Nonbacktracking expansion of finite graphs

G. Timár, R. A. da Costa, S. N. Dorogovtsev, and J. F. F. Mendes

1Departamento de Física da Universidade de Aveiro & I3N, Campus Universitário de Santiago, 3810-193 Aveiro, Portugal
2A. F. Ioffe Physico-Technical Institute, 194021 St. Petersburg, Russia

Message passing equations yield a sharp percolation transition in finite graphs, as an artifact of the locally treelike approximation. For an arbitrary finite, connected, undirected graph we construct an infinite tree having the same local structural properties as this finite graph, when observed by a nonbacktracking walker. Formally excluding the boundary, this infinite tree is a generalization of the Bethe lattice. We indicate an infinite, locally treelike, random network whose local structure is exactly given by this infinite tree. Message passing equations for various cooperative models on this construction are the same as for the original finite graph, but here they provide the exact solutions of the corresponding cooperative problems. These solutions are good approximations to observables for the models on the original graph when it is sufficiently large and not strongly correlated. We show how to express these solutions in the critical region in terms of the principal eigenvector components of the nonbacktracking matrix. As representative examples we formulate the problems of the random and optimal destruction of a connected graph in terms of our construction, the nonbacktracking expansion. We analyze the limitations and the accuracy of the message passing algorithms for different classes of networks and compare the complexity of the message passing calculations to that of direct numerical simulations. Notably, in a range of important cases, simulations turn out to be more efficient computationally than the message passing.

I. INTRODUCTION

The Bethe lattice is a valuable substrate, on top of which many cooperative models are solved exactly and show critical behaviors with mean-field critical exponents and Gaussian critical fluctuations. A regular Bethe lattice is equivalent to an infinite random regular graph, which is locally treelike and has only infinite cycles (loops) and no boundary. This is in contrast with an infinite Cayley tree in which the boundary suppresses phase transitions, e.g., in the Ising model on a Cayley tree, finite or infinite, long-range order never emerges. The configuration model of a sparse uncorrelated network and its generalizations represent a heterogeneous version of the Bethe lattice. This basic model also provides exact solutions for percolation and other problems in infinite random networks.

Self-consistency equations for the probability of reaching a finite branch by following a link yield the solution of percolation problems for infinite locally treelike random networks. Recently it was found in Ref. that such self-consistency equations produce a good approximation of the percolation behavior of large but still finite, sparse graphs. This conclusion was supported by numerical simulations. The equations are written for a given graph from which a fraction \( 1 - p \) of links is removed. These self-consistency equations—termed message passing equations in the context of finite graphs—predict a sharp continuous phase transition for any finite graph, e.g., for the 4-clique (four interconnected nodes). In reality, a percolation continuous phase transition cannot be defined in finite systems. The same is true for other models on finite graphs. In this paper, instead of describing a given finite system with some unknown accuracy, we address the issue of this useful approximation from a complementary perspective. Namely, if the message passing equations in principle cannot precisely describe physical models on finite graphs, then what do they describe? For each given finite graph, we construct a related infinite network for which these equations provide an exact description of the cooperative models. This infinite network is a generalization of the Bethe lattice with percolation properties that are described exactly by the message passing equations on the finite graph. The closer the original finite graph is to this construction, the better the approximation provided by the message passing equations. We therefore give a physical interpretation of the solutions of these equations for percolation on a finite graph. We express these solutions in the critical region in terms of the principal eigenvector components of the nonbacktracking matrix. As another example we consider the problem of optimal percolation. We show that using our network construction, an approximate algorithm for finding the so-called decycling number can be formulated in simple terms of the principal eigenvector of the nonbacktracking matrix. We introduce a way to quantify the accuracy of the message passing algorithm and indicate two classes of networks for which the approximation fails. We calculate the time complexity associated with message passing algorithms and show that in a wide range of important situations this method is less efficient than direct numerical simulations.

II. CONSTRUCTION

To build the nonbacktracking expansion (NBE) of an arbitrary finite, connected, undirected graph \( G \), we first construct the tree \( T(\ell, i) \) in the following way. For any given node \( i \) of \( G \), \( T(1, i) \) is a star graph consisting of a node labeled \( i \) in the center, and for any node \( j \) in \( G \).
In Fig. 1(a), the construction $T$ defined as $B$ in $T$ in $k$ of graph $G$ we see that the branches emanating from the root both directions must be taken into account for each link.

FIG. 1. Construction of (b) the tree $T(3, 1)$ and (c) the branch $B(2, 2 \leftarrow 1)$ for an example graph (a). The root (starting node) in $T(3, 1)$ is shown by the open circle and the starting link in $B(2, 2 \leftarrow 1)$ is represented by the dotted line.

A more explicit definition of the NBE can be given using the “$m$-cloned” network [12] of the given graph $G$. This network is constructed as follows. Let the nodes of $G$ be labeled $i = 1, 2, \ldots, N$. We make $m$ replicas of all these nodes, still labeled $i = 1, 2, \ldots, N$ within each of the $m$ replicas. The $mN$ nodes obtained in this way are the nodes of the $m$-cloned network $N(m)$ defined as follows. Let $N(m)$ be the network that is maximally random under the constraint that if and only if nodes $i$ and $j$ are connected in $G$, then any of the $m$ nodes labeled $i$ in $N(m)$ is connected to exactly one node labeled $j$ (of, in total, $m$ such nodes). In other words, make $m$ copies of graph $G$ and consider all possible rewirings that preserve the labels of end nodes of all links, with equal statistical weights (uniform randomness). Figure 2 explains this construction for $m = 2$. Let us take the limit $m \to \infty$ and consider the infinite network $N(m \to \infty)$. This network is locally treelike: any finite neighborhood of any node labeled $i$ is given by the tree $T(\ell, i)$. The relative number of finite loops in this infinite network of any finite neighborhood of each node of label $i$ (and of each link of label $j \to i$) is the same. This construction is a generalization of the Bethe lattice and we call it the nonbacktracking expansion of graph $G$. It is a specific Bethe lattice preserving the local topology of connections of the graph $G$. For example, the NBE of the 4-clique is simply a regular Bethe lattice of coordination number 3.

for which a link $j \to i$ exists, a node labeled $j$ is attached to the center node of $T(1, i)$. For general $\ell > 1$ we define $T(\ell, i)$ in a recursive way. To any leaf labeled $k$ in $T(\ell - 1, i)$, we attach a node labeled $l$ if and only if $B_{l \to k, k \to j} = 1$, where $B$ is the nonbacktracking matrix of graph $G$ and $j$ is the label of the parent node of leaf $k$ in $T(\ell - 1, i)$. The nonbacktracking matrix $B$ is defined as $B_{i \to j, k \to l} = \delta_{jk}(1 - \delta_{il})$, where $i, j$ and $k, l$ are end node pairs of links in the original graph. Note that both directions must be taken into account for each link. We see that the branches emanating from the root $i$ in $T(\ell, i)$ correspond exactly to nonbacktracking walks of, at most, length $\ell$, in the original graph $G$ starting from node $i$. (Note that some of these walks may be shorter than $\ell$, when a dead end in graph $G$ is encountered.) In this construction the nodes of the resulting tree have only labels of the original finite graph, $i = 1, 2, \ldots, N$, so different nodes in this tree can have the same labels. The same is true for nodes in each individual branch of the tree. Figure 1(b) explains $T(\ell, i)$ for the finite graph in Fig. 1(a). The construction $T(\ell, i)$ is also known as a computation tree in the computer science community [11]. Taking the limit $\ell \to \infty$ and formally excluding the boundary, we obtain a network in which all nodes of the same label $i$ and all links of the same label $j \to i$ are topologically equivalent, in the sense that the structure of any finite neighborhood of each node of label $i$ (and of each link of label $j \to i$) is the same. This construction is a generalization of the Bethe lattice and we call it the nonbacktracking expansion of graph $G$. It is a specific Bethe lattice preserving the local topology of connections of the graph $G$. For example, the NBE of the 4-clique is simply a regular Bethe lattice of coordination number 3.

