# Approximating the largest eigenvalue of network adjacency matrices

Juan G. Restrepo,[1,*] Edward Ott,[2] and Brian R. Hunt[3]
[1]*Center for Interdisciplinary Research in Complex Systems, Northeastern University, Boston, Massachusetts 02115, USA*
[2]*Department of Physics, Department of Electrical and Computer Engineering, and Institute for Research in Electronics and Applied Physics, University of Maryland, College Park, Maryland 20742, USA*
[3]*Department of Mathematics and Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742, USA*

(Received 30 May 2007; published 29 November 2007)

The largest eigenvalue of the adjacency matrix of a network plays an important role in several network processes (e.g., synchronization of oscillators, percolation on directed networks, and linear stability of equilibria of network coupled systems). In this paper we develop approximations to the largest eigenvalue of adjacency matrices and discuss the relationships between these approximations. Numerical experiments on simulated networks are used to test our results.

DOI: 10.1103/PhysRevE.76.056119 PACS number(s): 89.75.Hc, 05.45.Xt

## I. INTRODUCTION

In recent years, there has been much interest in the study of the structure of networks arising from real world systems [1]. Another concern has been dynamical processes taking place on networks, and the impact of network structure on such dynamics. The largest eigenvalue of the network adjacency matrix has emerged as a key quantity important for the study of a variety of different dynamical network processes. For example, large ensembles of heterogeneous dynamical systems can undergo a transition to synchronization as the coupling strength $k$ between the systems is increased. For a large class of networks and dynamical systems, the value of $k$ at which the transition to synchronization takes place is given by $k_c=k_0/\lambda$, where $k_0$ depends only on the dynamics of the uncoupled dynamical systems and $\lambda$ is the largest eigenvalue of the network adjacency matrix [2]. The largest eigenvalue $\lambda$ is also important in percolation on directed networks [3], linear stability of the fixed points of systems of network-coupled ordinary differential equations [4], and several other examples in physics and chemistry [5,6]. In this paper we study methods of obtaining approximations to $\lambda$ for the case of large complex networks.

We consider a network as a directed graph with $N$ nodes, and we associate to it an $N\times N$ adjacency matrix whose elements $A_{ij}$ are one if there is a directed edge from $i$ to $j$ and zero otherwise. (We require no self-edges $A_{ii}=0$ but allow bidirectional edges $A_{ij}=A_{ji}=1$.) We denote the largest eigenvalue of $A$ by $\lambda$ (assuming that the graph is connected, the eigenvalue of $A$ with the largest magnitude is unique, real, and positive by the Perron-Frobenius theorem [6]). Furthermore, we note that it is often the case that the largest eigenvalue is well separated from the second largest eigenvalue (see Fig. 1).

The properties of $\lambda$ have been studied in the context of small or regular graphs [5] and in classical Erdös-Renyi random graphs [7]. However, the structure of real world networks is usually more complex, as demonstrated by the fact that the degree distribution in a large number of examples has been found to be highly heterogeneous (often following a power law [8]), where the out-degree and in-degree of a node $i$ are defined by $d_i^{\text{out}}=\Sigma_{j=1}^{N}A_{ij}$ and $d_i^{\text{in}}=\Sigma_{j=1}^{N}A_{ji}$. The "degree distributions" $P(d^{\text{in}},d^{\text{out}})$, $P_{\text{in}}(d^{\text{in}})=\Sigma_{d^{\text{out}}}P(d^{\text{in}},d^{\text{out}})$, and $P_{\text{out}}(d^{\text{out}})=\Sigma_{d^{\text{in}}}P(d^{\text{in}},d^{\text{out}})$ are defined as the probabilities that a randomly chosen node has degree $d^{\text{in}}$ and $d^{\text{out}}$, $d^{\text{in}}$, and $d^{\text{out}}$, respectively. If $A=A^T$ we say the graph is undirected. For an undirected graph each edge serves as both an in and an out edge for each of the two nodes it joins, and for each node $i$ we have $d_i^{\text{in}}=d_i^{\text{out}}\equiv d_i$. Thus in the undirected case we write $P(d)$ to denote the corresponding degree distribution. The effect of the degree distribution on the largest eigenvalue of the adjacency matrix has been explored recently by Chung *et al.* [9], who considered a particular ensemble of random uncorrelated, undirected networks whose number of nodes $N$ is large (see also Refs. [10–15]). Here, by uncorrelated we mean that we regard the network to be a random draw from some ensemble of networks for which the joint probability distribution of the node degrees $Q(d_1,d_2,\ldots,d_N)$ factors $Q(d_1,d_2,\ldots,d_N)=P(d_1)P(d_2),\ldots,P(d_N)$. Chung *et al.* found that in the limit $N\to\infty$ these networks yield an expected largest eigenvalue that is determined by the ratio $\hat{\lambda}$ of the second to first moment of the average degree distribution

$$\hat{\lambda}=\langle d^2\rangle/\langle d\rangle, \tag{1}$$

where $\langle x\rangle=N^{-1}\Sigma_{i=1}^{N}x_i$, and by the expected largest degree $\bar{d}_{\max}$ (the maximum degree in the network averaged over

