langle k\rangle+1}{Np_l}\approx\frac{\sqrt{nS}+np_s+mnp_l}{mnp_l},$$

where $S=[p_s(1-p_s)+mp_l]\ln(mn)$. We see that Q is essentially independent of m and n when they become large. But when we fix n and increase m, for instance, in order to maintain the clustered structure, m should be smaller than $m_{\max}=p_s/p_l+1$ (see Sec. V). Substituting this expression of $m_{\max}$ in Eq. (15), we get

$$Q\approx 2+\sqrt{\frac{2-p_s}{np_s}\ln\frac{np_s}{p_l}}\ \text{for fixed } n.$$

For fixed m, the asymptotical behavior of Q can be given as

$$Q\approx\frac{p_s}{mp_l}+1\ \text{for fixed } m,$$

which does not depend on n and tends to a constant. The size $\Delta\varepsilon$ of the synchronizable region is then given by

$$\Delta\varepsilon\approx\frac{K_2p_l-K_1[\sqrt{S/(m^2n)}+p_s/m+p_l]}{p_l(\sqrt{nS}+np_s+mnp_l)}.$$

For fixed n, the leading term of $\Delta\varepsilon$ scales with m as m^{-1}. Making use of the expression for $m_{\max}$, we get

$$\Delta\varepsilon\approx\frac{K_2np_s-K_1[\sqrt{np_s(1-p_s)\ln(np_s/p_l)}+2np_s]}{np_s[\sqrt{np_s(1-p_s)\ln(np_s/p_l)}+2np_s]}.$$

For fixed m we then have

$$\Delta\varepsilon\approx\frac{K_2mp_l-K_1(p_s+mp_l)}{mnp_l(p_s+mp_l)},$$

which scales with n as n^{-1}.

3. Type-II coupling

Case 3: Fixing n and varying m. In this case, λ_N decreases as there are more clusters in the network, versus the cases associated with the type-I coupling where this eigenvalue increases as the network grows. Meanwhile, λ_2 increases with m, the eigenratio Q actually decreases with m, indicating that larger networks are more synchronizable. Both theoretical and numerical results show that the synchronizable coupling interval increases with m. Since the numerical results appear quite similar to those in Figs. 1 and 2, here we shall provide a scaling theory for type-II coupling.

Case 4: Fixing m and varying n. For the eigenvalues λ_N and λ_2 and the ratio Q, behaviors similar to those in case 3 have been observed. As a result, the synchronizable region in the parameter plane (ε,n) shows a similar pattern too: the underlying oscillator network is scalable.

4. Type-II coupling: Scaling theory

For type-II coupling, we have

$$Q=\frac{(1+2/\sqrt{\langle k\rangle})\langle k\rangle}{Np_l}\approx\frac{\langle k\rangle+2\sqrt{\langle k\rangle}}{Np_l}\approx\frac{np_s+nmp_l+2\sqrt{np_s+nmp_l}}{nmp_l},\qquad(15)$$

where $\langle k\rangle$ is the average degree of the network that can be calculated from Eq. (8). Apparently, Q depends neither on n nor on m when the system size becomes infinite. Let $m=m_{\max}$ be the critical value of the number of clusters above which the clustered structure cannot be maintained. We have

$$Q\approx 2[1+\sqrt{2/(np_s)}]\ \text{for fixed } n.$$

When we fix m and increase n, the asymptotic value of Q can be obtained as

$$Q\approx\frac{p_s}{mp_l}+1,\ \text{for fixed } m.$$

The synchronizable coupling-parameter interval $\Delta\varepsilon$ can then be calculated as

