Critical tipping point distinguishing two types of transitions in modular network structures

Saray Shai,[1,2,*] Dror Y. Kenett,[3] Yoed N. Kenett,[4] Miriam Faust,[4,5] Simon Dobson,[1] and Shlomo Havlin[6]
[1]*School of Computer Science, University of St Andrews, St Andrews, Fife KY16 9SX, Scotland, United Kingdom*
[2]*Department of Mathematics, University of North Carolina, Chapel Hill, North Carolina 27599, USA*
[3]*Center for Polymer Studies and Department of Physics, Boston University, Boston, Massachusetts 02215, USA*
[4]*Gonda Brain Research Center, Bar-Ilan University, 52900 Ramat-Gan, Israel*
[5]*Department of Psychology, Bar-Ilan University, 52900 Ramat-Gan, Israel*
[6]*Department of Physics, Bar-Ilan University, 52900 Ramat-Gan, Israel*
(Received 9 April 2015; revised manuscript received 22 October 2015; published 2 December 2015)

Modularity is a key organizing principle in real-world large-scale complex networks. The relatively sparse interactions between modules are critical to the functionality of the system and are often the first to fail. We model such failures as site percolation targeting interconnected nodes, those connecting between modules. We find, using percolation theory and simulations, that they lead to a "tipping point" between two distinct regimes. In one regime, removal of interconnected nodes fragments the modules internally and causes the system to collapse. In contrast, in the other regime, while only attacking a small fraction of nodes, the modules remain but become disconnected, breaking the entire system. We show that networks with broader degree distribution might be highly vulnerable to such attacks since only few nodes are needed to interconnect the modules, consequently putting the entire system at high risk. Our model has the potential to shed light on many real-world phenomena, and we briefly consider its implications on recent advances in the understanding of several neurocognitive processes and diseases.

DOI: 10.1103/PhysRevE.92.062805 PACS number(s): 89.75.Hc, 89.75.Fb, 64.60.ah

I. INTRODUCTION

Network science has become a leading approach to the study of emergent collective phenomena in complex systems, with a wide range of applications to fundamental real-world systems [1–4]. Many real-world systems have been shown to exhibit a modular structure, where nodes in smaller groups (called modules or communities) are connected more to each other than to the network at large, which is key to their behavior and functioning [5]. The modular organization of the Internet, and other large-scale infrastructures, tremendously enhances scalability and diffusion processes [6,7]. Modules of protein complexes and dynamic functional units constitute the building blocks of molecular networks [8]. Social and geographical regions with strong local ties promote the development of socioeconomic systems [9]. Finally, the nonrandom modular architecture of neural networks is considered crucial for the brain's functional demands of segregation and integration of information [10–12].

Detecting modules in networks has been an active research branch for many years, resulting in an extensive set of algorithms ubiquitously used to analyze and visualize large-scale data [13,14]. However, far less attention has been given to the implication of modularity to the function of networks and the physical mechanisms underlying such a structure. For example, modular structure has been suggested to arise in natural systems through the optimization of stability [15], efficiency [12], and evolvability [16] among others. The division of networks into modules has been suggested to enhance their robustness to certain cascading dynamics [17,18], as well as diffusion properties [19]. Several analytical frameworks have been developed to study networks with assortativity structure, where group of nodes with similar properties constitutes homogenous modules [20–22].

In the current study, we are interested in the implication of modularity to the resilience of networks. In a recent work, Bagrow *et al.* [23] showed that modular networks exhibit surprising percolation properties, such as the decoupling of modules (i.e., modules become nonoverlapping) as a result of random failure of nodes well before the network falls apart. Here we address analytically and by simulations the effect of failures of interconnected nodes, those connecting different modules.

Interconnected nodes play a key role in modular structures and their removal can have a deleterious effect on the network integrity [24], efficiency [25], and stability [26]. A recent study by da Cunha *et al.* empirically shows that module-based attacks targeting interconnected nodes ordered by betweenness centrality can be highly damaging, even more than attacks based solely on betweenness [27]. Masuda developed an efficient method to identify "globally important" nodes (quantified by their contribution to the connectivity in a coarse-grained network among modules) whose removal can fragment networks into small parts, thus providing an efficient immunization strategy [28]. In cases where transmissibility is high and/or when communities are dense, it has been shown that vaccinating interconnected nodes can be more efficient in controlling the spread of an epidemic than vaccinating high-degree nodes [29,30]. Moreover, the interconnected nodes are often the first to fail, as, for example, the case in the disruptive effect neurodegenerative diseases, such as schizophrenia and Alzheimer's disease, have on intermodular connectivity [31,32]. Also, changes in the concentration of interconnected nodes can be associated with functional transitions, for example, between physiological states [33]. Finally, it is often the case that interconnected nodes are considered to be important; for example, the New York City

*Corresponding author: sshai@live.unc.edu

and London airports interconnect modules of cities in these countries and provide an attractive target for attacks [7].

Utilizing recent advances in the understanding of interdependent networks [34–37], we introduce an analytical framework for studying the robustness of modular networks under attacks on interconnected nodes. We study a percolation process on networks consisting of a varying number of modules, m, and a varying degree of interconnected nodes. The analytical solution reveals two distinct percolation regimes separated by a critical number of modules m^*: For $m < m^*$ the system collapses abruptly as a result of the modules becoming disconnected from each another, while their internal structure is almost unaffected. In contrast, for $m > m^*$, the interconnected nodes play an important role also in the internal structure of modules. Therefore, the attack causes the modules themselves to collapse which in turn breaks continuously the entire system. Put another way, m^* represents the threshold above which modular structure itself becomes diffuse and the network returns to behaving as a single system.