FIG. 2. Construction of network $N(2)$ for the finite graph in Fig. 1(a), according to Ref. [12]. Three of the members of the statistical ensemble of graphs constituting $N(2)$ are shown as examples.
is negligible, almost all loops are infinite, and almost all nodes of label \(i\) (and links of label \(j \rightarrow i\)) are topologically equivalent. The network \(\mathcal{N}(m \rightarrow \infty)\) gives exactly the NBE of graph \(\mathcal{G}\).

To discuss the properties of the NBE it is helpful to define the branch \(B(\ell, j \leftarrow i)\) as the tree that is recursively generated by the nonbacktracking walks of length \(\ell\) in \(\mathcal{G}\) starting from the link \(j \leftarrow i\), as shown in Fig. 1(c). Clearly, \(T(\ell + 1, i)\) consists of all branches \(B(\ell, j \leftarrow i)\) (all \(j\), for which there is a link \(j \leftarrow i\) in graph \(\mathcal{G}\)).

III. PROPERTIES OF THE NONBACKTRACKING EXPANSION

Let us consider properties of these constructions. Let \(n_{l \leftarrow k}(\ell, j \leftarrow i)\) be the number of links of label \(l \leftarrow k\) on the surface of \(B(\ell, j \leftarrow i)\) (these are links from parent \(k\) on the surface of sphere \(\ell - 1\) to child \(l\) on the surface of sphere \(\ell\)). Let \(n(\ell, j \leftarrow i)\) be the size (number of nodes) of the surface of \(B(\ell, j \leftarrow i)\), so that \(f_{l \leftarrow k}(\ell, j \leftarrow i) \equiv n_{l \leftarrow k}(\ell, j \leftarrow i)/n(\ell, j \leftarrow i)\) is the relative number of links of label \(l \leftarrow k\) on the surface of \(B(\ell, j \leftarrow i)\). Every link on the surface of \(B(\ell + 1, j \leftarrow i)\) must emanate from the end node of a link on the surface of \(B(\ell, j \leftarrow i)\), so

\[
n_{l \leftarrow k}(\ell + 1, j \leftarrow i) = \sum_{l' \leftarrow k'} n_{l' \leftarrow k'}(\ell, j \leftarrow i)B_{l \leftarrow k', l' \leftarrow k'}, \tag{1}
\]

that is,

\[
n(\ell + 1, j \leftarrow i)f_{l \leftarrow k}(\ell + 1, j \leftarrow i) = n(\ell, j \leftarrow i) \sum_{l' \leftarrow k'} f_{l' \leftarrow k'}(\ell, j \leftarrow i)B_{l \leftarrow k', l' \leftarrow k'}. \tag{2}
\]

Taking \(\ell \to \infty\), the ratio \(n(\ell + 1, j \leftarrow i)/n(\ell, j \leftarrow i)\) converges to an asymptotic mean branching \(\overline{b}\) and the numbers \(f_{l \leftarrow k}(\ell, j \leftarrow i)\) must also converge to some asymptotic values \(f_{l \leftarrow k}^*(j \leftarrow i)\). For \(\ell \to \infty\), therefore, we get

\[
\overline{b}f_{l \leftarrow k}^*(j \leftarrow i) = \sum_{l' \leftarrow k'} f_{l' \leftarrow k'}^*(j \leftarrow i)B_{l \leftarrow k', l' \leftarrow k'} \tag{3}
\]

or, in vector form,

\[
\overline{b}f^*(j \leftarrow i) = Bf^*(j \leftarrow i). \tag{4}
\]

Equation (4) is a right eigenvector equation for the nonbacktracking matrix of the original graph \(\mathcal{G}\). According to the Perron-Frobenius theorem [12], the nonbacktracking matrix \(B\) has exactly one strictly non-negative right eigenvector, and this is the eigenvector corresponding to the leading eigenvalue. All of the components of \(\overline{b}f^*(j \leftarrow i)\) are, by definition, non-negative, therefore the mean branching \(\overline{b}\) is equal to the leading eigenvalue \(\lambda_1\) of \(B\). We note that the leading eigenvalue and the corresponding right eigenvector are independent of the starting link \(j \leftarrow i\), so we have the following general result. The asymptotic mean branching of \(B(\ell, j \leftarrow i)\), for any \(j \leftarrow i\), is given by the leading eigenvalue \(\lambda_1\) of the nonbacktracking matrix of graph \(\mathcal{G}\). From the construction of \(T(\ell, i)\), it is clear that \(T(\ell, i)\) must also have the same asymptotic mean branching for any root \(i\). The asymptotic relative numbers of links \(l \leftarrow k\) on the surface of \(B(\ell, j \leftarrow i)\) or \(T(\ell, i)\) are given by the components of the principal right eigenvector \(v^{(\lambda_1)}\) of \(B\). These asymptotic relative numbers are, again, independent of the starting link (or node).

Now we can find the asymptotic relative number \(g_j\) of nodes of label \(i\) on the surface of \(B(\ell, r^\prime \leftarrow r)\) or \(T(\ell, r)\), independent of the starting link \(r^\prime \leftarrow r\) or root \(r^\prime\):

\[
g_j = \frac{\sum_{j \in \partial i} v^{(\lambda_1)}(j, i)}{\sum_{k=1}^N \sum_{j \in \partial k} v^{(\lambda_1)}(k, j)}, \tag{5}
\]

where \(\partial i\) denotes the set of nearest neighbors of node \(i\) in graph \(\mathcal{G}\). The asymptotic relative number of nodes \(i\) must also be the same on the layer just beneath the surface, and each node in this layer (apart from dead ends) must be the starting node of a link on the surface, so we can also write

\[
g_j = \frac{\frac{1}{q_i} \sum_{j \in \partial i} v^{(\lambda_1)}(j, i)}{\sum_{k=1}^N \left(\frac{1}{q_k} \sum_{j \in \partial k} v^{(\lambda_1)}(i, j)\right)}, \tag{6}
\]

where \(q_i > 1\) is the degree of nodes \(i\). Equations (5) and (6) give a general relationship between the components of the eigenvector \(v^{(\lambda_1)}\). Note that \(g_j\) is the asymptotic probability of finding a nonbacktracking random walker at node \(i\) after an infinitely long walk. This quantity coincides with the nonbacktracking centrality of \(i\).

