Eigenvalue spectra of spatial-dependent networks

Joris Billen, Mark Wilson, and Arlette Baljon
Department of Physics, San Diego State University, San Diego, California 92128, USA

Avinoam Rabinovitch
Department of Physics, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel
(Received 8 May 2009; revised manuscript received 14 August 2009; published 21 October 2009)

Many real-life networks exhibit a spatial dependence; i.e., the probability to form an edge between two nodes in the network depends on the distance between them. We investigate the influence of spatial dependence on the spectral density of the network. When increasing spatial dependence in Erdös-Rényi, scale-free, and small-world networks, it is found that the spectrum changes. Due to the spatial dependence the degree of clustering and the number of triangles increase. This results in a higher asymmetry (skewness). Our results show that the spectrum can be used to detect and quantify clustering and spatial dependence in a network.

DOI: 10.1103/PhysRevE.80.046116 PACS number(s): 89.75.Hc, 64.60.aq, 02.60.−x, 05.10.−a

I. INTRODUCTION

An increase in computational resources has led to a considerable interest in complex networks over the last decade. Initially most studies handled networks in the dimensionless network space. Many real-life networks, however, live in a geographic space in which it is more favorable to form edges between nodes that are close to each other. For this reason interest in these spatial networks (SN) has increased in the last few years. SN can be found in the fields of communication (Internet [1]), biology (neural networks [2]), transportation (airport [3,4], rail [5], and road networks [3]), social networks [6] (friendships), and disease spreading [7]. To determine the existence of spatial dependency of a network one can look at several measures [3]. A spatial measure is the distribution of the Euclidean distance (ED) between nodes. In contrast, the so-called graph distance measures the number of edges traversed along the shortest path from one vertex to another (path length). In this work we propose a more prominent method to detect spatial dependence based on the spectral density of a network.

The eigenvalue spectrum of the adjacency matrix of a graph contains information related to important topological features of the graph. Therefore, it could also reflect the structural changes induced by spatial dependence. Eigenvalue spectra have been extensively studied for most common network models such as Erdös-Rényi (ER) random graphs, small-world (SW) networks, and scale-free networks [8]. In this work we study the influence of spatial dependence on the spectra of these networks in Euclidean space. In Sec. II we define the spectral density and its properties. In Sec. III, the models are introduced. The resulting spectra are presented in Sec. IV. Section V discusses the results.

II. EIGENVALUE SPECTRUM OF NETWORKS

The spectrum of a graph is the set of eigenvalues, λ_j, of a graph's adjacency matrix A [8]. The graphs under investigation are undirected and devoid of loops and multiple edges. Hence, the adjacency matrix is real and symmetric, possessing real eigenvalues [9]. The spectral density of a graph with N nodes can be defined as

$$\rho(\lambda) = \frac{1}{N}\sum_{j=1}^{N} \delta(\lambda - \lambda_j). \tag{1}$$

Since the spectrum contains all the topological information of the graph, it can be used to classify the network. The spectra of ER random, scale-free, and small-world structures in dimensionless network space have been studied extensively [8]. For the ER random network, the spectral density exhibits a Wigner semicircle. The scale-free network displays a symmetric triangular bulk spectrum. A small-world network is constructed by placing nodes on a circle, connecting the k nearest neighbors (NN), and then randomly rewiring each edge with a probability p [10]. At $p=0$ the small-world network has a regular structure and at $p=1$ it becomes ER. The spectrum of the small-world network exhibits several peaks for small p values because of its regularity [8] and it transitions to the semicircular shape when p approaches 1.

The moments of the spectral density of a graph are related to its topology and the sth moment of $\rho(\lambda)$ can be written as

$$m_s = \frac{1}{N}\sum_{j=1}^{N} (\lambda_j - \mu)^s = \int (\lambda - \mu)^s \rho(\lambda) d\lambda, \tag{2}$$

with μ as the mean eigenvalue. Since the adjacency matrix contains no loops, its trace will be zero and hence $\mu=0$. $D_s=Nm_s$ is the so-called "number of directed paths of the graph that return to their starting vertex after s steps" [11].