II. MODEL

We consider a modular network with N nodes divided into m modules and we define p_{intra} and p_{inter} as the intra- and interlinks probabilities, respectively. Let β_i denote the fraction of nodes in module i, the average number of intramodule (intermodule) links connected to a node in module i is then given by

$$k_{\text{intra}} = p_{\text{intra}}(N\beta_i - 1), \tag{1}$$

$$k_{\text{inter}} = p_{\text{inter}}N(1-\beta_i). \tag{2}$$

We define α to be the ratio between the probabilities for an intra- and intermodule link,

$$\alpha = \frac{p_{\text{intra}}}{p_{\text{inter}}}. \tag{3}$$

In Figs. 1(a)–1(c) we present examples of modular networks consisting of $m = 5$ equal-size modules and different values of α. Note that the ratio between the number of intermodules links and intramodule links depends not only on α but also on the number of modules, in this case $\beta_i = \frac{1}{m}$, yielding

$$\frac{k_{\text{intra}}}{k_{\text{inter}}} = \frac{p_{\text{intra}}\left(\frac{N}{m}-1\right)}{p_{\text{inter}}N\left(1-\frac{1}{m}\right)} \sim \frac{\alpha}{m-1}. \tag{4}$$

Thus, our model considers systems comprised of more modules to have more interlinks, as illustrated in Fig. 1(d).

III. FORMALISM

Given the model for generating random modular networks described above, we proceed to study their percolation properties. In particular, we use the multivariate generating functions formalism [36] to derive the percolation threshold and the size of the giant component. The formalism assumes a system of interconnected modules described by a set of multidegree distributions, $\{p^i_{k_1k_2\ldots k_m}\}$, where $p^i_{k_1k_2\ldots k_m}$ is the fraction of all nodes in the module i that have k_1 links to nodes in module 1, k_2 links to nodes in module 2, etc. The multidegree distribution for each module may be written in the form of a generating function:

$$G_i(x_1x_2\ldots x_m) = \sum_{k_1k_2\ldots k_m=0}^{\infty} p^i_{k_1k_2\ldots k_m} {x_1}^{k_1}{x_2}^{k_2}\ldots{x_m}^{k_m}. \tag{5}$$

Note that this function is simply an extension of the single-network generating function presented in Ref. [39]. The partition of a node's degree into m degrees, corresponding to its number of connections in each module, allows a finer analysis of heterogeneous systems, which is essentially the objective of studying modular or interacting networks.

Following Leicht and D'Souza [36], the distribution of the sizes of components reached by following a randomly chosen link between modules j and i to a node in module i and is generated by

$$H_{ij}(\mathbf{x}) = x_i G_{ij}(H_{1i}(\mathbf{x}), H_{2i}(\mathbf{x}), \ldots, H_{mi}(\mathbf{x})) \tag{6}$$

and the distribution of the sizes of components of a randomly chosen i node is generated by

$$H_i(\mathbf{x}) = x_i G_i[H_{1i}, H_{2i}, \ldots, H_{mi}], \tag{7}$$

where $\mathbf{x} = (x_1, x_2, \ldots, x_m)$, $\mathbf{1} = (1, 1, \ldots, 1)$, and $G_{ij}(\mathbf{x})$ is the generating function for the branching process defined as

$$G_{ij}(\mathbf{x}) = \left[\frac{\partial G_i}{\partial x_j}(\mathbf{1})\right]^{-1} \frac{\partial G_i}{\partial x_j}(\mathbf{x}). \tag{8}$$

Finally, using the generating functions given in Eqs. (6) and (7), the average number of nodes from module j in the component of a randomly chosen node in module i is given by

$$\langle s_i \rangle_j = \frac{\partial H_i}{\partial x_j}(\mathbf{x})\Big|_{\mathbf{x}=\mathbf{1}} = \delta_{ij} + \sum_{l=1}^{m} \frac{\partial G_i}{\partial x_l}(\mathbf{1}) \frac{\partial H_{li}}{\partial x_j}(\mathbf{1}), \tag{9}$$

where δ_{ij} denotes the Kronecker δ.

In the case of a modular Erdős-Rényi (ER) network [40,41] consisting of equal-size modules where both the intra- and

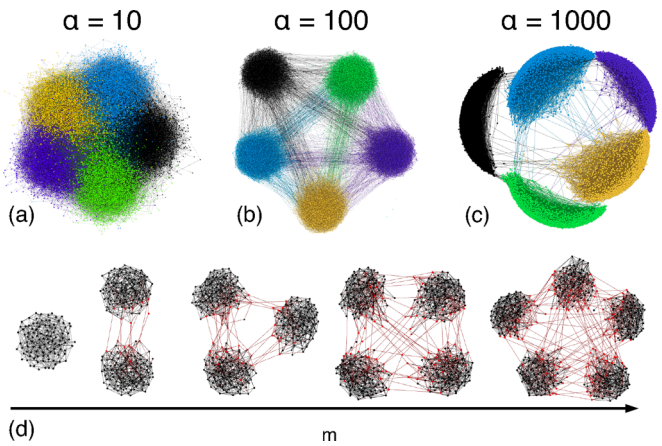

FIG. 1. (Color online) Visualization of the model for generating random modular networks. [(a)–(c)] Illustration of the effect of α on the obtained modular network using Gephi [38] on a Erdős-Rényi (ER) network of size $N = 10\,000$ with mean degree $k = 8$ divided into $m = 5$ modules. (d) Illustration of the effect of the number of modules m on the obtained network with a number of intermodule links increasing with the number of modules. Interconnected nodes and links are shown in red (light gray).