Let us return to \(n(\ell, j \leftarrow i)\), the size of the surface of \(B(\ell, j \leftarrow i)\). Let \(n(\ell) \equiv \sum_{j \rightarrow i} n(\ell, j \leftarrow i)\) be the sum of the sizes of all such surfaces, and \(h_{j \rightarrow i}(\ell) \equiv n(\ell, j \leftarrow i)/n(\ell)\) be the relative size of the surface of \(B(\ell, j \leftarrow i)\). Then

\[
n(\ell + 1, j \leftarrow i) = \sum_{l \leftarrow k} n(\ell, l \leftarrow k)B_{l \leftarrow k, j \leftarrow i} \tag{7}
\]

or

\[
n(\ell + 1)h_{j \leftarrow i}(\ell + 1) = n(\ell) \sum_{l \leftarrow k} h_{l \leftarrow k}(\ell)B_{l \leftarrow k, j \leftarrow i}. \tag{8}
\]

In the limit \(\ell \to \infty\), the ratio \(n(\ell + 1)/n(\ell)\) converges to the asymptotic mean branching \(\overline{b}\) and the numbers \(h_{j \rightarrow i}(\ell)\) must also converge to some asymptotic values \(h_{j \rightarrow i}^*\). We can write

\[
\overline{b}h_{j \rightarrow i}^* = \sum_{l \leftarrow k} h_{l \leftarrow k}^*B_{l \leftarrow k, j \leftarrow i} \tag{9}
\]

or, in vector form,

\[
\overline{b}h^* = h^*B, \tag{10}
\]

which is just a left eigenvector equation for \(B\). The same reasoning applies as before, so \(\overline{b}\) is equal to the leading eigenvalue of \(B\), \(\lambda_1\). The vector \(h^*\) is equal to the corresponding left eigenvector \(w^{(\lambda_1)}\). Then the asymptotic
relative sizes of the surfaces of $B(\ell, j \leftarrow i)$ are given by the components of $w^{(\lambda_1)}$. We express the asymptotic relative sizes $s_i$ of the surfaces of $T(\ell, i)$ in terms of these components:

$$s_i = \frac{\sum_{j \in \partial_i} w_{i \leftarrow j}^{(\lambda_1)}}{\sum_{k=1}^{N} \sum_{j \in \partial_k} w_{j \rightarrow k}^{(\lambda_1)}}.$$  \hspace{1cm} (11)

From the symmetric relationship between the left and the right eigenvectors we see that $w_{j \leftarrow i}^{(\lambda_1)} = v_{i \leftarrow j}^{(\lambda_1)}$, therefore

$$s_i = g_i. \hspace{1cm} (12)$$

Thus the asymptotic relative size of the surface of $T(\ell, i)$ is equal to the asymptotic relative number of nodes of label $i$ on the surface of any $T(\ell, r)$ independently of $r$. This quantity is expressed in terms of the components of the principal (left or right) eigenvector of the nonbacktracking matrix of the original graph $G$.

It is clear from the above constructions that for an arbitrary graph $G$, the interior of the infinite network $\mathcal{N}(m \rightarrow \infty)$ is given exactly by $T(\ell \rightarrow \infty, r)$, independently of $r$. Note, however, that the distribution of the numbers of nodes of different labels in $T(\ell \rightarrow \infty, r)$ is given by Eq. (6), while the same distribution for $\mathcal{N}(m \rightarrow \infty)$ is uniform—we made the same $m$ number of copies of every node. This discrepancy is reconciled by the presence of infinite loops: nodes in the infinite network $\mathcal{N}(m \rightarrow \infty)$ are “matched up” at infinity to provide the overall uniform distribution of node labels. It is informative to consider the dispersion $\sigma(d)$ of the relative number of node labels within a distance $d$ of an arbitrarily chosen starting node in both $T(\ell, r)$ and $\mathcal{N}(m)$. Suppose that $\ell$ and $m$ are such that the number of nodes in both $T(\ell, r)$ and $\mathcal{N}(m)$ is $V$ and that $V$ is large. The dispersion $\sigma(d)$ in both cases will converge to

$$\sigma^* = \langle g^2 \rangle - \langle g \rangle^2 \hspace{1cm} (13)$$

but will go to 0 for $\mathcal{N}(m)$ as $d$ approaches $\ln V$, while it remains the same for any $d$ in the case of $T(\ell, r)$. Here the variables $g$ are the asymptotic values given by Eq. (6). Figure 3 qualitatively illustrates this behavior.

**IV. EXAMPLE PROBLEMS**

We now consider the problem of approximating percolation on finite graphs. The critical singularity in percolation occurs only in infinite networks where the giant connected component can be strictly defined. In a large though finite random graph the so-called scaling window, in which the behavior of the largest connected cluster noticeably deviates from the singular behavior of the giant connected component, is narrow. In a recent work Karrer et al. [5] considered this problem of “approximate” percolation on a finite graph by solving self-consistency equations for the probabilities of reaching finite clusters by following a link in a given direction. Such self-consistency equations provide an exact solution for percolation in infinite locally treelike random networks, and in finite graphs they give an approximation to the behavior of the largest connected cluster. The larger and more “locally treelike” the finite graph is, the better this approximation is expected to be. For an arbitrary connected finite graph the “approximate critical point” is given by the inverse of the leading eigenvalue of the nonbacktracking matrix of the graph. It was shown in Ref. [5] that for infinite graphs, this is a tight lower bound on the exact percolation threshold. The question arises: More than treating it purely as an approximation, what is the exact meaning of the solutions of self-consistency equations on a finite graph? The results in Ref. [5] rely on solving the following set of basic equations:

$$H_{i \leftarrow j}(x) = 1 - p + px \prod_{k \in \partial j \setminus i} H_{j \rightarrow k}(x), \hspace{1cm} (14)$$

where $H_{i \leftarrow j}(x)$ is the probability generating function for the distribution of the sizes of finite clusters reachable by following links $i \leftarrow j$ from $i$ to $j$. (In the context of finite graphs, the meaning of finite clusters is: not the largest cluster.) Here $\partial j \setminus i$ denotes the set of the nearest neighbors of node $j$ excluding node $i$, and $p$ is the existence probability of any link. Clearly, in finite graphs, these equations do not hold exactly, but may be approximately correct if the graph in question is large and sparse. In our nonbacktracking expansion [the network $\mathcal{N}(m \rightarrow \infty)$], however, Eqs. (14) hold exactly, and by solving them we actually obtain the exact percolation results for $\mathcal{N}(m \rightarrow \infty)$. In Ref. [5] it is shown that the value of $p$, at which a non-trivial solution to Eqs. (14) appears, is given by the inverse of the leading eigenvalue of the nonbacktracking matrix of the given graph. This coincides with our result that the mean branching of the network $\mathcal{N}(m \rightarrow \infty)$ is also given by the leading eigenvalue of the nonbacktracking matrix. The inverse of this value is a true critical point in our NBE.
component, and is given by

\[ S = \frac{1}{N} \sum_i \left( 1 - \prod_{j \in \partial_i} H_{i \leftarrow j}(1) \right), \]  

(15)

where \( H_{i \leftarrow j}(1) \) are the solutions of Eq. (14) at \( x = 1 \), for a given value of \( p \). The degree distribution of any NBE coincides with the degree distribution of the original finite graph, therefore it necessarily has a cutoff—the maximum degree in the original graph. Consequently the order parameter exponent \( \beta = 1 \), \( S \approx \Omega(p - p_c) \), for the NBE of any finite graph.

We express the slope \( \Omega \) of the order parameter at \( p_c \) in terms of the components of the principal eigenvector of the nonbacktracking matrix which naturally emerges when linearizing Eq. (14). We cannot do this directly, unlike, e.g., in Refs. [16] and [17], since the nonbacktracking matrix is not symmetric, and so the full set of its eigenvectors does not form an orthogonal basis. To surpass this difficulty, we introduce a new symmetric matrix

\[ M = A - \lambda_1^{-1}(D - I), \]  

(16)

where \( \lambda_1 \) is the leading eigenvalue of the nonbacktracking matrix and matrices \( A, D, \) and \( I \) are the adjacency matrix, the degree matrix \( (D_{ij} = q_i \delta_{ij}) \), and the identity matrix, respectively. Matrix \( M \) is a special case of the Bethe Hessian matrix \( [18] \). Its eigenvectors already constitute a complete orthogonal basis. The leading eigenvalue of the nonbacktracking matrix, \( \lambda_1 \), is also an eigenvalue of \( M \), and the corresponding eigenvector components can be expressed in terms of the components of the principal eigenvector of the nonbacktracking matrix. These useful properties of matrix \( M \) enable us to calculate the slope of the order parameter at \( p_c = \lambda_1^{-1} \), (see the Appendix for details of the derivation):

\[ \Omega \approx \frac{\lambda_1 \left( \sum_i \sum_{j \in \partial_i} v^{(\lambda_1)}_{i \leftarrow j} \right) \sum_i (q_i - 1 - \lambda_1^2) \left( \sum_{j \in \partial_i} v^{(\lambda_1)}_{i \leftarrow j} \right)^2}{N \sum_i \sum_{j \in \partial_i} v^{(\lambda_1)}_{i \leftarrow j} \sum_{k,k' > k \in N_j \setminus i} q_i^{(\lambda_1)} v^{(\lambda_1)}_{i \leftarrow k} v^{(\lambda_1)}_{j \leftarrow k'} - \lambda_1 \sum_{k,k' > k \in \partial_i \setminus i} v^{(\lambda_1)}_{i \leftarrow k} v^{(\lambda_1)}_{j \leftarrow k'}}. \]  

(17)

Second approach requires only that the relevant matrix be diagonalizable, therefore it may be used to treat a wide range of problems in nonsymmetric systems, e.g., percolation on directed networks. A detailed discussion of these issues will be presented elsewhere.