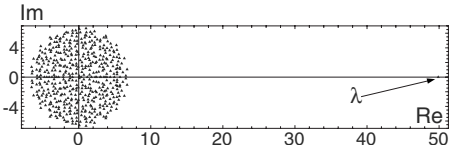


FIG. 1. All the eigenvalues of a randomly generated adjacency matrix are shown plotted in the complex plane for a case where $N=500$ and $A_{ij}=1$ with probability $p=0.1$ and $A_{ij}=0$ otherwise. We see that, aside from the largest eigenvalue $\lambda\approx 49.9$, all the other eigenvalues are contained within a disk of radius 7. (Furthermore, as $N$ increases with $p$ fixed, it is found that $\lambda$ scales as $N$, but the disk radius scales as $N^{1/2}$.)

*juanga@neu.edu

 

many network realizations). Specifically, they found that

$$\lambda \approx \begin{cases} \hat{\lambda} & \hat{\lambda} > \bar{d}_{\max} \ln N, \\ \sqrt{\bar{d}_{\max}} & \sqrt{\bar{d}_{\max}} > \hat{\lambda} \ln^2 N. \end{cases} \quad (2)$$

Some previous results for dynamical processes in networks have been stated in terms of the quantity $\hat{\lambda}$, for example, the synchronization threshold in the mean-field theory of coupled oscillators in networks [2,16,17] and the network percolation and epidemic spreading thresholds [18,19].

Real world networks often have some amount of edge degree correlations [20], i.e., a node of a given degree is more likely to be connected to nodes with certain other degrees than would be expected on the basis of chance. Networks in which high degree nodes connect preferentially to high (low) degree nodes, and vice versa, are called assortative (disassortative). Such correlations can affect dynamical processes on networks, as has been demonstrated, for example, in epidemic spreading models and percolation [21–23].

We also emphasize that the in- and out-degrees at a node can have different distributions [i.e., $P_{\text{in}}(d^{\text{in}}) \neq P_{\text{out}}(d^{\text{out}})$], as has been noted for some corporate information and genetic networks [24,25], and that there are potential correlations between the in and out degrees at the same node, which can also significantly affect the largest eigenvalue. We call these correlations node degree correlations.

The rest of this paper is organized as follows. Section II reviews the characterization of degree correlations. Section III develops the theory of the maximum eigenvalue $\lambda$ for the case of networks that satisfy a certain Markovian property. Some of the considerations of Sec. III are similar to theory in previous papers, where, however, those previous considerations were application specific, and the more general applicability to the largest eigenvalue was not apparent. Section IV tests our results on numerical constructions of network adjacency matrices of different types.

## II. DEGREE CORRELATIONS

### A. Node degree correlations

As we discussed in the Introduction, we define $P(d^{\text{in}}, d^{\text{out}})$ as the probability that a randomly chosen node has in-degree $d^{\text{in}}$ and out-degree $d^{\text{out}}$. As we shall see later, for networks without edge degree correlations, a first order approximation to the eigenvalue $\lambda$, generalizing Eq. (1), is given by

$$\hat{\lambda} = \langle d^{\text{in}} d^{\text{out}} \rangle / \langle d \rangle, \quad (3)$$

where $\langle \cdots \rangle$ denotes an average over nodes and $\langle d \rangle$ means either $\langle d_i^{\text{in}} \rangle$ or $\langle d_i^{out} \rangle$, which are equal: $\langle d_i^{\text{in}} \rangle = N^{-1} \Sigma_i d_i^{\text{in}} = N^{-1} \Sigma_{i,j} A_{ij} = \langle d_i^{\text{out}} \rangle$. If $d_i^{\text{in}}$ and $d_i^{\text{out}}$ are independent, the largest eigenvalue is approximately given by $\hat{\lambda} = \langle d \rangle$, and if $d_i^{\text{in}}$ and $d_i^{\text{out}}$ are perfectly correlated, so that $d_i^{\text{in}} = d_i^{\text{out}} \equiv d_i$ (e.g., as in an undirected network), the largest eigenvalue is approximately $\hat{\lambda} = \langle d^2 \rangle / \langle d \rangle$. We see that correlations between $d^{\text{in}}$ and $d^{\text{out}}$ can crucially affect the eigenvalue, especially if the second moment of either degree distribution diverges with increasing $N$, while the first moment converges: in such a case independence leads to a finite eigenvalue estimate, and perfect correlation to a diverging eigenvalue estimate.