interconnectivity are Poisson distributed with means k_{intra} and k_{inter}, respectively, we can derive an exact solution for the critical percolation. In particular, we obtain

$$G_i(x) = G_{ij}(x) = e^{k_{\text{intra}}(x_i-1)} e^{\frac{k_{\text{inter}}}{m-1}\sum_{l\neq i}(x_l-1)} \tag{10}$$

for any $1 \leqslant j \leqslant m$. Thus, Eq. (9) can be written as

$$\langle s_i \rangle_j = \delta_{ij} + k_{\text{intra}} \frac{\partial H_i}{\partial x_j}(\mathbf{1}) + \frac{k_{\text{inter}}}{m-1} \sum_{\substack{l=1 \\ l\neq i}}^{m} \frac{\partial H_l}{\partial x_j}(\mathbf{1}), \tag{11}$$

where

$$\frac{\partial H_i}{\partial x_j}(\mathbf{1}) = \delta_{ij} + k_{\text{intra}} \frac{\partial H_i}{\partial x_j}(\mathbf{1}) + \frac{k_{\text{inter}}}{m-1} \sum_{\substack{l=1 \\ l\neq i}}^{m} \frac{\partial H_l}{\partial x_j}(\mathbf{1}). \tag{12}$$

Solving the system for $\langle s_1 \rangle_1$ we obtain

$$\begin{aligned}
\frac{\partial H_1}{\partial x_1}(\mathbf{1}) &= 1 + k_{\text{intra}} \frac{\partial H_1}{\partial x_1}(\mathbf{1}) \\
&\quad + \frac{k_{\text{inter}}}{m-1}\left[\frac{\partial H_2}{\partial x_1}(\mathbf{1}) + \cdots + \frac{\partial H_m}{\partial x_1}(\mathbf{1})\right] \\
&= 1 + k_{\text{intra}} \frac{\partial H_1}{\partial x_1}(\mathbf{1}) + \frac{k_{\text{inter}}}{m-1}\left[\frac{k_{\text{inter}} \frac{\partial H_1}{\partial x_1}(\mathbf{1})}{1 - k_{\text{intra}} - \frac{m-2}{m-1} k_{\text{inter}}}\right] \\
&\Rightarrow \frac{\partial H_1}{\partial x_1}(\mathbf{1})\left[1 - k_{\text{intra}} - \frac{\frac{k_{\text{inter}}^2}{m-1}}{1 - k_{\text{intra}} - \frac{m-2}{m-1} k_{\text{inter}}}\right] = 1 \\
&\Rightarrow \frac{\partial H_1}{\partial x_1}(\mathbf{1}) \\
&= \frac{1 - k_{\text{intra}} - \frac{m-2}{m-1} k_{\text{inter}}}{(1 - k_{\text{intra}})\left(1 - k_{\text{intra}} - \frac{m-2}{m-1} k_{\text{inter}}\right) - \frac{k_{\text{inter}}{}^2}{m-1}}.
\end{aligned} \tag{13}$$

Thus, $\frac{\partial H_1}{\partial x_1}(\mathbf{1})$ diverges when $(1 - k_{\text{intra}})[1 - k_{\text{intra}} - \frac{(m-2)k_{\text{inter}}}{m-1}] - \frac{k_{\text{inter}}{}^2}{m-1} = 0$. This is also where all $\frac{\partial H_i}{\partial x_1}(\mathbf{1})$ diverges, and therefore the giant component emerges when the following equation is satisfied:

$$(1 - k_{\text{intra}})\left[1 - k_{\text{intra}} - \frac{(m-2)k_{\text{inter}}}{m-1}\right] - \frac{k_{\text{inter}}{}^2}{m-1} = 0. \tag{14}$$

This condition yields $k = k_{\text{intra}} + k_{\text{inter}} = 1$ for every m, recovering the standard result for single networks without modules. Thus, in the case of random node failures, the percolation threshold as well as the size of the giant component only depends on the average degree, k, and the modular structure has no effect on the percolation transition. The same result holds for scale-free networks [42], as shown by Leicht and D'Souza [36]; however, a closed form equation using the generating function approach is not available.

A more realistic case of attacks on the network is the preferential removal of interconnected nodes. This type of attacks can be studied using Callaway *et al.*'s [43] approach developed for studying the robustness of single networks to intentional attacks. Here, we extend the formalism from single to multiple networks in a similar manner that was done by Leicht and D'Souza [36]. In this approach, the occupation probability of nodes is not constant as before, but is a function of the node's degree. Let q denote the probability that a randomly chosen interconnected node is occupied, and $1 - q$ the probability that it is removed. For ER networks with average intra- and interdegree k_{intra}, k_{inter}, respectively, this probability is related to the general occupation probability, p, according to

$$q = \frac{p - e^{-k_{\text{inter}}}}{1 - e^{-k_{\text{inter}}}}. \tag{15}$$

In this case, the degree distribution of occupied nodes in module i is given by

$$\begin{aligned}
F_i(\mathbf{x}) &= e^{-k_{\text{inter}}} e^{k_{\text{intra}}(x_i-1)} + q[G_i(\mathbf{x}) - e^{-k_{\text{inter}}} e^{k_{\text{intra}}(x_i-1)}] \\
&= e^{k_{\text{intra}}(x_i-1)-k_{\text{inter}}}(1-q) + q G_i(\mathbf{x})
\end{aligned} \tag{16}$$

and the average number of occupant nodes from module j in the component of a randomly chosen node from module i is given by

$$\begin{aligned}
\langle s_i \rangle_j &= \delta_{ij} F_i(\mathbf{1}) + k_{\text{intra}} F_i(\mathbf{1}) \frac{\partial J_{ii}}{\partial x_j}(\mathbf{1}) \\
&\quad + q \frac{k_{\text{inter}}}{1 - \beta_i} \sum_{\substack{l=1 \\ l\neq i}}^{m} \beta_j \frac{\partial J_{li}}{\partial x_j}(\mathbf{1}).
\end{aligned} \tag{17}$$