Skewness and kurtosis are often used to describe shape characteristics of a distribution [12] and can be used to characterize the spectra. The skewness S is a measure of the asymmetry of a distribution and is defined as

$$S = \frac{m_3}{m_2^{3/2}} = \frac{N^{-1}\sum_i \lambda_i^3}{\sigma^3}, \tag{3}$$

with $\sigma^2=m_2=\langle k\rangle$ the standard deviation [9]. The kurtosis K is a measure of the peakedness of a distribution and is defined as

TABLE I. Average ED for spatial-dependent networks.

α	0	2	4	6	8	10
ER/SW, $p=1.00$	0.480	0.328	0.169	0.120	0.109	0.105
SW, $p=0.75$	0.386	0.278	0.278	0.278	0.107	0.104
SW, $p=0.50$	0.291	0.224	0.143	0.113	0.106	0.103
SW, $p=0.25$	0.195	0.165	0.125	0.109	0.104	0.102
Lowest cost/ER $\alpha\to\infty$	0.100	0.100	0.100	0.100	0.100	0.100
SF	0.480	0.339	0.196	0.147	0.133	0.128

$$K=\frac{m_4}{m_2^2}-3. \tag{4}$$

For a Gaussian distribution $m_4/m_2^2=3$ and $K=0$. A semicircular distribution is known to have $S=0$ and $K=-1$.

III. CONSTRUCTION OF SPATIAL ER, SCALE-FREE, AND SMALL-WORLD NETWORKS IN EUCLIDEAN SPACE

To construct spatial ER, scale-free, and small-world networks, N nodes are randomly placed in Euclidean space in a $1\times1\times1$ box. Then l edges are created, favoring ones of shorter ED's, leading to a network with average degree $\langle k\rangle=2l/N$.

Spatial dependence is obtained by choosing the probability $p_{i,j}$ to form an edge between two nodes i and j depending on the ED $d_{i,j}$ between them:

$$p_{i,j}\propto\frac{d_{i,j}^{-\alpha}}{\sum_{b,c} d_{b,c}^{-\alpha}}. \tag{5}$$

The strength of the spatial selection is determined by the value of the "proximity factor" α. When $\alpha=0$ there is no spatial dependence. For $\alpha\to\infty$ only the closest edges will be chosen. We now discuss in more detail the construction method for each of the three networks under investigation. In our simulations we use $N=1000$ and $l=5000$ resulting in $\langle k\rangle=10$. All data result from averages over 100 configurations.

A. Spatial ER network

In a spatial ER network, each possible edge receives a connection probability as in Eq. (5) favoring nearby nodes ($\alpha>0$). Once an edge is formed, the probability for creating that edge is set to zero and the new probabilities are calculated. Repeating this process l times, a network with a Poisson connectivity distribution with a peak at $\langle k\rangle=2l/N$ is obtained. Alternatively, we could have formed an edge between each pair of nodes with a certain probability p [9]. This would have given roughly the same results.

Table I shows the average ED for different values of α. For higher α, shorter edges are formed. The clustering coefficient [13] as a function of average path length over all nodes is shown in Fig. 1(a). The effect of proximity is visible by a strong increase in the clustering coefficient and a moderate increase in path length.

B. Spatial scale-free network

For constructing a spatial scale-free network we follow a procedure based on Ref. [14]:

(1) Select at random a subset of n_0 nodes and connect them. Nodes that have connections are called active.

(2) Take an inactive node i at random and connect it with an active node j with probability (up to a normalization factor)

$$p_{i,j}\propto(k_j+1)d_{i,j}^{-\alpha}, \tag{6}$$

where k_j is the degree of node j and $d_{i,j}$ is the ED between nodes i and j.

For each of the nodes, we repeat step (2) $m=5$ times until all nodes are active. The degree distribution for $\alpha=0$ shows a drop-off with a power law and $\langle k\rangle=2m=10$. For large values of α the proximity effect limits the choice of available connections thereby limiting the degree distribution, resulting in a deviation from a power-law behavior, as predicted in [14].