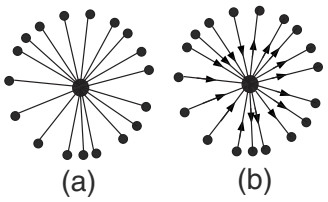


FIG. 2. Undirected (a) and directed (b) star networks illustrating the fact that, in the directed case, a node with many in and out edges does not necessarily constitute a subnetwork with a large eigenvalue (see text).

In order to quantify the correlations between $d^{\text{in}}$ and $d^{\text{out}}$ at a node, we define the node degree correlation coefficient

$$\eta \equiv \langle d^{\text{in}} d^{\text{out}} \rangle / \langle d \rangle^2. \quad (4)$$

Note that if there are no correlations $\eta = 1$ and $\eta$ is larger (smaller) than 1 for positive (negative) correlations. In terms of the correlation coefficient, $\hat{\lambda}$ is given by

$$\hat{\lambda} = \eta \langle d \rangle. \quad (5)$$

For undirected networks, $A$ is a symmetric positive matrix, and thus its largest eigenvalue satisfies

$$\lambda \geq \frac{q^T A q}{q^T q} \quad (6)$$

for any $N$ vector $q$. Choosing the components of $q$ to be zero for nodes not connected to the node of largest degree $d_{\max}$, to be one for nodes connected to the node of largest degree, and $\sqrt{d_{\max}}$ for the node of largest degree, Eq. (6) yields

$$\lambda \geq \sqrt{d_{\max}}. \quad (7)$$

If $\sqrt{d_{\max}} > \hat{\lambda}$, the mean-field approximation (3) must be incorrect. Thus, for undirected networks we use as a heuristic alternative to the mean-field approximation

$$\lambda \approx \max\{\hat{\lambda}, \sqrt{d_{\max}}\}, \quad (8)$$

which is consistent with both of the regimes considered in the rigorous result (2). (We remark, however, that Eq. (8) holds in principle only if the ratio $\hat{\lambda}/\sqrt{d_{\max}}$ or its inverse is large enough [see Eq. (2)]). One way of viewing Eq. (8) is that Eq. (6) implies that $\lambda$ is at least as large as the maximum eigenvalue of any subnetwork of the original network (by a subnetwork we mean one obtained by deleting edges of the original network). Considering a subnetwork consisting of the node of maximum degree and the nodes connecting to it, shown in Fig. 2(a), Eq. (7) corresponds to the fact that $\sqrt{d_{\max}}$ is the maximum eigenvalue of this star network. The regime $\lambda \approx \sqrt{d_{\max}}$ in Eq. (8) applies to networks whose largest eigenvalue is dominated by the node with largest degree.

We now contrast the above situation for undirected networks with what can happen for directed networks. We first note that the reasoning leading to Eq. (8) may not hold. For example, Eq. (6) no longer applies, and a node with many in

and out edges does not necessarily make a subnetwork with large eigenvalue. Regarding the latter point, consider Fig. 2(b) as compared to Fig. 2(a). This directed network has all its eigenvalues zero, because no pair of in and out edges connects the same two nodes.

### B. Edge degree correlations

For our subsequent analysis it is useful to introduce an edge degree correlation coefficient $\rho$ characterizing the correlation between the in-degree at node $i$ and the out-degree at node $j$, where a directed edge goes from $i$ to $j$,

$$\rho \equiv \langle d_i^{\text{in}} d_j^{\text{out}} \rangle_e / \langle d_i^{\text{in}} \rangle_e \langle d_j^{\text{out}} \rangle_e, \tag{9}$$

with $\langle Q_{ij} \rangle_e$ denoting an average over edges

$$\langle Q_{ij} \rangle_e \equiv \sum_{i,j} A_{ij} Q_{ij} / \sum_{i,j} A_{ij}. \tag{10}$$

Using Eq. (10) we have

$$\langle d_i^{\text{in}} \rangle_e = \sum_{i,j} A_{ij} d_i^{\text{in}} / \sum_{i,j} A_{ij} = \sum_i d_i^{\text{out}} d_i^{\text{in}} / \sum_{i,j} A_{ij} = \langle d^{\text{out}} d^{\text{in}} \rangle / \langle d \rangle \tag{11}$$

and so

$$\rho = \frac{\langle d_i^{\text{in}} d_j^{\text{out}} \rangle_e}{\eta^2 \langle d \rangle^2}. \tag{12}$$

(Our definition of the edge degree correlation coefficient $\rho$ is related to but slightly different from that of Newman [20]: we use the degrees $d_i^{\text{in}}$ and $d_i^{\text{out}}$, rather than the excess degrees $d_i^{\text{in}}-1$ and $d_i^{\text{out}}-1$. The qualitative behavior is not affected, and in most of our examples the degrees are typically large so that the difference is not important. On a directed edge from node $i$ to node $j$, we consider the product of the in-degree at node $i$ with the out-degree at node $j$, i.e., we sum $A_{ij} d_i^{\text{in}} d_j^{\text{out}}$ instead of $A_{ij} d_i^{\text{out}} d_j^{\text{in}}$. This will be convenient later and allows a natural interpretation of the numerator of $\rho$ as described below. Finally, our normalization is such that no correlations correspond to $\rho=1$ instead of zero.)