For equal-size modules where $\beta_i = \frac{1}{m}$, Eq. (17) can be solved analytically. Using similar algebra as before, we obtain

$$\begin{aligned}
&\frac{\partial J_{11}}{\partial x_1}(\mathbf{1}) + \cdots + \frac{\partial J_{mm}}{\partial x_1}(\mathbf{1}) \\
&= F_1(\mathbf{1}) + k_{\text{intra}} F_1(\mathbf{1}) + \frac{q^2\, k_{\text{inter}}}{1 - q\, k_{\text{inter}}} \\
&\quad + \frac{q^2\, k_{\text{intra}}\, k_{\text{inter}}\left[\frac{\partial J_{11}}{\partial x_1}(\mathbf{1}) + \cdots + j_{mm}\right]}{1 - q\, k_{\text{inter}}} \\
&\Rightarrow \frac{\partial J_{11}}{\partial x_1}(\mathbf{1}) + \cdots + \frac{\partial J_{mm}}{\partial x_1}(\mathbf{1}) \\
&= \frac{F_1(\mathbf{1})(1 - q\, k_{\text{inter}}) + q^2\, k_{\text{inter}}}{[1 - k_{\text{intra}} F_1(\mathbf{1})](1 - q\, k_{\text{inter}}) - q^2\, k_{\text{intra}}\, k_{\text{inter}}}.
\end{aligned} \tag{18}$$

Thus, the critical occupation probability of interconnected nodes, q_c, above which a giant component emerges, is given by

$$q_c = \begin{cases} 0 & m < m^*(k_{\text{inter}}), \quad \text{(19a)} \\ \frac{-b+\sqrt{b^2-4ac}}{2a} & m > m^*(k_{\text{inter}}), \quad \text{(19b)} \end{cases}$$

where $a = k_{\text{intra}} k_{\text{inter}} e^{-k_{\text{inter}}}$, $b = k_{\text{intra}} + k_{\text{inter}} - k_{\text{intra}} e^{-k_{\text{inter}}} - k_{\text{intra}} k_{\text{inter}} e^{-k_{\text{inter}}}$, and $c = k_{\text{intra}} e^{-k_{\text{inter}}} - 1$. Note that two possible solutions emerge, one from Eq. (19) and one trivial solution $q_c = 0$, corresponding to the removal of all the interconnected nodes. Each solution corresponds to a distinct percolation regime, separated by a critical number of modules, m^*, defined as the crossing of the two solution curves and is a function of the degree of interconnections, see Eq. (4). From these solutions, we obtain the critical occupation probability p_c, using Eq. (15).

The value of m^* is obtained when $q_c = 0$ is a solution of Eq. (19); in other words, by setting $q = 0$ in the denominator of Eq. (18), which yields

$$(k - k^*_{\text{inter}}) e^{-k^*_{\text{inter}}} = 1, \tag{20}$$

where k^*_{inter} is the mean interdegree at the point m^*. From Eq. (4), k^*_{inter} is related to m^* according to

$$\frac{k-k^*_{\text{inter}}}{k^*_{\text{inter}}}=\frac{\alpha}{m-1}\Rightarrow k^*_{\text{inter}}=\frac{k(m^*-1)}{\alpha+m^*-1}. \quad (21)$$

Replacing Eq. (21) in Eq. (20), we obtain the implicit equation

$$k\left(1-\frac{m^*-1}{\alpha+1-m^*}\right)e^{-k\frac{m^*-1}{\alpha+1-m^*}}=1 \quad (22)$$

from which we can derive m^* as a function of k and α.

Once the giant component emerges ($p>p_c$), the fraction of occupied nodes from module i in the giant component, S_i is given by

$$S_i=e^{-k_{\text{inter}}}(1-q)(1-e^{-k_{\text{intra}}S_i})+q\left[1-e^{-(k_{\text{intra}}S_i+\frac{k_{\text{inter}}}{1-\beta_i}\sum_{j\neq i}\beta_j S_j)}\right]. \quad (23)$$

Substituting $\beta_i=\frac{1}{m}$ in Eq. (23), yields

$$S=e^{-k_{\text{inter}}}(1-q)(1-e^{-k_{\text{intra}}S})+q[1-e^{-(k_{\text{intra}}S+k_{\text{inter}}S)}]. \quad (24)$$

For $k_{\text{intra}}=0$, only a fraction q of the nodes in the network are connected, and one obtains $S=q(1-e^{-kS})$, and for $k_{\text{inter}}=0, S=(1-e^{-kS})$, recovering the standard results for percolation in single ER networks [40,41].

For nonequal module size, we can solve Eq. (23) numerically. We obtain the same q_c in this case, and in particular that $S=S_i$ for every $1\leqslant i\leqslant m$, i.e., modules of different sizes percolate at the same p. This result can be explained by noting that the average inter- and interdegrees of nodes are independent of modules size. Thus, the number of interlinks connecting to a module (as well as the number of interlinks connecting nodes inside the module) is proportional to its size, see Eq. (23). Therefore, the fraction of nodes in the giant component from each module is equal to the total fraction of nodes in the giant component.

In the case of scale-free networks [1,2], the giant component and q_c can be obtained numerically from the analytical results. In particular, we consider the case of scale-free networks with average intra- and interdegree k_{intra}, k_{inter}, respectively, generated by [44]

$$G(x)=\sum_{k=s_{\text{intra}}}^{K}\frac{(k+2)^{1-\lambda}-(k+1)^{1-\lambda}}{(K+2)^{1-\lambda}-(s_{\text{intra}}+1)^{1-\lambda}}x^k+\sum_{k=s_{\text{inter}}}^{K}\frac{(k+2)^{1-\lambda}-(k+1)^{1-\lambda}}{(K+2)^{1-\lambda}-(s_{\text{inter}}+1)^{1-\lambda}}x^k, \quad (25)$$

where s_{intra} and s_{inter} are the minimal degree cutoff (chosen such that the desired mean degree is obtained) and $K=1000$ is the maximal degree cutoff. The degree distribution of occupied nodes is then generated by