$$\Delta\varepsilon\approx K_2-K_1\frac{np_s+nmp_l}{nmp_l},$$

where the leading term is independent of n and m. For $m=m_{\max}$, we have

$$\Delta\varepsilon\approx K_2-2K_1,\ \text{for fixed } n.$$

If m is fixed but n is increased, we have, asymptotically,

$$\Delta\varepsilon\approx K_2-K_1\frac{p_s+mp_l}{mp_l}.$$

For type-II coupling, when we fix m (or n) and increase n (or m), a finite interval in the coupling parameter always exists for which the network is synchronizable. The clustered networks are thus scalable. Taking into account the results for type-I coupling, we can conclude that, for fixed intercluster connecting probability, the networks are scalable for both type-I and type-II coupling schemes. In particular, the eigenratio Q tends to a constant value as the network grows, and the synchronizable parameter interval $\Delta\varepsilon$ is inversely proportional to the network size for type-I coupling but it tends to a constant for type-II coupling as the system size N is increased.

B. Scalability for fixed average number of intercluster connections

The average number of intercluster connections is

$$\mu = \frac{n^2 m(m-1)p_l}{mn} = n(m-1)p_l.$$

When we fix μ to grow the network, the actual intercluster connection probability

$$p_l = \frac{\mu}{n(m-1)} \qquad (16)$$

will be reduced, but our theoretical results in Sec. III for the eigenvalues are still applicable.

1. Type-I coupling

Cases 5 and 6: Fixing n (or m) and varying m (or n). In these cases, λ_N is still given by Eq. (10), except that σ and $\langle k\rangle$ now become

$$\sigma^2 = np_s(1-p_s) + \mu\left[1 - \frac{\mu}{n(m-1)}\right],$$
$$\langle k\rangle = np_s + \mu, \qquad (17)$$

and λ_2 can be calculated from

$$\lambda_2 = \frac{m\mu}{m-1}. \qquad (18)$$

When we fix μ, there is no required maximum value of m (see Sec. IV C). We thus only need to discuss the behavior of Q as the network size is increased. From Eqs. (10), (17), and (18), we have

$$Q \approx \frac{1}{\mu}\{\sqrt{[p_s(1-p_s)+mp_l]\ln(mn)} + np_s\}, \qquad (19)$$

which scales as $\sqrt{m\ln m}$ for fixed n and as n for fixed m. Thus, for type-I coupling, when the network grows, Q always increases. As the network size increases through a critical value, there exists no interval in the coupling parameter for which the network can be synchronized, indicating a loss of scalability. These behaviors have been verified numerically.

2. Type-II coupling

Cases 7 and 8: Fixing n (or m) and varying m (or n). Substituting Eq. (16) into both Eq. (13) and Eq. (14), we get

$$\lambda_N \approx 1 + \frac{2}{\sqrt{\langle k\rangle}} = 1 + \frac{2}{\sqrt{np_s+\mu}} \qquad (20)$$

and

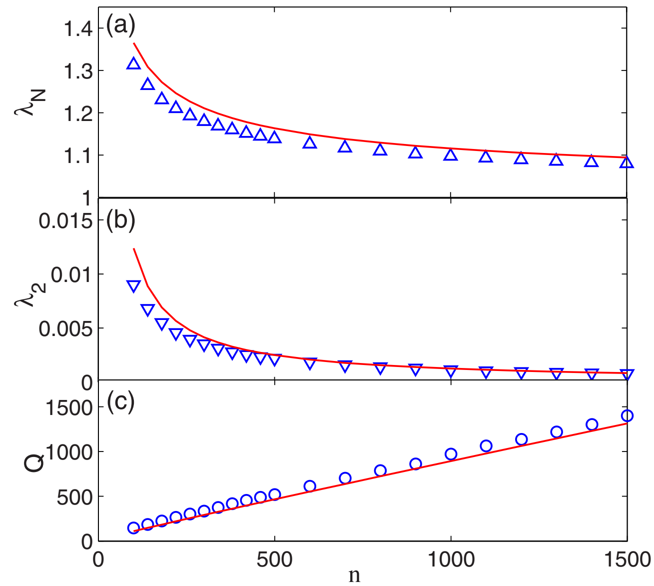

FIG. 3. (Color online) For fixed μ=0.3, type-II coupling, clustered networks of m=5 clusters [(a)–(c)] λ_N, λ_2 and Q vs n, respectively.