The effect of proximity is an increase in the clustering coefficient [Fig. 1(b)]. The path length is smaller than that of

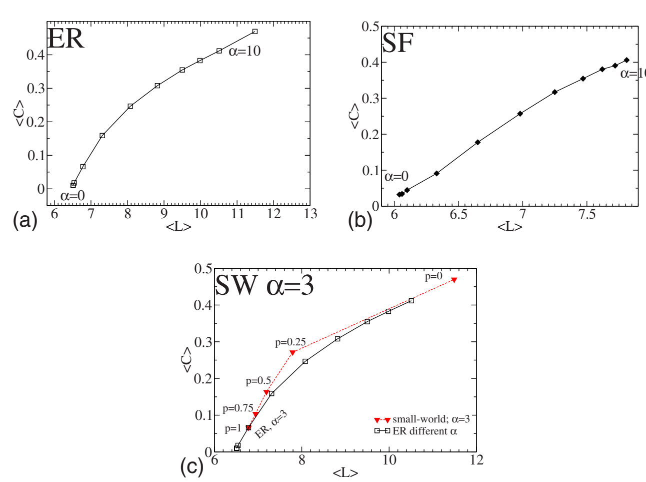

FIG. 1. (Color online) The average clustering coefficient over all nodes as a function of the path length for spatial-dependent networks. (a) Random ER for different α. The data point for $\alpha\to\infty$ corresponds to the lowest cost network. (b) Spatial-dependent scale-free network for different α. (c) Small-world network for $\alpha=3$ and $\alpha=6$ for different p. For comparison, data of Fig. 1(a) are shown as well.

an ER network due to the presence of hubs with a very high degree.

C. Spatial small-world network

In a dimensionless SW network, a regular structure is obtained by placing N nodes on a ring and connecting each of them to its $\langle k\rangle$ nearest neighbors [10]. Then each edge is rewired with a probability p. For intermediate p the system shows the small-world property characterized by a small path length and a high clustering coefficient. For $p=1$ the ER random network is obtained, exhibiting a small path length and a low clustering coefficient. For $p=0$ the regular network stays intact.

For the small-world network in 3D Euclidean space, we create a lowest cost configuration as the analogy of the regular structure. A lowest cost structure is found in brain networks where the cost of edges is optimized [15]. This structure is obtained by calculating the distance between all possible pairs of nodes and connecting the $N\langle k\rangle/2$ ones that are closest to each other. The construction method for the spatial SW is therefore as follows:

(1) Calculate the ED between every possible pair of nodes and connect the $N\langle k\rangle/2$ closest ones. This is the lowest cost network.

(2) Consider rewiring each edge with a certain probability p. If rewiring takes place, the current edge is destroyed and replaced by a new edge, while closer edges are favored depending on α according to Eq. (5). During the rewiring process, creation of every possible edge is allowed, but only once [16]. At the end of this rewiring cycle double edges can exist (except for $p=0$ or $p=1$).

(3) The "doubles" that exist after the rewiring cycle are now rewired, but this time only nonexisting edges are allowed. Hence, in the final configuration there are no multiple edges.

The network is completely determined by two parameters: the rewiring probability p and the proximity factor α. Note that the SW with proximity network for $p=1$ and a certain α will correspond to the spatial ER for the same α. We see from Fig. 1(c) for $\alpha=3$ that the clustering coefficient for the small-world network is always higher than that for the ER for intermediate values of p. For $p=0$ and $p=1$ both networks have the same clustering coefficient. $p=0$ is the lowest cost network with the shortest edges possible and is identical with the ER network with $\alpha\to\infty$. For $\alpha=5$ similar results are found. Again, Table I shows that for higher proximity, the average ED decreases.