$$F(x)=p_{\text{not}}(1-q)\sum_{\|=\int_{\text{intra}}}^{K}\frac{(k+2)^{1-\lambda}-(k+1)^{1-\lambda}}{(K+2)^{1-\lambda}-(s_{\text{intra}}+1)^{1-\lambda}}x^k+qG(x), \quad (26)$$

where p_{not} is fraction of nodes with no interconnections. From Eqs. (25) and (26), we obtain the size of the giant component

$$S=F(1)-F(u),\quad \text{where } u=1-F'(1)+F'(u). \quad (27)$$

By solving numerically the set of analytic Eqs. (25)–(27), we obtain a critical occupation probability of interconnected nodes, q_c, for which u is smaller than 1. Comparing with the trivial occupation probability $q_c=0$ for which the modules are disconnected, we obtain the critical number of modules, m^*.

IV. RESULTS

In Fig. 2, we provide support to the analytical solution given in Eqs. (15)–(25) by extensive numerical simulations of ER modular networks of size $N=600\,000$ nodes. First, we show the percolation threshold as a function of the number of modules m for ER networks where the mean degree is kept fixed $k=4$ and $\alpha=100$, see Fig. 2(a). In the regime $m<m^*$ the attack on interconnected nodes mainly breaks the connectivity between the modules leaving their internal structure intact. Therefore, only the removal of all the interconnected nodes ($q_c=0$) breaks down the giant component. Note that, as demonstrated in Figs. 2(c) and 2(d), for $m<m^*$ the percolation transition is abrupt, while for $m>m^*$ it is continuous. Note that for $m<m^*$, p_c is defined as the point where the discontinuous jump occurs, i.e., where the modules become separated.

In order to demonstrate this phenomenon, in Fig. 3 we visualize the giant component at $S=0.1$ (close to full collapse) with interconnected nodes shown in black and all other nodes colored according to the module they belong to. For a network with $m=4<m^*$, random node failure destroys the internal structure of the modules evenly, see Fig. 3(a). In this random failure case, all the modules always appear in the giant component (i.e., there is always at least one node from each module in the giant component) as

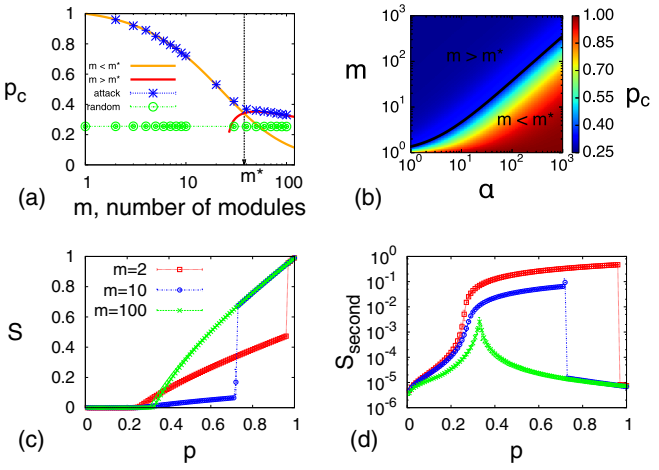

FIG. 2. (Color online) Two percolation regimes when attacking interconnected nodes. (a) p_c as a function of m calculated for ER networks with $k=4$, $\alpha=100$. Simulation points obtained from at least 1000 simulation runs of networks of size $N=600\,000$. Solid lines represent the analytical result obtained in Eqs. (15) and (19). (b) Two-parameter (α, m) phase diagram for a fixed mean degree $k=4$. The black line corresponds to the critical number of modules m^* obtained from Eq. (22). [(c) and (d)] Fraction of nodes in the largest cluster S and second largest cluster S_{second}, respectively, with solid lines representing the analytical result obtained in Eq. (24) and symbols are from simulations.