The above methods can be applied to the general class of problems defined by the set of message passing equations

\[ a_{i \leftarrow j} = F_{i \leftarrow j}(p, \{a_{j \leftarrow k} : k \in \partial j \setminus i\}), \]  

(19)

with the condition that

\[ \frac{\partial F_{i \leftarrow j}}{\partial a_{j \leftarrow k}}|_{p=p_c} = c, \]  

(20)

where \( c \) is a constant, independent of \( i \leftarrow j \) and \( j \leftarrow k \). For any such model the solutions \( a_{i \leftarrow j} \) of Eq. (19) and the order parameter \( S \) near \( p_c \) can be calculated using our schemes (see the Appendix).

As another example of the application of our results, we consider an algorithm approximating optimal percolation, suggested in Ref. [6]. It was shown that an efficient way of destroying a network (removing the lowest number of nodes in order to disconnect the giant component) is by sequentially removing nodes according to their collective influence \( CI(i) \). This characteristic of node \( i \) is defined as

\[ CI(i) = (q_i - 1) \sum_{j \in \partial \text{Ball}(i,\ell)} (q_j - 1), \]  

(21)
Influences on a finite graph. We see that in this case the collective
node under consideration. Let us consider
the same problem on the nonbacktracking expansion of
the graph, respectively. It was also shown in [7] analytically
that removing a node according to Eq. (22) results in
the biggest decrease in $\lambda_1$, the leading eigenvalue of the
nonbacktracking matrix, i.e., the mean branching of the
NBE. Thus, this method produces "step-
wise optimal" percolation. From the result $w^{(\lambda_1)}_{i \to j} = v^{(\lambda_1)}_{i \to j}$
we see that using the centrality index of Eq. (23) is actu-
al equivalent to using
\[ CI_p(i) = \sum_{j \in \partial i} w^{(\lambda_1)}_{i \to j} v^{(\lambda_1)}_{i \to j} = \sum_{j \in \partial i} v^{(\lambda_1)}_{i \to j} v^{(\lambda_1)}_{j \to i}. \] (23)

Considering the nonbacktracking expansion we can give
a physical interpretation of the product $w^{(\lambda_1)}_{i \to j} v^{(\lambda_1)}_{i \to j}$. This
is the relative number of infinite paths in the NBE passing
through a link $i \to j$ (or link $j \to i$). The relative
number of infinite paths arriving at a link $i \to j$ is given
by $v^{(\lambda_1)}_{i \to j}$ and the relative number of infinite paths start-
ing from a link $i \to j$ is given by $w^{(\lambda_1)}_{i \to j}$ (see Sec. III).

Summing the product $w^{(\lambda_1)}_{i \to j} v^{(\lambda_1)}_{i \to j}$ over all neighbors $j$ of
node $i$, we get the relative number of infinite paths in the
NBE passing through a node $i$, which is exactly the $CI_p$
measure. One can consider the stepwise optimal
removal of edges instead of nodes. For this problem, the
expression (23) should be replaced by the edge centrality
measure
\[ CI_{ij} = w^{(\lambda_1)}_{i \to j} v^{(\lambda_1)}_{i \to j} = v^{(\lambda_1)}_{i \to j} v^{(\lambda_1)}_{j \to i}. \] (24)

where $q_i$ is the degree of node $i$ and $\partial \text{Ball}(i, \ell)$ is the set
of the nodes that are at a distance $\ell$ from node $i$. In this
algorithm, $\ell$ must be large but smaller than the diam-
er of the graph under consideration. Let us consider
the same problem on the nonbacktracking expansion of
a finite graph. We see that in this case the collective
influences $CI_\ell(i)$—taking $\ell \to \infty$—are just the asymp-
totic relative sizes of the surfaces of $T(\ell, i)$, multiplied
by $q_i - 1$. We have already expressed these relative sizes
in terms of the components of the principal eigenvector
of the nonbacktracking matrix of the original graph [see
Eq. (11)]. If the finite graph under consideration is large
and has no short loops, we can assume that using our
expression is a good approximation to the optimal way
of disconnecting the graph. An improvement on the co-
collective influence method is suggested in Ref. [7]. It was
shown empirically that better attack performance can be
achieved by sequentially removing nodes according to
the following centrality measure (collective influence
propagation)
\[ CI_p(i) = \sum_{j \in \partial i} w^{(\lambda_1)}_{i \to j} v^{(\lambda_1)}_{i \to j} + w^{(\lambda_1)}_{j \to i} v^{(\lambda_1)}_{j \to i}. \] (22)

V. ACCURACY AND LIMITATIONS OF THE MESSAGE PASSING APPROACH

We have shown that when we are using message pass-
ing equations to approximately solve any given problem
on a finite graph, we are actually solving the problem
exactly on the nonbacktracking expansion of the given
finite graph. At present there is no theory available to
determine how good this approximation is in the gen-
cral case. We now briefly discuss the accuracy of this
approximation, through the example of random percola-
tion, and how well it may be expected to work in certain
situations.

Note that although we defined the NBE for finite
graphs, it may also be defined for infinite networks as a
limit of finite networks. From Refs. [8] and [15] we know
that for infinite networks the critical point (for percola-
tion) of the NBE is a lower bound on that of the original
network. An even stronger statement is also true: the
relative size of the giant component in the NBE is an
upper bound on that of the original network (for any

FIG. 4. (Color online) (a) The order parameter $S$ as a func-
tion of $p$, computed using Eq. (15), for the 4-clique, an
Erdős–Rényi graph, the Zachary Karate Club network [19],
a subgraph of the Gnutella peer-to-peer file sharing network [20],
and a snapshot of the Internet at the autonomous sys-
tems level [21]. (b) Normalized slopes of the same example
networks, obtained using Eq. (14) [or Eq. (15)]. All normal-
ized slopes go to 1 at $p = p_c$. 

\begin{align*}
&\text{Giant component } S/\text{Activ. prob. } p \quad p - p_c \\
&\text{Activ. prob. } p \quad p - p_c \\
&\text{Activation probability } p \\
&\text{Giant component } S
\end{align*}
link activation probability \( p \). This is easily seen considering the following. The solutions \( H_{i \rightarrow j}(1) \) of Eq. (14) are a lower bound on the actual probabilities of reaching a finite cluster when following link \( i \leftrightarrow j \) in the direction of \( j \), due to the presence of loops in the original network (see Ref. [2]). According to Eq. (15), therefore, the order parameter of the NBE is an upper bound on that of the original network. With this result in mind we can quantify the \textit{badness} \( R \) of the NBE approximation as

\[
R = \frac{\int_0^1 |S_{\text{NBE}}(p) - S(p)| \, dp}{\int_0^1 S(p) \, dp}, \tag{25}
\]

where \( S(p) \) is the order parameter of the original network and \( S_{\text{NBE}}(p) \) is the order parameter of the corresponding NBE, for a given link activation probability \( p \). For infinite networks the absolute value in Eq. (25) is not necessary, as \( S_{\text{NBE}}(p) - S(p) > 0 \) for any \( p \) in this case. For finite graphs, however, \( S_{\text{NBE}}(p) - S(p) < 0 \) for \( p < p_{\text{c,NBE}} \), where \( p_{\text{c,NBE}} \) is the critical point of the NBE and \( S(p) \) is the relative size of the largest component. Therefore the absolute value in Eq. (25) is required to make this a universally applicable, meaningful measure of the badness of the approximation. For uncorrelated random networks the NBE coincides with the network itself (in the infinite size limit), therefore \( R = 0 \).