We note that the term $\Sigma_{i,j} A_{ij} d_i^{\text{in}} d_j^{\text{out}}$ that appears in the definition of $\langle d_i^{\text{in}} d_j^{\text{out}} \rangle_e$ is the number of directed paths of length three, i.e., the number of edges of the form $n \to i \to j \to m$. Similarly, $\langle d_i^{\text{in}} \rangle_e N \langle d \rangle = \langle d_j^{\text{out}} \rangle_e N \langle d \rangle = N \langle d^{\text{in}} d^{\text{out}} \rangle = \Sigma_k d_k^{\text{out}} d_k^{\text{in}} = \Sigma_{i,j,k} A_{ik} A_{kj}$ is the number of paths of length 2, $N\langle d \rangle$ is the number of paths of length 1 (or the number of edges), and $N$ is the number of paths of length 0. Accordingly, let us write $\Sigma_{i,j} A_{ij} d_i^{\text{in}} d_j^{\text{out}} \equiv n_3$, $N \langle d^{\text{in}} d^{\text{out}} \rangle \equiv n_2$, $N\langle d \rangle \equiv n_1$, and $N \equiv n_0$. With this notation, the coefficients $\eta$ and $\rho$ can be rewritten as

$$\eta = \frac{n_2 n_0}{n_1^2}, \tag{13}$$

$$\rho = \frac{n_3 n_1}{n_2^2}. \tag{14}$$

As an example of this interpretation, we note that for networks with uncorrelated in- and out-degrees, the number of paths of length two is $n_2 \approx n_1 \langle d \rangle = n_1^2 / n_0$, and so $\eta \approx 1$. For networks where there are no edge degree correlations, the number of paths of length 3, $n_3$, can be obtained from the number of paths of length 2, $n_2$, times the average branching ratio given by $n_2/n_1$, and thus $\rho \approx 1$ for such networks.

## III. LARGEST EIGENVALUE OF MARKOVIAN NETWORKS

### A. Formulation

For generality, in this section we allow different types of network nodes, where we specify the node type by an index $\sigma = 1, 2, \ldots, M$, where $M$ is the total number of possible node types (e.g., in the case of social networks connecting people, $\sigma$ might label sex, race, social class, etc.). Furthermore, we introduce the quantity

$$z = (d^{\text{in}}, d^{\text{out}}, \sigma), \tag{15}$$

which we refer to as the degree. With this definition, we use $P(z)$ to denote the degree distribution, i.e., $P(z)$ is the probability that a randomly chosen node has degree $z$. This implies, for example, that

$$\sum_{d^{\text{in}}, d^{\text{out}}} P(z) = N_\sigma / N, \tag{16}$$

where $N_\sigma$ is the number of nodes of type $\sigma$ and $N = \Sigma_\sigma N_\sigma$. (While our numerical examples in Sec. IV are for the case of a single node type $M=1$ the subsequent considerations in the present section do not have this restriction.)

We consider a particular class of networks for which the only nontrivial correlations are between nodes that are directly connected by a single edge. Such Markovian networks have been considered in previous works on epidemic spreading and percolation [18,19,21,23]. Under this assumption, if we define $P(z'|z)$ to be the probability that a randomly chosen node with degree $z$ points to a node with degree $z'$, then if we choose a random outward path of length 2 from a node with degree $z$, the probability that the first hop ends on a node with degree $z'$ and the second on a node with degree $z''$ is by this assumption

$$P(z', z''|z) = P(z''|z') P(z'|z).$$

Let $\psi_z^{(m)}$ be the expected number of directed paths of length $m$ whose starting node has degree $z$. Using the assumption that the network is Markovian, we can express $\psi_z^{(m+1)}$ in terms of $\psi_z^{(m)}$ as

$$\psi_z^{(m+1)} = d^{\text{out}} \sum_{z'} P(z'|z) \psi_{z'}^{(m)}. \tag{17}$$

The number of paths of length $m$ grows, in the limit of large $m$, as the largest eigenvalue $\lambda$, $\psi_z^{(m)} \sim \psi_z \lambda^m$ [5]. Therefore, we can associate to Eq. (17) the eigenvalue problem