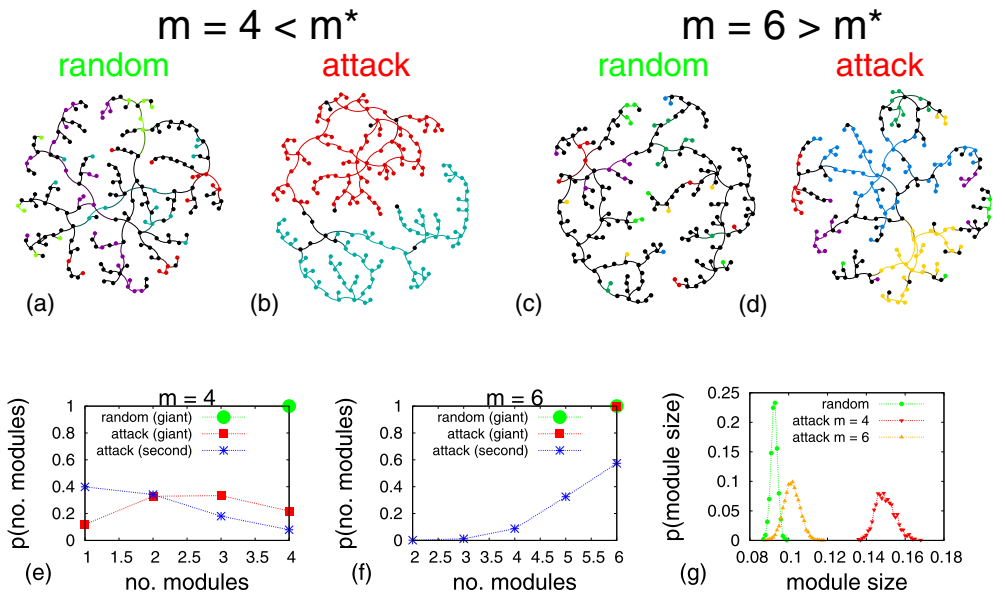

FIG. 3. (Color online) Size of modules in the giant component at $S = 0.1$. Visualization is shown for networks of size $N = 2000$ with mean degree $k = 4$ and $\alpha = 10$, at the point where the giant component contains 10% of the nodes ($S = 0.1$) [(a) and (c)] for random node removal and [(b) and (d)] for attack on interconnected nodes. [(e) and (f)] Distribution of the number of modules in the giant component and second-largest component at $S = 0.1$. A module is considered to be part of a component if at least one of its nodes are part of the component. (g) Distribution of the size of modules in the giant component at $S = 0.1$, normalized by the initial module size. Note that in (g), the size of modules is measured by reconstructing the graph of each module in the giant component and counting its number of nodes in this graph. In other words, interconnected nodes that have been detached from their original module are not considered. Results obtained by at least 1000 simulation runs of networks of size $N = 600\,000$ with mean degree $k = 4$.

shown in Fig. 3(e), and the size of modules is very narrowly distributed, see Fig. 3(g). In contrast to random failure, when attacking the interconnected nodes (at $S = 0.1$), see Fig. 3(b), not all the modules remain in the giant component [for example, in Fig. 3(b) there are only two of them]. However, the modules that do remain are almost intact, containing 14.6% of their initial nodes, significantly more than in the random case (10%). Thus for $m < m^*$ the network collapses by the breakdown of internally connected modules from the network.

In contrast, for $m > m^*$, the interconnected nodes play an important role also in the internal structure of modules and therefore their removal breaks down both the internal and external connectivity. Nevertheless, the attack still leaves them slightly more complete than in the case of random removal, see Fig. 3(c) and 3(d). Furthermore, in the case of attack, almost *all* modules appear in the giant component [see Fig. 3(f)], and thus their relative size is smaller compared to the $m < m^*$ case [see Fig. 3(g)]. As m increases, the difference between the attack and random cases becomes smaller since the network becomes more homogeneous and in the very large m limit, the modules stop being a factor at all, resulting in the convergence of the percolation threshold to the one obtained for random failures. When compared with degree-based attacks, studies suggest a phase transition at which removing interconnected nodes is more efficient when failure probability is low and/or when modules are dense, but targeting high-degree nodes is more efficient otherwise [29,30].

In Figs. 2(c) and 2(d) we show the size of the giant component, S, and the second giant component, S_{second}, as a function of p, observing an abrupt decrease in S due to failures of entire modules for $m < m^*$. In addition, while for $m = 100 > m^*$ we observe a regular second-order percolation transition characterized by the continuous decrease of S and the sharp peak in S_{second}, the case of $m < m^*$ demonstrates an abrupt transition characterized by the sudden collapse of S_{second}. Note that in this case, p_c is considered as the point where the discontinuous jump occurs.

Figure 2(b) shows the two-parameter (α, m) phase diagram for a fixed mean degree $k = 4$. The black line corresponds to m^* as a function of α obtained from Eq. (22). Below the black line, the percolation threshold is considerably higher, indicating that systems with number of modules below m^* can be very fragile. As the number of modules increase, the system becomes more balanced, and p_c values are decreasing until converging to the percolation threshold of single ER networks for large m.

In Fig. 4, we show analytical and numerical results for scale-free networks with scaling exponent λ and $\alpha = 100$, for both intra- and interdegree distributions. Here the critical point, m^*, is significantly larger compared to ER networks, with $m^* = 169$ for $\lambda = 2.9$, and m^* goes to infinity for $\lambda = 2.5$, see Figs. 4(a) and 4(b). In other words, for $\lambda = 2.5$, the system always collapses as a result of the modules becoming disconnected. This is due to the resilience of scale-free networks, which makes them very hard to fragment the modules internally. However, while the modules themselves are robust, in scale-free networks, only a few nodes are needed to interconnect the modules, consequently putting the entire system at high risk, see Figs. 4(a)–4(d).

Finally, it is possible to investigate p_c as a function of k_{inter} for fixed m and k_{intra}. A similar transition in p_c is observed

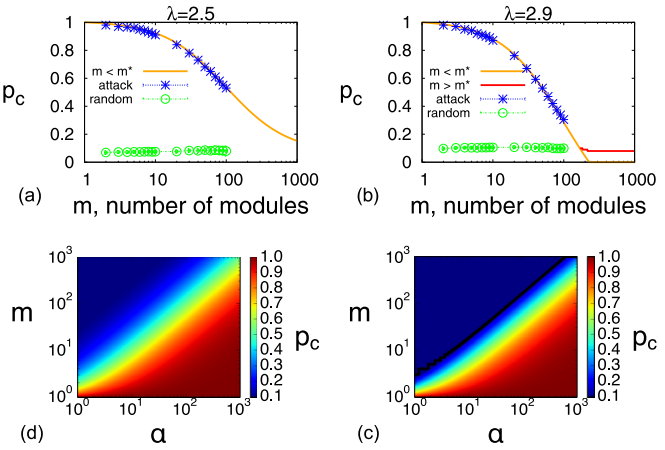

FIG. 4. (Color online) Fragility of modular scale-free networks [(a) and (b)] p_c as a function of m calculated for scale-free networks with mean degree $k = 4$, $\alpha = 100$, and scaling exponents $\lambda = 2.5$ and $\lambda = 2.9$, respectively. Simulation points obtained from at least 1000 simulation runs of networks of size $N = 600\,000$. Solid lines represent the numerical solution for Eqs. (25)–(27). Two-parameter (α, m) phase diagram for a fixed mean degree $k = 4$. The black line in (d) corresponds to the critical number of modules, m^* obtained by solving Eqs. (25)–(27) numerically and finding the point where the two possible solutions for p_c cross. Note that in (c) there is no black line since m^* approach to infinite for $\lambda = 2.5$.