D. Eigenvalue spectrum of spatial networks

The spectral densities for the ER random network for different values of α are shown in Fig. 2(a). As expected for $\alpha=0$ we find a semicircle. For $\alpha=5$ the spectrum is asymmetric. The peak shifts to the left and the right tail becomes fat. For $\alpha=10$, -1 is the most abundant eigenvalue.

The spectrum for the scale-free network is shown in Fig. 2(b). Without proximity a triangular shape is found. For increased α the peak of the spectrum shifts to the left while the right-hand tail becomes fatter, and for $\alpha=10$ the peak is at -1.

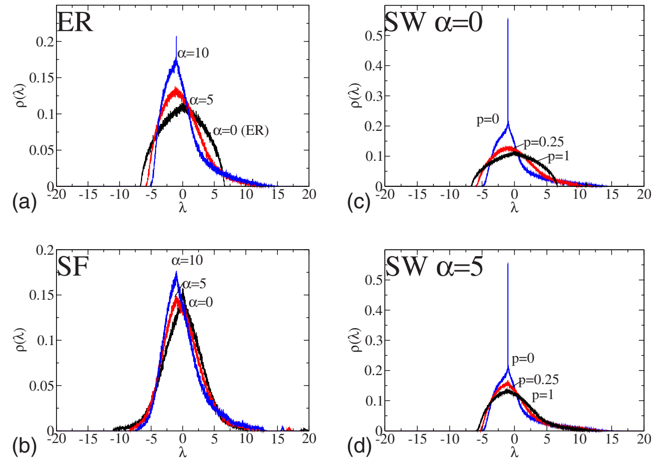

FIG. 2. (Color online) Eigenvalue spectra for spatially dependent networks: (a) ER, (b) scale-free, (c) small-world for $\alpha=0$, and (d) small-world for $\alpha=5$.

The small-world network with no proximity is shown in Fig. 2(c). $p=1$ corresponds to an ER network. For lower p values the peak shifts to the left and the right tail becomes heavier. For the lowest cost network, $p=0$, we find a peak at -1 and a very fat tail to the right. The small-world network for $\alpha=5$ is shown in Fig. 2(d). For $p=1$ the small-world network again corresponds to the ER network with $\alpha=5$. For decreasing p there is a transition to the lowest cost network ($p=0$) and the peak at -1 becomes more prominent.

All spectra show an increased asymmetry with increased α. In order to quantify this effect we investigate the skewness S of the spectra in Fig. 3. For the semicircle the skewness is slightly higher than 0. This is due to the finite system size. As we will argue in Sec. IV, for $N\to\infty$, $S\to 0$ in an ER network. For increased α the network becomes more positively skewed and S increases. The lowest cost network shows the highest S. Next, we look at the peakedness in terms of the kurtosis K. For the small-world and the ER random networks, the kurtosis increases with α. For the semicircle K is close to -1 as expected. The lowest cost network has the highest K. For the scale-free network the kurtosis decreases for small α and then increases for higher α.

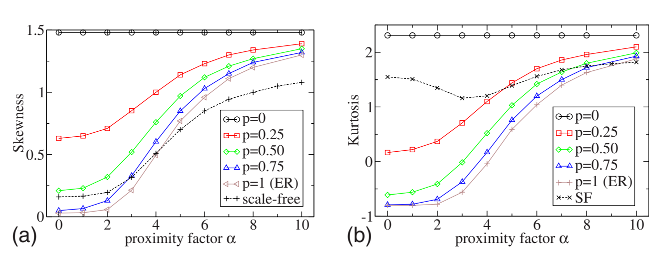

FIG. 3. (Color online) Skewness (a) and kurtosis (b) for spatial-dependent networks.

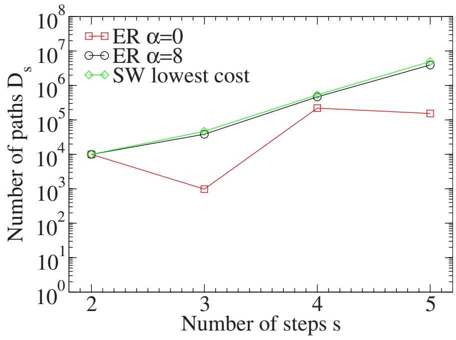

FIG. 4. (Color online) The number of directed paths D_s in an ER network without proximity (squares), with strong proximity (α=8, circles), and the lowest cost network (diamonds).