We now consider two representative examples of networks, having the same NBE, where the approximation fails: a two-dimensional square lattice and a synthetic modular network given by the following construction. Consider \( n \) subnetworks, each of which is a random regular network of \( m \) nodes and degree 4. In each of the subnetworks, let us remove two randomly selected links, leaving four nodes of degree 3. Then, let us connect these affected nodes to similar affected nodes in other subnetworks randomly, restoring the uniform degree 4 for every node in the network. In the limit \( n, m \rightarrow \infty \) this network has no finite loops and has modularity \( Q = 1 \), according to the definition of modularity in Ref. [24]. The corresponding NBE is exactly the same as the one for a two-dimensional square lattice. This NBE also coincides with that of a 5-clique; it is a random regular network of degree 4. The above examples are particularly badly approximated by the NBE as shown in Fig. 5 (a).

These two examples represent two classes of networks that are not well approximated by the NBE: low-dimensional networks and highly modular networks. Low-dimensional networks have many short loops and, therefore, are essentially non-treelike. A modification of the original message passing method has been suggested in Ref. [27] to deal with clustering and in Ref. [28] for networks with loopy motifs. Low dimensionality, however, implies many intertwined loops of many different sizes, which appears to be a significant factor in rendering such systems untreatable by the locally treelike approximation. Interestingly, even if there are no finite loops in the network, the NBE may still be a bad approximation if the modularity is very high [see the model modular network “Communities” in Fig. 5 (a)]. A message passing method for modular networks is suggested in Ref. [29], however, to use it, one needs to know the modular structure of the network \textit{a priori}. The problem is, that there is no unique definition of modularity in networks. Also, finding the community structure of a network—using any meaningful definition of modularity—typically leads to NP-hard decision problems [30]. In Fig. 5 (b) we present real-world examples of the two classes of networks discussed above: a planar network (road network of California [25]) and a modular network (Amazon copurchase network [20]). In all cases the simulation results differ significantly from the message passing solutions. See Table I for details.

![Simulation results and message passing solutions for the size of the giant connected component; examples of networks badly approximated by the non-backtracking expansion. (a) Synthetic networks: a two-dimensional square lattice and a model network of maximal modularity \( Q = 1 \). (b) Real-world networks: the road network of California [25] and an Amazon copurchase network [20]. In all cases the simulation results differ significantly from the message passing solutions. See Table I for details.](image)

FIG. 5. (Color online) Simulation results and message passing solutions for the size of the giant connected component; examples of networks badly approximated by the non-backtracking expansion. (a) Synthetic networks: a two-dimensional square lattice and a model network of maximal modularity \( Q = 1 \). (b) Real-world networks: the road network of California [25] and an Amazon copurchase network [20]. In all cases the simulation results differ significantly from the message passing solutions. See Table I for details.
module networks are ubiquitous and high modularity are two attributes that cause the modularities. We showed above that low dimensionality are small worlds, i.e., infinite-dimensional, and have low approximation are very close to 0 (see Table I). These networks studied in [5]—where the badness values of the approximation. (Note the very small coefficients (obtained by cutting the two links separating the modules). The NBE approximation is not particularly good when the whole network is considered, but is surprisingly good one for the modules separately. (Note the very small size, $N = 51$, of the pharynx network!)

in nature, therefore such excellent agreement as in Fig. 7 (or Ref. [5]) may not actually be a very common case.

Finally, we discuss another limitation of message passing methods, their time complexity. A significant advantage of using message passing equations as opposed to doing simulations is that, supposedly, the computational time required to find the solutions is much less than the time involved in doing the actual simulations and averaging over many realizations. This is not always true, however. The number of operations required in message passing methods (i.e., when solving the message passing equations by iterations) is

\[ T_{MP} \sim n_{conv} N(q) \sim n_{conv} L, \]

where the number of iterations to convergence is denoted $n_{conv}$, and $L$ is the number of links in the network. To achieve this complexity one may do the following. Considering the percolation problem, Eq. 14, at each iteration, for every node, first compute the product of all incoming messages. This operation has time complexity $\sim L$. To update each outgoing message from every node, divide the precalculated product for the starting node of every other part of the network. Therefore, $n_{conv}$ cannot be less than the network diameter, i.e., the longest of all the shortest paths between node pairs in the network. In a typical network demonstrating the small-world effect the diameter grows logarithmically with the system size $N$. In this case $n_{conv}$ can be expected to grow at least as $n_{conv} \sim \ln N$, and so for networks with sufficiently rapidly decaying degree distributions $T_{MP} \sim N \ln N$. (Note that in networks with a divergent second moment

### TABLE I. Size $N$ (number of nodes), mean degree $\langle q \rangle$, and badness value $R$ [defined in Eq. (25)] for the networks considered in Sec. IV

| Network                | $N$  | $\langle q \rangle$ | $R$    |
|------------------------|------|---------------------|--------|
| Random regular         | $10^7$ | 4                  | 0.00103|
| Square lattice         | $10^7$ | 4                  | 0.19752|
| Communities            | $10^7$ | 4                  | 0.31575|
| Roads                  | 1965206 | 2.82               | 0.47778|
| Amazon                 | 334836  | 5.53               | 0.08517|
| C. elegans full        | 448   | 21.17              | 0.02769|
| C. elegans pharynx     | 51    | 9.22               | 0.01745|
| C. elegans main body   | 397   | 22.69              | 0.00380|
| Erdős-Rényi           | $10^7$ | 5                  | 0.00064|
| Gnutella               | 62586  | 4.73               | 0.00037|
| Internet               | 22963  | 4.22               | 0.00071|

a Computer generated.
b Downloaded from: http://snap.stanford.edu/data/

c Downloaded from: http://www-personal.umich.edu/~mejn/netdata/
d Downloaded from: http://www.wormatlas.org/
of the degree distribution, the diameter may increase
with $N$ even more slowly than logarithmically. At the other extreme, namely in the case of a chain, the
diameter grows as $N$. Let us compare these results with
the time complexity of doing actual simulations.

Say we want to determine the size of the “giant” percolating
cluster in a large graph to a given accuracy (given
standard error). The quantity we are interested in is
an average over $n$ samples, $S = \langle s \rangle$, where $s$ is a bi-
ary quantity ($s = 1$ or $s = 0$) indicating whether or
not a given node of the graph is in the giant compo-
ent. The standard error of the estimate for $S$ is simply
$\sigma/\sqrt{n}$, where $\sigma$ is the standard deviation of $s$. There-
fore, for a given desired error in $S$ the number of node
samples needed is independent of the system size. To de-
termine whether a node belongs to the giant component,
we must explore the neighborhood of this node (e.g., us-
ing a simple breadth first search) visiting up to $l$ other
nodes, where $l \gg l^*$. Here $l^*$ is the typical size of a finite
cluster. The typical cluster size depends on the distance
from the critical point but not the system size. Consid-
ering the above, we arrive at the following conclusion.
At a given distance from the critical point, for a given
desired accuracy, the order parameter can be determined
using simulations in constant time for any network ar-
tecture. This is in strong contrast with the message
passing approach, where the accuracy is unknown and
the time complexity significantly higher. While the mes-
sage passing method may be faster in certain small net-
works, it is always inferior to simple simulations for large
enough networks.