[see Figs. 5(a)–5(c)], but the critical point is now a function of the concentration of interconnected nodes, k^*_{inter}. In order to find k^*_{inter} we compare the value of p_c when $k_{\text{inter}} = 0$ to the p_c for a single-module network with mean degree $k = k_{\text{intra}}$. For ER networks we obtain

$$k_{\text{intra}} e^{-k^*_{\text{inter}}} = 1. \tag{28}$$

In other words, k^*_{inter} is the point where it is equally easy to disconnected the modules (by removing all interconnected nodes) and to break the modules internally. For scale-free networks, k^*_{inter} is obtained by solving Eqs. (25)–(27) numerically and finding the point where the two possible solutions for p_c cross. In addition, we show the two-parameter (k_{intra}, k_{inter}) phase diagram for a fixed number of modules $m = 10$, see Figs. 5(d)–5(f). For ER networks, we obtain $k^*_{\text{inter}}(1) = 0$, meaning that modules with mean intradegree 1 become fragmented as soon as k_{inter} is positive, in agreement with the known percolation threshold for ER networks. Scale-free networks exhibit a high-risk area depending almost solely on k_{inter} and independent of k_{intra}. While broader degree distributions result in more robust single-module networks, in this case the opposite is true. The scale-free network with $\lambda = 2.5$ has a very large high-risk area compared with the other more homogenous networks. The reason is that a broader distribution with the same average degree implies that there are more low-degree nodes. Thus although the modules themselves remain robust, there are less interconnected nodes, which make it easier to break the network.

V. DISCUSSION

In summary, motivated by examples from real-world modular networks, we consider attack on interconnected nodes, those connecting between modules. Unlike the case of random failure, in such attacks we find that for each α there exists a critical point m^* below which the system first separates abruptly into modules before being completely destroyed. Our analysis reveals rich phase transition phenomena that could be applied to a wide variety of systems, from the optimal design of infrastructure, the efficient immunization approach in modular networks (where epidemic spreading can be prevented at a low cost by immunizing interconnected nodes), and new insights and understandings of brain disorders. In particular, the modular architecture of neural structural and functional networks is considered a fundamental principle of the brain [11], and disrupted brain modular organization is related to neuropathology [31,32]. For example, while schizophrenia has been related to a breakdown in brain modules, AlzheimerÕs disease has been related to disrupted connectivity between modules [31]. Similarly, recent studies have empirically

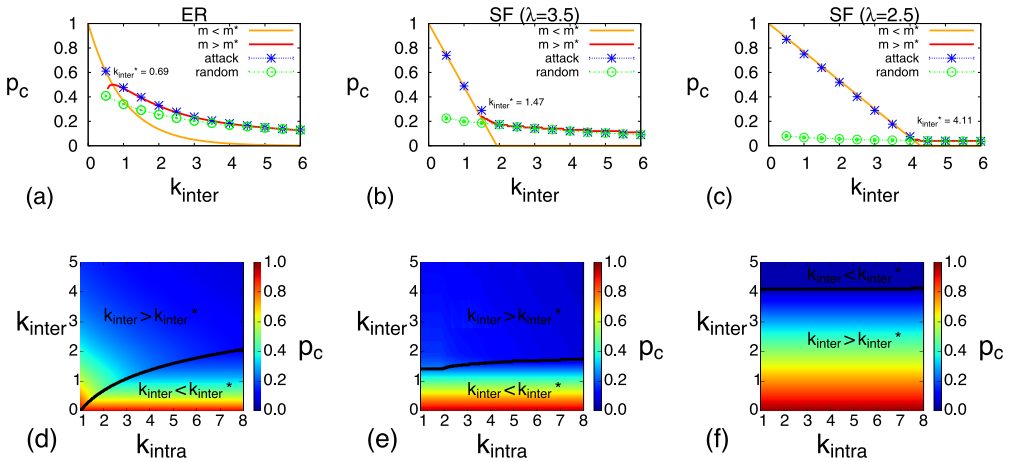

FIG. 5. (Color online) Critical concentration of interconnected nodes [(a)–(c)] p_c as a function of k_{inter} calculated for (a) ER and [(b) and (c)] SF networks with $m = 10$ and $k_{\text{intra}} = 2$. Simulation points obtained from at least 1000 simulation runs of networks of size $N = 600\,000$. Solid lines represent the analytical result obtained in Eqs. (19), (25)–(27). [(d)–(f)] Two-parameter (k_{intra}, k_{inter}) phase diagram for a fixed number of modules $m = 10$. The black line corresponds to the critical mean interdegree k^*_{inter} at which the two p_c solutions cross.

uncovered differences in the modular structure of semantic networks of high versus low creative individuals, suggesting a new perspective on the effect of modularity on memory and language [45,46]. Thus, our findings of the two different regimes may shed further light on the role modularity plays in neurocognitive processes.

ACKNOWLEDGMENTS

S.H. and D.Y.K. thank Defense Threat Reduction Agency, Office of Naval Research, Binational Science Foundation, the LINC (Grant No. 289447) and the Multiplex (Grant No. 317532) EU projects, the DFG, and the Israel Science Foundation for support. D.Y.K., Y.N.K., and M.F. thank BSF (Grant No. BSF2013106) for financial support. Y.N.K. and M.F. also acknowledge support from the I-CORE Program of the Planning and Budgeting Committee and the Israel Science Foundation (Grant No. 51/11). S.S. thanks the James S. McDonnell Foundation 21st Century Science Initiative—Complex Systems Scholar Award (Grant No. 220020315) and the Scottish Informatics and Computer Science Alliance for financial support.