IV. DISCUSSION

A. Skewness

To understand the observed asymmetry in Fig. 2 and the increase in skewness with α in Fig. 3(a), we go back to the definition of skewness, Eq. (3). This can be transformed into

$$S = \frac{D_3}{N\sigma^3}. \tag{7}$$

Since $\sigma^2 = m_2 = \langle k \rangle$ [9] is independent of α, the skewness is directly dependent on the number of directed paths (DP) starting from a vertex and returning to that vertex after three steps, D_3. We studied the general D_s behavior for an ER network, a spatial ER network, and the lowest cost network (Fig. 4). The ER network with α=0 has a significantly lower number of DP with odd s than DP of even s. This zigzag pattern is a consequence of the fact that, except for a few connections, a random graph looks like a tree and a tree has no DP of odd length. An ER network has no DP of odd length for $N \to \infty$ [8]. The number of DP with two steps $D_2 = Nm_2 = N\langle k \rangle$ is independent of α. As a result of the spatial dependence there is a strong increase in D_3 and D_5. We can understand the increase in D_3 with α (Fig. 4) by comparing the number of triangles T for the ER network with the number for a spatial network using a regular lattice. On a triangle one can define six directed paths (starting from each of the tree nodes and going either clockwise or counterclockwise). Hence,

$$D_3 = 6T, \tag{8}$$

where T is the number of triangles. A triangle consists of three nodes all having degree k=2. Consider an ER network and start from a specific node. There are on the average $\binom{\langle k \rangle}{2}$ possible choices to pick two of its neighbors. The probability that these two neighbors are connected and a triangle is formed equals the total number of links $l = N\langle k \rangle / 2$ divided by the total number of possible links $N(N-1)/2$. Therefore, the number of triangles in the ER network is

$$T = \frac{N}{3} \frac{\langle k \rangle!}{(\langle k \rangle - 2)!\,2!} \frac{N\langle k \rangle}{2} \frac{2}{N(N-1)} \approx \frac{1}{6} \langle k \rangle^2 \langle k-1 \rangle. \tag{9}$$

The 1/3 comes from overcounting (each triangle has three corners). Combining Eqs. (7)–(9) we find

$$S = \frac{\langle k \rangle^2 \langle k-1 \rangle}{N\sigma^3} = \frac{\langle k \rangle^{1/2} \langle k-1 \rangle}{N}. \tag{10}$$

This shows that the nonzero skewness for the ER with α =0 in Fig. 3 is due to the finite system size. We have verified this finite-size effect numerically. We have also observed that in an "antiproximity" network, where α is negative and distant edges are favored, $S \to 0$ for $\alpha \to -\infty$.

We now consider the number of triangles in a network with proximity. Assume the nodes are placed on a 2D triangular lattice with coordination number z=6 and that only NN can be connected (lowest cost network). The number of triangles is different than the one for a simple ER network since the probability p that the two chosen neighbors are connected is not the same. First of all the probability that the two chosen neighbors are NN of each other equals $\frac{2z}{z(z-1)} = \frac{2}{z-1}$. Second, the number of possible links equals $Nz/2$ in this case. Hence, the number of triangles equals

$$T = \frac{1}{3} N \frac{\langle k \rangle!}{(\langle k \rangle - 2)!\,2!} \frac{2}{z-1} \frac{\langle k \rangle}{z} = \frac{N}{3z(z-1)} \langle k \rangle^2 \langle k-1 \rangle. \tag{11}$$

For large N this is substantially higher than for the ER network. From Eqs. (7), (8), and (10) we find here

$$S = \frac{2\langle k \rangle^2 \langle k-1 \rangle}{\sigma^3 z(z-1)} = \frac{2\langle k \rangle^{1/2} \langle k-1 \rangle}{z(z-1)}. \tag{12}$$