VI. DISCUSSION AND EXTENSIONS

Message passing equations often provide a good ap-
proximation to the behavior of interacting models on
large sparse networks. In this paper, for any finite graph,
we have presented a network construction for which the
message passing equations are exact, thereby giving a
physical interpretation of this approximation. For a wide
range of problems, represented by message passing equa-
tions, we have shown how to express the solutions near
the critical point in terms of the principal eigenvector
components of the nonbacktracking matrix.

It is well established that this approximation should
work well for infinite-dimensional (small-world) networks
which are sparse and have no strong degree-degree cor-
relations. It is not clear how well it should work for
networks of low dimensionality, many loops, and corre-
lations. We have introduced a way of quantifying the
accuracy of the approximation and elaborated on two
distinct classes of networks where this approach fails:
low-dimensional and highly modular networks. As real-
world examples of these cases we considered the road net-
work of California and an Amazon copurchase network,
where the message passing equations, indeed, proved a
bad approximation.

Message passing methods based on the locally tree-like
approximation have a wide range of applicability and, in
some cases, provide good results. However, as we have
shown, for increasingly large systems these methods have
a high computational cost compared with simple simul-
ations. Also, while one can always find an estimate of
the error associated with simulations in a straightforward
way, the accuracy of message passing methods is
entirely unknown.

All the results in this paper apply also to weighted
graphs. In this case one must work with a weighted non-
backtracking matrix, obtained by multiplying each row
of the unweighted nonbacktracking matrix by the weight
of the corresponding link. Using this modified matrix,
the critical point and the solutions near the critical point
are determined exactly as in the unweighted case.

The network construction presented here may be gen-
eralized straightforwardly to directed graphs. The di-
rected percolation critical point is also given by the in-
verse of the leading eigenvalue of the nonbacktracking
matrix. Solutions near the critical point can be found
using the method we introduced for nonsymmetric sys-
tems (see Method 2 in the Appendix).

The nonbacktracking expansion can be generalized
to various percolation problems for interdependent and
multiplex networks. In the case of multiplex networks,
the only difference from single-layer (simplex) networks
is that the resulting infinite network (the NBE) will
consist of links of different colors, corresponding to
links of different layers. The message passing equations
written for any given model on finite multiplex networks
will be exact in the corresponding NBE, just as in
simplex networks. A number of percolation problems on
multiplex and interdependent networks exhibit discon-
tinuous transitions. In order to retain the interesting
physics in these systems, one has to account for the non-
linear terms in the basic equations. Finding the
analytical solutions similar to our Eqs. (17) and (18)
at discontinuous transitions in multiplex networks is an
interesting challenge for the future.

We thank L. Zdeborová for useful comments. This
work was supported by the FET proactive IP project
MULTIPLEX 317532.

APPENDIX: FINDING SOLUTIONS NEAR $p_c$

We stated that in the class of physical problems on undirected networks defined by Eqs. (19) and (20) the solution
near $p_c$ can be expressed in terms of the components of the principal eigenvector of the nonbacktracking matrix.
Here we derive this expression in the particular case of the message passing equations, Eq. (14), for percolation. We expand the message passing equations, Eq. (14), as

$$a_{i\rightarrow j} = p \left( \sum_{k \in \partial j \setminus i} a_{j\rightarrow k} - \sum_{k, k' > k \in \partial j \setminus i} a_{j\rightarrow k} a_{j\rightarrow k'} + \ldots \right), \quad \text{(A1)}$$

where $a_{i\rightarrow j} = 1 - H_{i\rightarrow j}(1)$. The set of eigenvectors of the nonbacktracking matrix, $v^{(\lambda)}$, forms a complete basis enabling us to express the vector $a = \{a_{i\rightarrow j}\}$ as a linear combination of the eigenvectors

$$a = \sum_\lambda C_\lambda v^{(\lambda)}, \quad \text{(A2)}$$

Notice that both $a$ and the coefficients $C_\lambda$ vary with $p$, approaching 0 when $\delta = p - p_c \to 0$. Replacing $a_{i\rightarrow j}$ with $\sum_\lambda C_\lambda v^{(\lambda)}_{i\rightarrow j}$ in Eq. (A1), we obtain the equation

$$\sum_\lambda C_\lambda v^{(\lambda)}_{i\rightarrow j}(p\lambda - 1) = p \sum_\lambda C_\lambda C_{\lambda'} \sum_{k, k' > k \in \partial j \setminus i} v^{(\lambda)}_{j\rightarrow k} v^{(\lambda')}_{j\rightarrow k'} + \ldots, \quad \text{(A3)}$$

where we have used the identity

$$\sum_{k \in \partial j \setminus i} v^{(\lambda)}_{j\rightarrow k} = \lambda v^{(\lambda)}_{j\rightarrow i},$$

since $v^{(\lambda)}$ is the eigenvector of the nonbacktracking matrix with eigenvalue $\lambda$. Since all $C_\lambda$ approach 0 as $\delta \to 0$, Eq. (A3) implies that there is an eigenvalue $\lambda_1 = 1/p_c$, such that $C_{\lambda_1} = O(\delta)$, and that for $\lambda \neq \lambda_1$ the coefficients $C_\lambda$ approach 0 as $\delta^2$ or more rapidly. Furthermore, by the Perron-Frobenius theorem if all components of $v^{(\lambda_1)}$ are non-negative, then the eigenvalue $\lambda_1$ is positive real and has the largest absolute value (hence, the subscript 1). In this way we find that at small $\delta$

$$a_{i\rightarrow j} \equiv C_{\lambda_1} v^{(\lambda_1)}_{i\rightarrow j}, \quad \text{(A4)}$$

where $v^{(\lambda_1)}$ are the components of the principal eigenvector of the nonbacktracking matrix. From this point onwards one may use two different approaches. Method 1 introduces an auxiliary symmetric matrix, to proceed with the derivation in the standard way. Method 2 exploits a useful property of the left and right eigenvectors of diagonalizable matrices, resulting in a much shorter derivation and a more compact final result.

**Method 1**

Let us consider the following sums, over the set of nearest neighbors $\partial i$ of node $i$,

$$b_{i\rightarrow} = \sum_{j \in \partial i} a_{j\rightarrow}, \quad b_{\rightarrow i} = \sum_{j \in \partial i} a_{j\rightarrow i}. \quad \text{(A5)}$$

We sum both sides of Eq. (A1):

$$b_{i\rightarrow} = p \left( \sum_{j \in \partial i} b_{j\rightarrow} - b_{\rightarrow i} - \sum_{j \in \partial i} \sum_{k, k' > k \in \partial j \setminus i} a_{j\rightarrow k} a_{j\rightarrow k'} + \ldots \right),$$

$$b_{\rightarrow i} = p \left( (q_i - 1) b_{i\rightarrow} - \sum_{j \in \partial i} \sum_{k, k' > k \in \partial j \setminus i} a_{i\rightarrow k} a_{i\rightarrow k'} + \ldots \right). \quad \text{(A6)}$$