$$\lambda_2 \approx \frac{m\mu}{(m-1)\langle k\rangle} \approx \frac{m\mu}{(m-1)(np_s+\mu)}. \qquad (21)$$

For large n (or m), the leading term of Q can be written as

$$Q \approx \frac{m-1}{m\mu}(np_s + 2\sqrt{np_s}),$$

which does not depend on m for sufficiently large values of m, but it increases with n as the network size is increased. Therefore, for fixed n, when m is increased, Q first increases and then approaches asymptotically a constant. But, for fixed m, the eigenratio increases linearly with n, indicating a quick loss of the network synchronizability as n becomes large. A representative example is shown in Fig. 3.

We see that n and m have different influence on Q for different cases. For example, when m is increased, Q increases as $\sqrt{m\ln(m)}$ for type-I coupling and as $(m-1)/m$ for type-II coupling. These increases are much slower than $Q \sim n$ when n is increased for fixed m. Thus, growing a clustered network by increasing the size of individual clusters can be much more effective to suppress synchronization than increasing the number of clusters.

From the above analysis, we can see that for the type of growing scheme defined by fixing μ, Q increases for both types of the coupling schemes.[33] Because of this, although small networks may be synchronizable, the synchronizability will be lost for larger networks. Clustered networks under the constraint of fixed μ are thus not scalable.

C. Scalability and deterioration of clustered characteristics

The above results are based on the assumption that the networks considered possess a clustered topology. An interesting question is whether the clustered structure can be retained when the network grows.

By definition, a clustered network requires that the intracluster connections be denser than the intercluster connections. Defining ν and μ as the average numbers per node of intra/intercluster connections, respectively, we need $\nu > \mu$. For our clustered network model, ν and μ are given by

$$\nu = (n-1)p_s,$$

$$\mu = n(m-1)p_l,$$

which leads to the following condition for clustered structure:

$$\frac{\nu}{\mu} = \frac{(n-1)p_s}{n(m-1)p_l} > 1. \tag{22}$$

According to Eq. (22), the presence of the clustered topology depends on four parameters: n, m, p_l, and p_s. (In this paper p_s is fixed.)

First consider the setting where p_l is fixed. If we fix n and increase m, the condition guaranteeing a clustered network structure becomes

$$m < \frac{(n-1)p_s}{np_l} + 1 \approx \frac{p_s}{p_l} + 1,$$

which depends only on the ratio of p_s and p_l. For $m \geqslant p_s/p_l+1$, the clustered structure no longer exists. For the typical numerical setting we have used, the parameters are $p_s=0.3$ and $p_l=0.01$. In order to ensure the clustered characteristics, the value of m should not exceed $p_s/p_l+1 = 0.3/0.01+1=31$. This rule has been followed in all our numerical examples.

If we fix m and increase n, the clustered condition becomes

$$n[p_s-(m-1)p_l] > p_s.$$

Since Eq. (22) implies $(m-1)p_l < [(n-1)/n]p_s < p_s$, we have $p_s-(m-1)p_l > 0$ and, hence,

$$n > \frac{p_s}{p_s-(m-1)p_l},$$

which can usually be satisfied. For example, for $m=5$, $p_s=0.3$, and $p_l=0.01$, the requirement is $n \geqslant 2$. (Typical values of n used in our simulations are two orders of magnitude larger.)

We remark, however, that for fixed p_l, the clustered topology can be maintained *if the number of clusters is small*. In realistic networked systems this number may be large. While networks generated for fixed value of p_l are scalable with respect to synchronization, the clustered topology is lost as the network becomes large if both the number of clusters and the number of nodes in each cluster grow.