We note that the increase in clustering coefficient with proximity observed in Fig. 1 is also due to an increase in the number of triangles T. The clustering coefficient of a node i with degree k_i is defined as the number of triangles t_i in which vertex i participates, normalized by the maximum possible number of such triangles [13]:

$$c_i = \frac{2t_i}{k_i(k_i - 1)}. \tag{13}$$

In a system with N nodes the average clustering coefficient

$$\bar{C} = \frac{1}{N} \sum_{i=1}^{N} \frac{2t_i}{(k_i)(k_i - 1)}. \tag{14}$$

For large $\langle k \rangle$, $k_i(k_i - 1)$ is sharply peaked around $\langle k_i(k_i - 1) \rangle$. In this case

$$\bar{C} = \frac{1}{N\langle k_i(k_i - 1) \rangle} \sum_{i=1}^{N} 2t_i = \frac{6T}{N\langle k_i(k_i - 1) \rangle}, \tag{15}$$

where we have used that $T = \frac{1}{3}\sum_{i=1}^{N} t_i$, since each triangle contributes to three nodes.

For an ER network with no spatial dependence, T is independent of N; hence, both S and $\bar{C}$ are inversely proportional to N according to Eqs. (10) and (14). The inverse

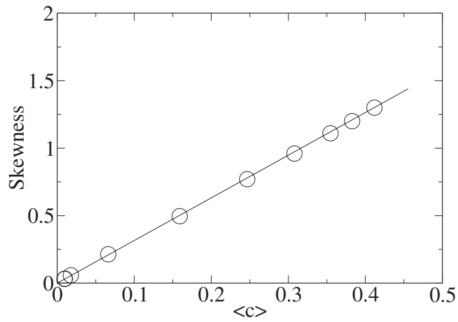

FIG. 5. Skewness as a function of average clustering coefficient for the spatial ER network (circles). Equation (16) is added as a solid line with a slope of 3.16.

dependence of $\bar{C}$ is a well known property [9]. For lowest cost networks, the number of triangles increases linearly with system size N and hence S and $\bar{C}$ are independent of system size. Further, we have performed our analysis in three dimensions, but since Eq. (15) is dependent solely on the number of triangles we expect similar results for different dimensions.

B. Relation between skewness and clustering coefficient

By combining Eqs. (7)–(15) and using $\sigma = \langle k \rangle^{1/2}$ it is observed that

$$S = \frac{6T}{N\sigma^3} = \frac{6T}{N\langle k \rangle^{3/2}} = \bar{C}\frac{\langle k(k-1)\rangle}{\langle k \rangle^{3/2}}. \tag{16}$$

This shows that the skewness is an alternative measurement of the clustering of a network. We have numerically verified Eq. (16) by calculating the skewness as a function of the clustering coefficient for different values of α for the spatial ER network (Fig. 5). The measured values (circles) are in good agreement with the values predicted by Eq. (16) (solid line). For the scale-free network the condition under which Eq. (15) is a good approximation does not hold and we find a slope different than the one expected by Eq. (16).

Asymmetry arises from the increase in the number of triangles in a network. Therefore, any method (not only the introduction of spatial dependence) which increases T would also increase asymmetry. For instance, in a scale-free network increased clustering can be achieved by constructing a network in which, with a certain probability, the step of adding a node with preferential attachment is replaced by the creation of a triangle as described in [17]. We have constructed such networks and found indeed an increase in S, similar to the increase in the spatial ones.

C. Peak at −1

For all spatial networks it is observed that the spectrum peaks at -1 for high α values. This peak is also related to the observed increase in the number of triangles in the spatial network. To verify this hypothesis, we created an ER network and connected some of the dead-end vertices (nodes with degree 1) in two different ways. First, only dead-end vertices that are both connected to a common node were connected hence creating a triangle. Next, dead-end vertices were chosen at random and connected. For the first method we observed that the spectrum was a semicircle with a distinct peak at -1. The second method did not alter the semicircular distribution of the regular ER network. We conclude that the peak at -1 is induced by the spatial nature of the network. For high α this leads to the connection of nodes that are close to each other, resulting in an increase in triangles.