Replacing $b_{\rightarrow i}$ in the first of these equations with the right-hand side of the second equation we obtain

$$b_{i\rightarrow} = p \left[ \sum_{j \in \partial i} b_{j\rightarrow} - p(q_i - 1) b_{i\rightarrow} + \sum_{j \in \partial i} \left( p \sum_{k, k' > k \in \partial j \setminus i} a_{i\rightarrow k} a_{i\rightarrow k'} - \sum_{k, k' > k \in \partial j \setminus i} a_{j\rightarrow k} a_{j\rightarrow k'} \right) + \ldots \right]. \quad \text{(A7)}$$
where \( q_i \) is the degree of node \( i \). Let us introduce the matrix \( M \) with elements \( M_{ij} \) defined as

\[
M_{ij} = \begin{cases} 
A_{ij} & \text{if } i \neq j, \\
-p_c(q_i - 1) & \text{if } i = j,
\end{cases}
\]

where \( A_{ij} \) are the elements of the adjacency matrix. The matrix \( M \) is symmetric, and so we can express \( b = \{b_{1+}, b_{2+}, \ldots, b_{N+}\} \) as a linear combination of its eigenvectors, \( u^{(\lambda)} \), which form a complete orthogonal basis,

\[
b = \sum_{\lambda} C_{\lambda} u^{(\lambda)}.
\]

In the following we denote the eigenvalues of \( M \) by \( \tilde{\lambda} \). These eigenvalues are different from the eigenvalues \( \lambda \) of the nonbacktracking matrix. We replace \( b_{i+} \) and \( a_{i+} \) in Eq. (A7) with their decompositions, Eqs. (A9) and (A4), respectively,

\[
\sum_{\lambda} C_{\lambda} u^{(\lambda)}(p_c - \lambda - \delta) = -p_c \sum_{\lambda, \lambda'} C_{\lambda} C_{\lambda'} \sum_{j \in \partial i} \left( p_c \sum_{k, k' > k \in \partial j \setminus i} v^{(\lambda)}_{i-k} v^{(\lambda')}_{i-k'} - \sum_{k, k' > k \in \partial j \setminus i} v^{(\lambda)}_{i-k} v^{(\lambda')}_{j-k'} \right) + \ldots.
\]

This equation implies that among the eigenvalues \( \tilde{\lambda} \) there is one, say \( \tilde{\lambda}^* \), equal to \( 1/p_c = \lambda_1 \). Note that the eigenvalue \( \tilde{\lambda}^* \) is not the leading eigenvalue of \( M \). Note also that according to Eqs. (A11), (A5), and (A9), \( C_{\lambda} u^{(\lambda)} = C_{\lambda_1} \sum_{j \in \partial i} v^{(\lambda_1)}_{i-j} \sim \delta \), while the rest of the coefficients \( C_{\lambda} \) and \( C_{\tilde{\lambda}} \) approach 0 as \( \delta^2 \) or more rapidly. To obtain \( C_{\lambda_1} \) we multiply both sides of Eq. (A10) by \( u^{(\lambda_1)} \) and sum over \( i \). The result is

\[
C_{\lambda_1} \cong \frac{\sum_{j \in \partial i} v^{(\lambda_1)}_{i-j} \sum_{j \in \partial i} \left( \sum_{k, k' > k \in \partial j \setminus i} v^{(\lambda_1)}_{i-k} v^{(\lambda_1)}_{i-k'} - \lambda_1 \sum_{k, k' > k \in \partial j \setminus i} v^{(\lambda_1)}_{j-k} v^{(\lambda_1)}_{j-k'} \right)}{(\sum_{j \in \partial i} v^{(\lambda_1)}_{i-j} \sum_{j \in \partial i} \left( \sum_{k, k' > k \in \partial j \setminus i} v^{(\lambda_1)}_{i-k} v^{(\lambda_1)}_{i-k'} - \lambda_1 \sum_{k, k' > k \in \partial j \setminus i} v^{(\lambda_1)}_{j-k} v^{(\lambda_1)}_{j-k'} \right))^2 \delta}.
\]

Note that the contribution of the first term on the left-hand side of Eq. (A10) is 0 because the eigenvectors \( u^{(\lambda)} \) are orthogonal. Finally, we express the order parameter near \( p_c \) in terms of the principal eigenvector components of the nonbacktracking matrix as

\[
S = \frac{1}{N} \sum_{i} \left( 1 - \prod_{j \in \partial i} H_{i-j}(1) \right) \cong \frac{1}{N} \sum_{i} \sum_{j \in \partial i} a_{i+} \cong \frac{C_{\lambda_1}}{N} \sum_{j \in \partial i} v^{(\lambda_1)}_{i-j} = \Omega \cdot (p - p_c),
\]

where

\[
\Omega \cong \frac{\lambda_1 \left( \sum_{j \in \partial i} v^{(\lambda_1)}_{i-j} \sum_{j \in \partial i} \left( \sum_{k, k' > k \in \partial j \setminus i} v^{(\lambda_1)}_{i-k} v^{(\lambda_1)}_{i-k'} - \lambda_1 \sum_{k, k' > k \in \partial j \setminus i} v^{(\lambda_1)}_{j-k} v^{(\lambda_1)}_{j-k'} \right) \right)^2}{N \sum_{j \in \partial i} v^{(\lambda_1)}_{i-j} \sum_{j \in \partial i} \left( \sum_{k, k' > k \in \partial j \setminus i} v^{(\lambda_1)}_{i-k} v^{(\lambda_1)}_{i-k'} - \lambda_1 \sum_{k, k' > k \in \partial j \setminus i} v^{(\lambda_1)}_{j-k} v^{(\lambda_1)}_{j-k'} \right)}.
\]

**Method 2**

Let us denote the right eigenvectors of the nonbacktracking matrix \( v^{(\lambda)} \) and the left eigenvectors \( w^{(\lambda)} \). The complex conjugate of any \( v^{(\lambda)} \) is orthogonal to all \( w^{(\lambda')} (\lambda \neq \lambda') \), and the conjugate of any \( w^{(\lambda)} \) is orthogonal to all \( v^{(\lambda')} (\lambda \neq \lambda') \). Indeed, let \( \lambda \neq \lambda' \) be two different eigenvalues of \( B \). Then clearly \( \lambda w^{(\lambda)} v^{(\lambda')} = w^{(\lambda)} B v^{(\lambda')} = \lambda' w^{(\lambda)} v^{(\lambda')} \), giving \((\lambda - \lambda') w^{(\lambda)} v^{(\lambda')} = 0\), and therefore, \( w^{(\lambda)} v^{(\lambda')} = 0 \).
Recall Eq. (A3),
\[
\sum_{\lambda} C_{\lambda} v^{(\lambda)}_{i \rightarrow j} (p\lambda - 1) = p \sum_{\lambda, \lambda'} C_{\lambda} C_{\lambda'} \sum_{k, k' : k > k' \in \partial_{\lambda} i} v^{(\lambda)}_{j \rightarrow k} v^{(\lambda')}_{j \rightarrow k'} + \ldots,
\]
and that \(v^{(\lambda)}_{i \rightarrow j}\) are components of the right eigenvector of the nonbacktracking matrix, corresponding to eigenvalue \(\lambda\). Multiplying both sides by \(w^{\lambda}_{i \rightarrow j}\) and summing over \(i \leftrightarrow j\) we see that the terms of \(\lambda \neq \lambda_1\) become 0. This procedure results in a considerably simpler expression for the coefficient \(C_{\lambda_1}\),
\[
C_{\lambda_1} \approx \frac{\lambda_1^2 \sum_{i \rightarrow j} v^{(\lambda_1)}_{i \rightarrow j} v^{(\lambda_1)}_{j \rightarrow i}}{\sum_{i \rightarrow j} v^{(\lambda_1)}_{i \rightarrow j} \sum_{k, k' : k > k' \in \partial_{\lambda_1} i} v^{(\lambda_1)}_{j \rightarrow k} v^{(\lambda_1)}_{j \rightarrow k'}} \delta,
\]
where we have used the property of the nonbacktracking matrix of undirected graphs \(w^{(\lambda)}_{i \rightarrow j} = v^{(\lambda)}_{j \rightarrow i}\), for simplification. The singularity of the giant component \(S = \Omega \delta\) has an amplitude, Eq. (A12):
\[
\Omega = \frac{\lambda_1^2 \left( \sum_{i \rightarrow j} v^{(\lambda_1)}_{i \rightarrow j} \right) \left( \sum_{i \rightarrow j} v^{(\lambda_1)}_{j \rightarrow i} v^{(\lambda_1)}_{j \rightarrow i} \right)}{N \sum_{i \rightarrow j} v^{(\lambda_1)}_{i \rightarrow j} \sum_{k, k' : k > k' \in \partial_{\lambda_1} i} v^{(\lambda_1)}_{j \rightarrow k} v^{(\lambda_1)}_{j \rightarrow k'}}.
\]
Equations (A14) and (A15) are equivalent to the expressions, Eqs. (A11) and (A13).