$$\lambda_C \psi_z = d^{\text{out}} \sum_{z'} P(z'|z) \psi_{z'} \tag{18}$$

or

$$\lambda_C \psi_z = C\psi_z, \tag{19}$$

where $C$ is the $d^{\text{out}}_{\max} d^{\text{in}}_{\max} \times d^{\text{out}}_{\max} d^{\text{in}}_{\max}$ matrix with entries $C_{zz'} = d^{\text{out}} P(z'|z)$, and we seek its largest eigenvalue $\lambda_C$ and its corresponding eigenfunction $\psi_z$. The eigenfunction $\psi_z$ determines how the number of directed paths starting at a node with degree $z$ depends on $z$. As we shall see, when there are no edge degree correlations ($\rho=1$), $\psi_z$ is proportional to the out-degree $\psi_z \propto d^{\text{out}}$. We consider $\lambda_C$ to be an approximation to $\lambda$ that is more accurate than the mean-field result in that it includes correlations between connected nodes $\lambda \approx \lambda_C$. Additionally, an approximation $u_C$ to the right eigenvector $u$ of $A$ can be obtained in terms of $\psi_z$ in a similar way

$$(u_C)_i = \psi_{z_i} \tag{20}$$

and an analogous expression holds for the left eigenvector $v$ of $A$, using the left eigenvector of the matrix with entries $d^{\text{out}} P(z'|z)$ instead of $\psi_z$. We will refer to Eqs. (18) and (20) as the Markovian approximation.

In the Markovian approximation, the matrix $C$ captures the properties of the matrix $A$. This correspondence arises because the largest eigenvalue $\lambda$ is determined only by the rate of growth of the number of directed paths. Under the Markovian approximation, the average number of directed paths depends only on the way nodes with different degrees $z$, $z'$ are connected to each other. This information is contained in the entries of $C$, $C_{zz'} = d^{\text{out}} P(z'|z)$, the expected number of edges pointing from a node of degree $z$ to a node of degree $z'$. As shown in our analysis, the information contained in the matrix $C$ is enough to determine $\lambda$.

### B. $\epsilon$ expansion

While in many cases Eq. (18) can be solved directly, it is also of interest to explore approximations to its solution. Thus, we will here expand Eq. (18) about a zeroth order approximation in which edge degree correlations are neglected and in which the first order correction gives the perturbation to the uncorrelated case due to a small amount of edge degree correlations. In particular, we will be interested in zeroth order and first order approximations to the largest eigenvalue of $A$ and to the corresponding right and left eigenvectors $u$ and $v$, where

$$Au = \lambda u \quad \text{and} \quad v^T A = \lambda v^T. \tag{21}$$

Following previously used terminology, we refer to the zeroth order approximation as the mean-field theory.

When there are no edge degree correlations, the probability that an edge outgoing from a node with degree $z$ points to a node with degree $z'$, $P(z'|z)$ becomes independent of $z$. This probability is proportional to the fraction of edges incoming into nodes of degree $z'$:

$$P(z'|z) = \hat{P}(z') = (d^{\text{in}})' P(z')/\langle d \rangle. \tag{22}$$

Notice that $\hat{P}(z)$ is the probability that a randomly chosen link points to a node of degree $z$. We will expand Eq. (18) about the uncorrelated case. For this purpose, we will write

$$P(z'|z) = \hat{P}(z') + \epsilon \delta P(z'|z), \tag{23}$$

where

$$\delta P(z'|z) = P(z'|z) - \hat{P}(z') \tag{24}$$

and $\epsilon$ is an expansion parameter that we formally consider small, although, in reality, $\epsilon=1$. Introducing expansions for $\lambda_C$ and $\psi_z$,

$$\lambda_C = \hat{\lambda} + \epsilon \delta\lambda + \cdots, \tag{25}$$

$$\psi_z = \hat{\psi}_z + \epsilon \delta\psi_z + \cdots. \tag{26}$$

Expanding Eq. (18) to zero order in $\epsilon$, we obtain

$$\hat{\psi}_z = d^{\text{out}} \tag{27}$$

and

$$\hat{\lambda} = \langle d^{\text{in}} d^{\text{out}} \rangle / \langle d \rangle. \tag{28}$$

Thus in the zeroth order approximation, the right eigenvector of $A$ has components

$$u_i = d^{\text{out}}_i. \tag{29}$$

To obtain the left eigenvector, we follow the same steps as above but with $A$ replaced by $A^T$. This interchanges the roles of $d^{\text{in}}$ and $d^{\text{out}}$, thus yielding

$$v_i = d^{\text{in}}_i \tag{30}$$

for the left eigenvector.