D. Kurtosis

We observed an increase in kurtosis with α for all SW and ER networks [Fig. 3(b)]. For increasing α the peak grows and shifts to the left (Fig. 2).

For the scale-free network, the kurtosis first decreases [Fig. 3(b)]. This is the result of a combination of two effects. First, the pure scale-free network already has a sharp peak and hence high kurtosis around $\lambda = 0$. Second, for high α the degree distribution deviates from a power law (Fig. 1 in [14]). Hence, the network is not scale-free anymore and the sharp peak decreases. Only at sufficiently high α values the peak at -1, that is characteristic of spatial dependence, appears and K increases again.

V. CONCLUSIONS

We have performed a study on the eigenvalue spectrum of spatial networks by modeling spatial ER, scale-free, and small-world networks. It was found that a positively skewed spectrum is a universal property of all spatial networks. We have shown that the increase in skewness is related to the increase in number of triangles in the system. We believe that the observed peak at -1 is also due to the increase in triangles. Our results show that the eigenvalue spectrum can also be used as a tool to detect clustering in a network. One way to achieve such clustering is by spatial dependence. The spectrum asymmetry is, therefore, a tool to study the degree of spatial dependence. The spectrum sheds more information on the network structure than measures such as ED distribution or path length. For instance, in a recent work the authors have used the spectrum to determine the maximum length of an edge in a simulated polymeric gel [18].

ACKNOWLEDGMENTS

The authors gratefully acknowledge support by a grant from the NSF under Grant No. DMR0517201 and the anonymous referees for their valuable comments.

[1] J. Wang and G. Provan, Adv. Complex Syst. **12**, 45 (2009).
[2] N. Voges, A. Aertsen, and S. Rotter, Neurocomputing **70**, 1833 (2007).
[3] M. Gastner and M. Newman, Eur. Phys. J. B **49**, 247 (2006).
[4] R. Guimera, S. Mossa, A. Turtschi, and L. A. N. Amaral, Proc. Natl. Acad. Sci. U.S.A. **102**, 7794 (2005).
[5] P. Sen *et al.*, Phys. Rev. E **67**, 036106 (2003).
[6] R. Xulvi-Brunet and M. Sokolov, Phys. Rev. E **75**, 046117 (2007).
[7] C. Bauch and D. A. Rand, Proc. R. Soc. London, Ser. B **267**, 2019 (2000).
[8] I. J. Farkas, I. Derényi, A.-L. Barabási, and T. Vicsek, Phys. Rev. E **64**, 026704 (2001).
[9] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. Hwang, Phys. Rep. **424**, 175 (2006).
[10] D. Watts and S. Strogatz, Nature (London) **393**, 440 (1998).
[11] D. Cvetkovic, M. Doob, and H. Sachs, *Spectra of Graphs—Theory and Applications*, 3rd ed. (Deutscher Verlag der Wissenschaften, Berlin/Academic, New York, 1980).
[12] D. Joanes and C. Gill, Journal of the Royal Statistical Society (Series D): Statistician **47**, 183 (1998).
[13] J. Saramaki, M. Kivela, J. Onnela, K. Kaski, and J. Kertesz, Phys. Rev. E **75**, 027105 (2007).
[14] M. Barthélemy, Europhys. Lett. **63**, 915 (2003).
[15] D. W. Shattuck *et al.*, J. Neurosci. Methods **178**, 385 (2009).
[16] During the rewiring process, the formation of double edges is allowed. However, during the construction, each edge can only be made once. Doubles are formed when the original edge is not rewired.
[17] P. Holme and B. J. Kim, Phys. Rev. E **65**, 026107 (2002).
[18] J. Billen, M. Wilson, A. Baljon, and A. Rabinovitch, Europhys. Lett. **87**, 68003 (2009).