Second, we consider the case where the average number of intercluster links μ is fixed. In this case, $p_l=\mu/[n(m-1)]$ decreases with network size. The condition for the clustered structure is

$$\frac{\nu}{\mu} = \frac{(n-1)p_s}{\mu} > 1.$$

We see that ν/μ is independent of m. There is thus no requirement on m to ensure the clustered structure. If we fix m and increase n, the ratio of ν/μ will become larger. The condition becomes

$$n > \frac{\mu}{p_s} + 1.$$

For example, for the parameters used in our numerical examples ($\mu=0.3$ and $p_s=0.3$), the requirement is $n>2$, which is always guaranteed. The conclusion is that, although the clustered topology can be maintained by fixing μ, the scalability is lost.

V. CONCLUSIONS

We have addressed the scalability of complex clustered networks by investigating the size dependence of the network synchronizability. The general conclusion is that such networks are not scalable with respect to synchronization. In particular, if the probabilities of intracluster and intercluster connections are fixed, larger networks are actually more synchronizable. In this case, however, the number of intercluster links increases with the network size and, as such, a sufficiently large network may not exhibit the distinct feature of being clustered. On the other hand, if the average number of intercluster links is fixed, the network synchronizability deteriorates quickly as the network size becomes larger. A practical implication is that, for typical clustered networks, if synchronization is important for the system function, the clustered topology is undesirable.[34] We hope these results are useful for the exploration of dynamics on complex clustered networks.

An important issue concerns a possible time delay in the coupling function,[35] as interactions between dynamical units in realistic physical systems cannot be instantaneous. When the coupling is time-delayed, the synchronization manifold still exists, so its stability can be analyzed. In particular, while the master-stability function needs to be determined from a set of variational equations that contain time delay, one can still define a generalized coupling parameter as the product between the original coupling parameter and the eigenvalues of the coupling (Laplacian) matrix. Hence, although a time delay can cause a shift or a change in the interval where the master-stability function is negative, the eigenratio is determined solely by the network topology and can still be used to characterize the network synchronizability. We thus expect our results to hold for complex clustered networks where there is a time delay in the interactions among nodes.

ACKNOWLEDGMENTS

L.H. and Y.C.L. are supported by AFOSR under Grant No. FA9550-07-1-0045, by ONR through WVHTC (West Virginia High Technology Consortium Foundation), and by an ASU-UA Collaborative Program on Biomedical Research.

X.J.M. and Y.W.'s visit to Arizona State University is sponsored by the State Scholarship Fund of China. X.J.M., Y.W., and Z.G.Z. are partially supported by NNSF of China Grant Nos. 70431002 and 10575010, the 973 Program No. 2007CB814805, and the Foundation of Doctoral Training No. 20060027009.