Expanding to first order in $\epsilon$, we obtain

$$\hat{\lambda}\delta\psi_z + \delta\lambda\hat{\psi}_z = d^{\text{out}} \sum_{z'} \hat{P}(z')\delta\psi_{z'} + d^{\text{out}} \sum_{z'} \delta P(z'|z)\hat{\psi}_{z'}. \tag{31}$$

Multiplying by $\hat{P}(z)$ and summing over $z$ we obtain, after some simplification,

$$\delta\lambda = [(\hat{\lambda}\langle d \rangle)^{-1} \sum_{z,z'} (d^{\text{out}})' d^{\text{out}} d^{\text{in}} P(z'|z) P(z)] - \hat{\lambda}. \tag{32}$$

The probability that a randomly chosen edge starts at a node with degree $z$ is $\tilde{P}(z) = \langle d \rangle^{-1} P(z) d^{\text{out}}$ and, therefore, the term $\langle d \rangle^{-1} \Sigma_{z,z'} (d^{\text{out}})' d^{\text{out}} d^{\text{in}} P(z'|z) P(z)$ is equal to

$$\sum_{z,z'} d^{\text{in}} (d^{\text{out}})' \tilde{P}(z) P(z'|z) = \langle d^{\text{in}}_i d^{\text{out}}_j \rangle_e. \tag{33}$$

To first order, therefore, we obtain

$$\lambda_C \approx \hat{\lambda} + \delta\lambda = \frac{\langle d^{\text{in}}_i d^{\text{out}}_j \rangle_e}{\hat{\lambda}} = \hat{\lambda}\rho, \tag{34}$$

where $\rho$ is defined in Sec. II, and we call $\lambda = \hat{\lambda}\rho$ the linear approximation.

We note that the successively more refined approximations $\langle d \rangle$, $\hat{\lambda}$ [Eq. (5)], and the linear approximation [Eq. (34)] correspond, respectively, to $n_1/n_0$, $n_2/n_1$, and $n_3/n_2$, and $\lambda$

$=\lim_{k\to\infty} n_{k+1}/n_k$. We now briefly comment on the range of validity of the expansion results. As the rigorous result Eq. (2) shows, the mean-field result for undirected networks is not valid when the network is dominated by the node with maximum degree. A network in which the maximum degree is too large compared with the bulk of the degree distribution cannot satisfy Eq. (22). For example, the star network of Fig. 2(a) is very degree-degree correlated because all of its outer nodes connect only to the high degree hub node and, indeed, application of Eq. (22) to this simple star network yields an incorrect result. For this network, we have $P(d'|d) = \delta_{d',d_{\max}}\delta_{d,1} + \delta_{d',1}\delta_{d,d_{\max}}$, where $d_{\max}$ is the degree of the central node. This yields $\lambda_C = \sqrt{d_{\max}}$. However, the approximation (22) yields $\hat{P}(d') = 1/2$ for both $d' = 1$ and $d' = d_{\max}$, which gives incorrect values for $\langle d\rangle$, $\langle d^2\rangle$, and $\lambda$. Our expansions about an uncorrelated network therefore implicitly assumed that this network was not dominated by the node with maximum degree. A more rigorous delineation of the range of validity of the Markovian approximation [Eq. (18)] for correlated and/or directed networks along the lines of Ref. [9]. is open for further research.

## IV. NUMERICAL TESTS ON SIMULATED NETWORKS

In this section we numerically construct random networks and use them to compare with the predictions of Sec. III. In order to study the variation of the largest eigenvalue $\lambda$ of the adjacency matrix $A$, we first construct large ($N \gg 1$) approximately uncorrelated networks with a given expected degree distribution using a generalization of the random graph model of Chung *et al.* [9]. First we associate to each node $k$ target in- and out-degrees $(\hat{d}_k^{\text{in}}, \hat{d}_k^{\text{out}})$, where $\Sigma_k \hat{d}_k^{\text{in}} = \Sigma_k \hat{d}_k^{\text{out}} = N\langle\hat{d}\rangle$. Note that the choice of target node degrees can be made to correspond to a desired degree distribution (e.g., scale free or Poisson). We then choose each element of the adjacency matrix $A_{ij}$ randomly to be one with probability $\hat{d}_i^{\text{out}}\hat{d}_j^{\text{in}}/(N\langle\hat{d}\rangle)$ and zero otherwise [we restrict the maximum degrees so that $\hat{d}_i^{\text{out}}\hat{d}_j^{\text{in}}/(N\langle\hat{d}\rangle) \le 1$]. This determines a network realization with in- and out-degrees $d_i^{\text{out}} = \Sigma_j A_{ij}$ and $d_j^{\text{in}} = \Sigma_i A_{ij}$. In general, $(d_k^{\text{in}}, d_k^{\text{out}})$ can be different from the target values $(\hat{d}_k^{\text{in}}, \hat{d}_k^{\text{out}})$. Nevertheless, with high probability, for large $N$, the resulting degree distribution $P(d^{\text{in}}, d^{\text{out}})$ of this randomly chosen network will be approximately, in a suitable sense, the target distribution [9]. In particular, the moments $\langle d^{\text{in}} d^{\text{out}}\rangle$ and $\langle d\rangle$ will be approximately unchanged when calculated using either $P$ or the target degrees. Furthermore, $\rho \approx 1$ for $N \gg 1$. The network generated by this algorithm is directed (i.e., $A$ is asymmetric); symmetric networks can be generated by first considering only $i < j$ and then setting $A_{ji} = A_{ij}$.