[1] R. J. Baxter, *Exactly Solved Models in Statistical Mechanics* (Courier Corp., Mineola, NY, 2007).
[2] E. A. Bender and E. R. Canfield, “The asymptotic number of labeled graphs with given degree sequences,” J. Combin. Theory Ser. A 24, 296 (1978).
[3] B. Bollobás, “A probabilistic proof of an asymptotic formula for the number of labelled regular graphs,” Eur. J. Combin. 1, 311 (1980).
[4] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, “Random graphs with arbitrary degree distributions and their applications,” Phys. Rev. E 64, 026118 (2001).
[5] B. Karrer, M. E. J. Newman, and L. Zdeborová, “Percolation on sparse networks,” Phys. Rev. Lett. 113, 208702 (2014).
[6] F. Morone and H. A. Makse, “Influence maximization in complex networks through optimal percolation,” Nature 524, 65 (2015).
[7] F. Morone, B. Min, L. Bo, R. Mari, and H. A. Makse, “Collective influence algorithm to find influencers via optimal percolation in massively large social media,” Sci. Rep. 6, 30062 (2016).
[8] S. Bau, N. C. Wormald, and S. Zhou, “Decycling numbers of random regular graphs,” Random Struct. Algor. 21, 397 (2002).
[9] K. Hashimoto, “Zeta functions of finite graphs and representations of p-adic groups,” in *Automorphic Forms and Geometry of Arithmetic Varieties: Advanced Studies in Pure Mathematics*, Vol. 15, edited by K. Hashimoto and Y. Namikawa (Kinokuniya, Tokyo, 1989) p. 211.
[10] F. Krzakala, C. Moore, E. Mossel, J. Neeman, A. Sly, L. Zdeborová, and P. Zhang, “Spectral redemption in clustering sparse networks,” Proc. Natl. Acad. Sci. USA 110, 20935 (2013).
[11] Y. Weiss and W. T. Freeman, “On the optimality of solutions of the max-product belief-propagation algorithm in arbitrary graphs,” IEEE Trans. Info. Theory 47, 736 (2001).
[12] A. Faqeeh, S. Melnik, and J. P. Gleeson, “Network cloning unfolds the effect of clustering on dynamical processes,” Phys. Rev. E 91, 052807 (2015).
[13] H. Minc, *Nonnegative Matrices* (John Wiley and Sons, New York, 1988).
[14] T. Martin, X. Zhang, and M. E. J. Newman, “Localization and centrality in networks,” Phys. Rev. E 90, 052808 (2014).
[15] K. E. Hamilton and L. P. Pryadko, “Tight lower bound for percolation threshold on an infinite graph,” Phys. Rev. Lett. 113, 208701 (2014).
[16] A. V. Goltsev, S. N. Dorogovtsev, and J. F. F. Mendes, “Percolation on correlated networks,” Phys. Rev. E 78, 051105 (2008).
[17] A. V. Goltsev, S. N. Dorogovtsev, J. G. Oliveira, and J. F. F. Mendes, “Localization and spreading of diseases in complex networks,” Phys. Rev. Lett. 109, 128702 (2012).
[18] A. Saade, F. Krzakala, and L. Zdeborová, “Spectral clustering of graphs with the Bethe Hessian,” in *Adv. Neural Inf. Process. Syst.* (2014) p. 406.
[19] W. W. Zachary, “An information flow model for conflict and fission in small groups,” J. Anthropol. Res. 33, 452 (1977).
[20] M. Ripeanu, A. Iamnitchi, and I. Foster, “Mapping the Gnutella network,” IEEE Internet Comput. 6, 50 (2002).
[21] D. Meyer, “University of Oregon Route Views Archive Project,” (2001).
[22] A. Braunstein, L. DallAsta, G. Semerjian, and L. Zdeborová, “Network dismantling,” Proc. Natl. Acad. Sci. USA 113, 12368 (2016).
[23] L. Zdeborová, P. Zhang, and H. J. Zhou, “Fast and simple decycling and dismantling of networks,” Sci. Rep.
13

[24] M. E. J. Newman and M. Girvan, “Finding and evaluating community structure in networks,” Phys. Rev. E 69, 026113 (2004).

[25] J. Leskovec, K. J. Lang, A. Dasgupta, and M. W. Mahoney, “Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters,” Internet Math. 6, 29 (2009).

[26] J. Yang and J. Leskovec, “Defining and evaluating network communities based on ground-truth,” Knowl. Inf. Syst. 42, 181 (2015).

[27] F. Radicchi and C. Castellano, “Beyond the locally tree-like approximation for percolation on real networks,” Phys. Rev. E 93, 030302 (2016).

[28] S. Yoon, A. V. Goltsev, S. N. Dorogovtsev, and J. F. F. Mendes, “Belief-propagation algorithm and the Ising model on networks with arbitrary distributions of motifs,” Phys. Rev. E 84, 041144 (2011).

[29] A. Faqeeh, S. Melnik, P. Colomer-de Simón, and J. P. Gleeson, “Emergence of coexisting percolating clusters in networks,” Phys. Rev. E 93, 062308 (2016).

[30] S. E. Schaeffer, “Graph clustering,” Comput. Sci. Rev. 1, 27 (2007).

[31] T. A. Jarrell, Y. Wang, A. E. Bloniarz, C. A. Brittin, M. Xu, J. N. Thomson, D. G. Albertson, D. H. Hall, and S. W. Emmons, “The connectome of a decision-making neural network,” Science 337, 437 (2012).

[32] R. Cohen and S. Havlin, “Scale-free networks are ultrasmall,” Phys. Rev. Lett. 90, 058701 (2003).

[33] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, “Metric structure of random networks,” Nucl. Phys. B 653, 307 (2003).

[34] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, “Catastrophic cascade of failures in interdependent networks,” Nature 464, 1025 (2010).

[35] G. Bianconi and S. N. Dorogovtsev, “Multiple percolation transitions in a configuration model of a network of networks,” Phys. Rev. E 89, 062814 (2014).

[36] S. Boccaletti, G. Bianconi, R. Criado, C. I. del Genio, J. Gómez-Gardeñes, M. Romance, I. Sendiña-Nadal, Z. Wang, and M. Zanin, “The structure and dynamics of multilayer networks,” Phys. Rep. 544, 1 (2014).

[37] M. Kivelä, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter, “Multilayer networks,” J. Complex Networks 2, 203 (2014).

[38] F. Radicchi, “Percolation in real interdependent networks,” Nature Phys. 11, 597 (2015).

[39] F. Radicchi and G. Bianconi, “Redundant interdependencies boost the robustness of multiplex networks,” Phys. Rev. X 7, 011013 (2017).

[40] G. J. Baxter, G. Bianconi, R. A. da Costa, S. N. Dorogovtsev, and J. F. F. Mendes, “Correlated edge overlaps in multiplex networks,” Phys. Rev. E 94, 012303 (2016).