[1]L. F. Lago-Fernandez, R. Huerta, F. Corbacho, and J. A. Siguenza, Phys. Rev. Lett. **84**, 2758 (2000).
[2]M. B. Barahona and L. M. Pecora, Phys. Rev. Lett. **89**, 054101 (2002).
[3]X. F. Wang and G. Chen, Int. J. Bifurcation Chaos Appl. Sci. Eng. **12**, 187 (2002); IEEE Trans. Circuits Syst., I: Fundam. Theory Appl. **49**, 54 (2002).
[4]T. Nishikawa, A. E. Motter, Y.-C. Lai, and F. C. Hoppensteadt, Phys. Rev. Lett. **91**, 014101 (2003).
[5]A. E. Motter, C. Zhou, and J. Kurths, Europhys. Lett. **69**, 334 (2005); Phys. Rev. E **71**, 016116 (2005); C. Zhou, A. E. Motter, and J. Kurths, Phys. Rev. Lett. **96**, 034101 (2006).
[6]L. Donetti, P. I. Hurtado, and M. A. Muñoz, Phys. Rev. Lett. **95**, 188701 (2005).
[7]F. M. Atay, T. Bıyıkoğlu, and J. Jost, IEEE Trans. Circuits Syst., I: Regul. Pap. **53**, 92 (2006).
[8]T. Nishikawa and A. E. Motter, Phys. Rev. E **73**, 065106 (2006).
[9]L. Huang, K. Park, Y.-C. Lai, L. Yang, and K. Yang, Phys. Rev. Lett. **97**, 164101 (2006).
[10]X. G. Wang, Y.-C. Lai, and C.-H. Lai, Phys. Rev. E **75**, 056207 (2007).
[11]S. Strogatz, *Sync: The Emerging Science of Spontaneous Order* (Hyperion, New York, 2003).
[12]D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998).
[13]A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[14]S. H. Strogatz, Nature (London) **410**, 268 (2001).
[15]R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
[16]M. Girvan and M. E. J. Newman, Proc. Natl. Acad. Sci. U.S.A. **99**, 7821 (2002).
[17]D.-H. Kim and A. E. Motter, Phys. Rev. Lett. **98**, 248701 (2007).
[18]L. Huang, Y.-C. Lai, and R. A. Gatenby, Chaos **18**, 013101 (2008).
[19]D. J. Watts, P. S. Dodds, and M. E. J. Newman, Science **296**, 1302 (2002).
[20]E. Ravasz, A. L. Somera, D. A. Mongru, Z. Oltvai, and A.-L. Barabási, Science **297**, 1551 (2002).
[21]V. Spirin and L. A. Mirny, Proc. Natl. Acad. Sci. U.S.A. **100**, 12123 (2003).
[22]A. E. Motter, T. Nishikawa, and Y.-C. Lai, Phys. Rev. E **68**, 036105 (2003).
[23]G. Palla, I. Derényi, I. Farkas, and T. Vicsek, Nature (London) **435**, 814 (2005).
[24]E. Oh, K. Rho, H. Hong, and B. Kahng, Phys. Rev. E **72**, 047101 (2005).
[25]B. Bollobás, *Random Graphs* (Academic, London, 1985).
[26]L. M. Pecora and T. L. Carroll, Phys. Rev. Lett. **80**, 2109 (1998).
[27]To illustrate the applicability of Eqs. (4) and (5) in determining the network synchronizability, we have performed the following set of numerical experiments. For the Rössler-type of chaotic node dynamics, the coupling-parameter ratio K_2/K_1 has approximately the value of 23.1. Our idea is then to consider two networks, one with an eigenratio less than 23.1 and another with an eigenratio larger than this value. For example, for a ring network of 10 nodes, the eigenratio is $Q=\lambda_N/\lambda_2 \approx 4.0/0.382 \approx 10.48 < 23.1$ but for a ring network of 20 nodes, the eigenratio becomes $Q \approx 40.9 > 23.1$. Thus, according to the criterion in Eq. (5), the first network is synchronizable and the second is not. Direct numerical simulations of the network dynamics indicate that this is indeed so.
[28]L. F. Costa, F. A. Rodrigues, G. Travieso, and P. R. Villas Boas, Adv. Phys. **56**, 167 (2007).
[29]F. R. K. Chung, *Spectral Graph Theory* (AMS, Providence, 1994).
[30]E. P. Wigner, Ann. Math. **53**, 36 (1951).
[31]F. Chung, L. Lu, and V. Vu, Proc. Natl. Acad. Sci. U.S.A. **100**, 6313 (2003).
[32]J. M. Blair, C. A. Edwards, and J. H. Johnson, Math. Comput. **30**, 827 (1976).
[33]For type-II coupling and fixed n, although our theory predicts that Q tends to a constant value for sufficiently large m, Q increases for small values of m.
[34]In a typical situation, the increase of the total size (N) of a clustered network is accompanied by the growth both of the number of clusters and of the number of nodes in an individual cluster.
[35]M. Dhamala, V. K. Jirsa, and M. Ding, Phys. Rev. Lett. **92**, 074104 (2004).