Starting from an uncorrelated network generated by this algorithm, we then rewire the connections in such a way that the degree distribution is preserved. On doing so, the correlation coefficient changes and we calculate the largest eigenvalue for different values of $\rho$. The rewiring algorithm we use is a simplified version of that used in Ref. [22] and consists in the iteration of the following steps. (1) Two edges

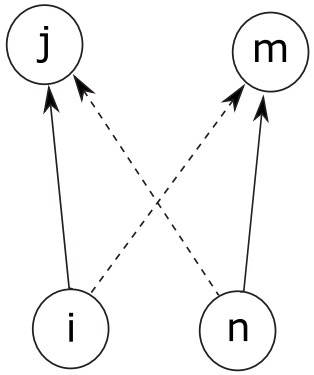


FIG. 3. Schematic representation of the rewiring algorithm. Edges $i \to j$ and $n \to m$ were chosen at random. When creating a network with assortative (disassortative) correlations, they are replaced with edges $i \to m$ and $n \to j$ if the resulting network has a larger (smaller) value of $\rho$.

are chosen at random. Assume one connects node $n$ to node $m$ and the other connects node $i$ to node $j$. (2) Let

$$H(i,j;n,m) = d_n^{\text{in}} d_m^{\text{out}} + d_i^{\text{in}} d_j^{\text{out}} - d_n^{\text{in}} d_j^{\text{out}} - d_i^{\text{in}} d_m^{\text{out}}. \quad (35)$$

The two edges chosen in step 1 are replaced with two edges connecting node $n$ to node $j$ and node $i$ to node $m$ (see Fig. 3) if $sH(i,j;n,m) < 0$, and are left alone otherwise. Setting $s = 1$ or $-1$ we produce assortative or disassortative networks, respectively.

As we iterate from $\rho = 1$ with $s = 1$ ($s = -1$), $\rho$ steadily increases (decreases). Thus, we produce a sequence of networks with successively larger (smaller) $\rho$.

In Fig. 4(a) we show the largest eigenvalue $\lambda$ (solid line), the linear approximation given by Eq. (34) (dashed line), and the Markovian approximation $\lambda_C$ obtained by exact diagonalization of the matrix $C = d^{\text{out}} P(z'|z)$ in Eq. (18) (boxes) as a function of the correlation coefficient $\rho$ for an undirected

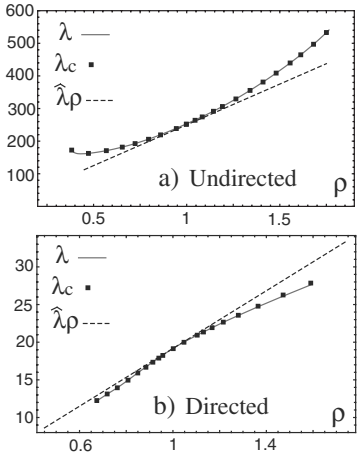


FIG. 4. Largest eigenvalue $\lambda$ (solid line), linear approximation [Eq. (34), dashed line], and Markovian approximation $\lambda_C$ [Eq. (18), boxes] for (a) an undirected network with $N = 25\,000$, $\gamma = 2.5$, and $\langle d\rangle = 100$ and (b) a directed network with $N = 10\,000$, $\gamma = 2.5$, and $\langle d\rangle = 20$.

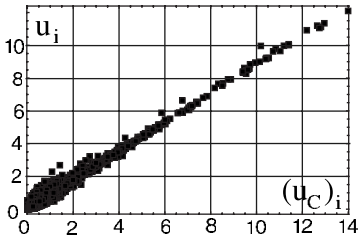


FIG. 5. Entries of the eigenvector $u$ corresponding to $\lambda$ versus the entries of its Markovian approximation $u_C$ [Eq. (20)] in an arbitrary scale for the network of Fig. 4(b) at $\rho \approx 0.9$.

network with power law degree distribution generated as described above with $N=25\,000$, $\langle d \rangle = 100$, and exponent $\gamma = 2.5$. In Fig. 4(b) we plot the same quantities for a directed network with $N=10\,000$, $\langle d \rangle = 20$, and exponent $\gamma = 2.5$. When constructing the directed network, we chose the target $\hat{d}_k^{\text{in}}$ independently from $\hat{d}_k^{\text{out}}$ so that there are no node degree correlations ($\eta = 1$). In these plots there is no discernible difference between the approximation $\lambda_C$ and the actual values of $\lambda$. We observe that the largest eigenvalue depends strongly on the correlation coefficient: in the undirected case, it increases more than three times as $\rho$ varies from 0.4 to 1.7. Also, we see that in these examples the linear approximation works for $|\rho - 1| \lesssim 0.2$, but fails for larger values of $|\rho - 1|$. In the undirected case, $\lambda$ is larger than the linear approximation, which follows from Eq. (6) if we set $q_i = d_i$. We also note that in the undirected network there are strong node degree correlations ($\eta = \hat{\lambda}/\langle d \rangle \sim 2.5$), but this does not affect the quality of the approximations.