[1] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
[2] G. Caldarelli, *Scale-Free Networks: Complex Webs in Nature and Technology*, Oxford Finance (Oxford University Press, Oxford, 2007).
[3] R. Cohen and S. Havlin, *Complex Networks: Structure, Robustness and Function* (Cambridge University Press, Cambridge, 2010).
[4] M. Newman, *Networks: An Introduction* (Oxford University Press, New York, 2010).
[5] M. Girvan and M. E. J. Newman, Proc. Natl. Acad. Sci. USA **99**, 7821 (2002).
[6] K. A. Eriksen, I. Simonsen, S. Maslov, and K. Sneppen, Phys. Rev. Lett. **90**, 148701 (2003).
[7] R. Guimerà, S. Mossa, A. Turtschi, and L. A. N. Amaral, Proc. Natl. Acad. Sci. USA **102**, 7794 (2005).
[8] V. Spirin and L. A. Mirny, Proc. Natl. Acad. Sci. USA **100**, 12123 (2003).
[9] L. M. A. Bettencourt, J. Lobo, D. Helbing, C. Kühnert, and G. B. West, Proc. Natl. Acad. Sci. USA **104**, 7301 (2007).
[10] D. S. Bassett and M. S. Gazzaniga, Trends Cogn. Sci. **15**, 200 (2011).
[11] E. Bullmore and O. Sporns, Nat. Rev. Neurosci. **13**, 336 (2012).
[12] L. K. Gallos, H. A. Makse, and M. Sigman, Proc. Natl. Acad. Sci. USA **109**, 2825 (2012).
[13] S. Fortunato, Phys. Rep. **486**, 75 (2010).
[14] M. A. Porter, J.-P. Onnela, and P. J. Mucha, Not. Am. Math. Soc. **56**, 1082 (2009).
[15] E. A. Variano, J. H. McCoy, and H. Lipson, Phys. Rev. Lett. **92**, 188701 (2004).
[16] N. Kashtan and U. Alon, Proc. Natl. Acad. Sci. USA **102**, 13773 (2005).
[17] C. D. Brummitt, R. M. D'Souza, and E. A. Leicht, Proc. Natl. Acad. Sci. USA **109**, E680 (2012).
[18] J.-j. Wu, Z.-y. Gao, and H.-j. Sun, Phys. Rev. E **74**, 066111 (2006).
[19] A. Nematzadeh, E. Ferrara, A. Flammini, and Y.-Y. Ahn, Phys. Rev. Lett. **113**, 088701 (2014).
[20] J. P. Gleeson, Phys. Rev. E **77**, 046117 (2008).
[21] S. Melnik, M. A. Porter, P. J. Mucha, and J. P. Gleeson, Chaos **24**, 023106 (2014).
[22] M. E. J. Newman, Phys. Rev. E **67**, 026126 (2003).
[23] J. P. Bagrow, S. Lehmann, and Y.-Y. Ahn, Network Sci. **FirstView**, 1 (2015).
[24] J.-D. J. Han, N. Bertin, T. Hao, D. S. Goldberg, G. F. Berriz, L. V. Zhang, D. Dupuy, A. J. M. Walhout, M. E. Cusick, F. P. Roth, and M. Vidal, Nature **430**, 88 (2004).
[25] O. Sporns, C. J. Honey, and R. Kötter, PLoS ONE **2**, e1049 (2007).
[26] Y. He, J. Wang, L. Wang, Z. J. Chen, C. Yan, H. Yang, H. Tang, C. Zhu, Q. Gong, Y. Zang, and A. C. Evans, PLoS ONE **4**, e5226 (2009).
[27] B. R. da Cunha, J. C. González-Avella, and S. Gonçalves, PLoS ONE **10**, e0142824 (2015).
[28] N. Masuda, New J. Phys. **11**, 123018 (2009).
[29] L. Hébert-Dufresne, A. Allard, J.-G. Young, and L. J. Dubé, Sci. Rep. **3**, 2171 (2013).
[30] M. Salathé and J. H. Jones, PLoS Comput. Biol. **6**, e1000736 (2010).
[31] C. J. Stam, Nature Rev. Neurosci. **15**, 683 (2014).
[32] E. C. van Straaten and C. J. Stam, Eur. Neuropsychopharmacol. **23**, 7 (2013).
[33] A. Bashan, R. P. Bartsch, J. W. Kantelhardt, S. Havlin, and P. C. Ivanov, Nat. Commun. **3**, 702 (2012).
[34] S. Buldyrev, R. Parshani, G. Paul, H. Stanley, and S. Havlin, Nature **464**, 1025 (2010).
[35] J. Gao, S. V. Buldyrev, H. E. Stanley, and S. Havlin, Nat. Phys. **8**, 40 (2012).
[36] E. A. Leicht and R. M. D'Souza, arXiv:0907.0894 (2009).
[37] M. Kivelä, A. Arenas, M. Barthélemy, J. P. Gleeson, Y. Moreno, and M. A. Porter, J. Complex Networks **2**, 203 (2014).
[38] M. Bastian, S. Heymann, and M. Jacomy, in *Third International AAAI Conference on Weblogs and Social Media* (ICWSM'09, San Jose, California, 2009), http://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154.
[39] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. E **64**, 026118 (2001).
[40] P. Erdős and A. Rényi, Publ. Math. Inst. Hung. Acad. Sci **5**, 17 (1960).
[41] B. Bollobás, *Random Graphs*, 2nd ed. (Cambridge University Press, Cambridge, 2001).
[42] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[43] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. Lett. **85**, 5468 (2000).
[44] J. Gao, S. V. Buldyrev, H. E. Stanley, X. Xu, and S. Havlin, Phys. Rev. E **88**, 062816 (2013).
[45] S. B. Kaufman, "The controlled chaos of creativity", retrieved from http://blogs.scientificamerican.com/beautiful-minds/2014/06/25/the-controlled-chaos-of-creativity (2014).
[46] M. Faust and Y. N. Kenett, Front. Human Neurosci. **8**, 511 (2014).