In Fig. 5 we show the eigenvector $u_i$ for the network of Fig. 4(b) at $\rho \approx 0.9$ plotted against the corresponding approximation $(u_C)_i$, Eq. (20), using an arbitrary scale. Up to a normalization factor, there is good agreement between the true value and its Markov estimate [$(u_C)_i \propto u_i$].

## V. CONCLUSION

In this paper we have considered several approximations to the largest eigenvalue of the adjacency matrix of large, directed networks. The mean-field result (3) appears to apply well to networks whose neighboring nodes are uncorrelated in their degrees. The linear approximation (34) applies for sufficiently small correlation, while the Markov model (18) applies for arbitrarily strong degree correlations between neighbors. The price to be paid for a more refined approximation is the requirement of greater knowledge of the network [e.g., use of Eq. (18) requires knowledge of $P(z'|z)$ which is not required for the two other less refined approximations].

We caution that, although we have obtained good agreement between the theory and numerical results on simulated networks, this may not necessarily carry through for real networks encountered in practice. In particular, the Markov assumption of Eq. (17) may not always hold (e.g., due to community structure [26], clustering, or edge degree correlations extending over more than one edge between nodes). This remains a topic for further study.

## ACKNOWLEDGMENTS

This work was supported by ONR (Physics), by the NSF (Grants No. PHY 0456240 and No. DMS 0104-087), and by USAFOSR Grant No. FA95500410319.

[1] M. E. J. Newman, SIAM Rev. **45**, 167 (2003); R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002); S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D-U. Hwang, Phys. Rep. **424**, 175 (2006).
[2] J. G. Restrepo, E. Ott, and B. R. Hunt, Phys. Rev. E **71**, 036151 (2005); Chaos **16**, 015107 (2006); Phys. Rev. Lett. **96**, 254103 (2006).
[3] J. G. Restrepo, E. Ott, and B. R. Hunt, e-print arXiv:0704.0491v1; B. Bollobas, C. Borgs, J. Chayes, and O. Riordan, e-print arXiv:math/0701346v1.
[4] R. May, Nature (London) **238**, 413 (1972); J. Feng, V. K. Jirsa, and M. Ding, Chaos **16**, 015109 (2006); M. Brede and S. Sinha, e-print arXiv:cond-mat/0507710.
[5] D. Cvetkovic and P. Rowlinson, Linear Multilinear Algebra **28**, 3 (1990).
[6] C. R. MacCluer, SIAM Rev. **42**, 487 (2000).
[7] M. Krivelevich and B. Sudakov, Combin. Probab. Comput. **12**, 61 (2003).
[8] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[9] F. Chung, L. Lu, and V. Vu, Proc. Natl. Acad. Sci. U.S.A. **100**, 6313 (2003).
[10] I. J. Farkas, I. Derényi, A.-L. Barabási, and T. Vicsek, Phys. Rev. E **64**, 026704 (2001).
[11] K.-I. Goh, B. Kahng, and D. Kim, Phys. Rev. E **64**, 051903 (2001).
[12] S. N. Dorogovtsev, A. V. Goltsev, J. F. F. Mendes, and A. N. Samukhin, Phys. Rev. E **68**, 046109 (2003).
[13] M. A. M. de Aguiar and Y. Bar-Yam, Phys. Rev. E **71**, 016106 (2005).
[14] K. Iguchi and H. Yamada, Phys. Rev. E **71**, 037102 (2005).
[15] H. Yang, F. Zhao, and B. Wang, Chaos **16**, 043112 (2006).
[16] T. Ichinomiya, Phys. Rev. E **70**, 026116 (2004).
[17] D.-S. Lee, Phys. Rev. E **72**, 026208 (2005).
[18] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin, Phys. Rev. Lett. **85**, 4626 (2000).
[19] Y. Moreno, R. Pastor-Satorras, and A. Vespignani, Eur. Phys. J. B **26**, 521 (2002).
[20] M. E. J. Newman, Phys. Rev. E **67**, 026126 (2003).
[21] A. Vázquez and Y. Moreno, Phys. Rev. E **67**, 015101(R) (2003).
[22] M. E. J. Newman, Phys. Rev. Lett. **89**, 208701 (2002).
[23] M. Boguñá and R. Pastor-Satorras, Phys. Rev. E **66**, 047104 (2002).
[24] D. Braha and Y. Bar-Yam, Phys. Rev. E **69**, 016113 (2004).
[25] N. Guelzim, S. Bottani, P. Bourgine, and François Képès, Nat. Genet. **31** 60 (2002).
[26] M. E. J. Newman and M. Girvan, Phys. Rev. E **69**, 026113 (2004).