Spectral analysis and the dynamic response of complex networks

M. A. M. de Aguiar[1,2] and Y. Bar-Yam[1]
[1]*New England Complex Systems Institute, Cambridge, Massachusetts 02138, USA*
[2]*Instituto de Física Gleb Wataghin, Universidade Estadual de Campinas, 13083-970 Campinas, São Paulo, Brazil*
(Received 18 June 2004; revised manuscript received 3 September 2004; published 4 January 2005)

The eigenvalues and eigenvectors of the connectivity matrix of complex networks contain information about its topology and its collective behavior. In particular, the spectral density $\rho(\lambda)$ of this matrix reveals important network characteristics: random networks follow Wigner's semicircular law whereas scale-free networks exhibit a triangular distribution. In this paper we show that the spectral density of hierarchical networks follows a very different pattern, which can be used as a fingerprint of modularity. Of particular importance is the value $\rho(0)$, related to the homeostatic response of the network: it is maximum for random and scale-free networks but very small for hierarchical modular networks. It is also large for an actual biological protein-protein interaction network, demonstrating that the current leading model for such networks is not adequate.

DOI: 10.1103/PhysRevE.71.016106 PACS number(s): 89.75.Hc, 87.10.+e, 87.23.−n

The network concept has been gaining recognition as a fundamental tool in both biological and social sciences, where the theory of complex systems finds fertile ground. Biological examples include food webs in ecology [1], nervous systems [2], cellular metabolism [3], protein conformation [4], and a protein-protein interaction network [5]. Social networks include scientific collaboration, citation, problem solving, and linguistic networks [6]. Most biological and social networks studied are not randomly connected; they follow a *scale-free* behavior (see [7] and references therein). In random networks the probability that a node has k connections, $P(k)$, is Poisson distributed and, therefore, every node has about the same number of connections. In scale-free networks $P(k)$ follows a power law, a property that can be constructed by sequential *preferential attachment* of nodes, where new nodes are more likely to connect to already highly connected ones. The properties of such networks are often characterized by the presence of a few highly connected nodes, the *hubs*, whereas most of the remaining nodes have a small number of connections. The importance of such networks, originally couched in terms of robustness of static connectivity to failure despite sensitivity to attack [8], may perhaps be better characterized in terms of their response dynamics, that provides both robustness *and* sensitivity [9].

Although scale-free networks describe several statistical properties of biological networks, they fail to take into account one important aspect: namely, the modularity exhibited by most complex systems [10–12]. The concept of modularity assumes that the full network of interactions can be partitioned into a number of subnetworks or modules. Each module is composed of several elements which are more interconnected than they are connected to the rest of the network. Modular systems may be organized in a structural hierarchy, with multiple levels of modular decomposition. Molecules, organelles, cells, tissues, organs and organisms, families, communities, etc., are an example of such a hierarchy of structures. Networks incorporating both modular hierarchy and scale-free character were recently discussed by Ravasz *et al.* [13] (see also [14]). One property often used to characterize modular networks is their clustering coefficient—the degree to which neighbors of a node are connected to each other—which is larger than that of generic scale-free models.

In this work we investigate the spectral properties of modular networks. We show that the density of states of the connectivity matrix (particularly its randomized version where elements are set to ± 1) provides a connection between the structure and the dynamic response of a network. This enables us to distinguish between various models and actual systems in a manner that may be directly relevant to considering the behavior of system response to perturbations. In particular, we are able to distinguish clearly between random, scale-free, and modular networks. However, none of these standard model networks capture the properties of an actual protein-protein interaction network.

The connectivity (or adjacency) matrix A represents the topology of the system, indicating which variables are interconnected. It is defined as $A_{ij}=1$ if nodes i and j are connected and zero otherwise. If we consider the network as an influence network, where each link may have a strength and phase that is not specified, a model of the interactions between nodes A_R can be constructed from A by changing each of the entries 1 of A into -1 with 50% probability (keeping $A_{ij}=A_{ji}$, since they represent the same connection). The spectral properties of A_R contain information about the dynamics of the network. If the network is in equilibrium and a perturbation is introduced, this perturbation propagates through the nodes according to A_R. In a linear approximation the state of the nodes are updated according to $x_i^{t+1}=\Sigma_j A_{Rij}x_j^t$. Below we study the spectral properties of A and A_R and show that they are in many cases similar or otherwise can be related.

The smoothed density of states of the network is defined by

$$\overline{\rho_\epsilon}(\overline{\lambda}) = \frac{1}{N}\sum_i \delta_\epsilon(\overline{\lambda} - \lambda_i), \tag{1}$$

where λ_i are the eigenvalues of the connectivity matrix and N is the total number of nodes. Since A and A_R are symmet-

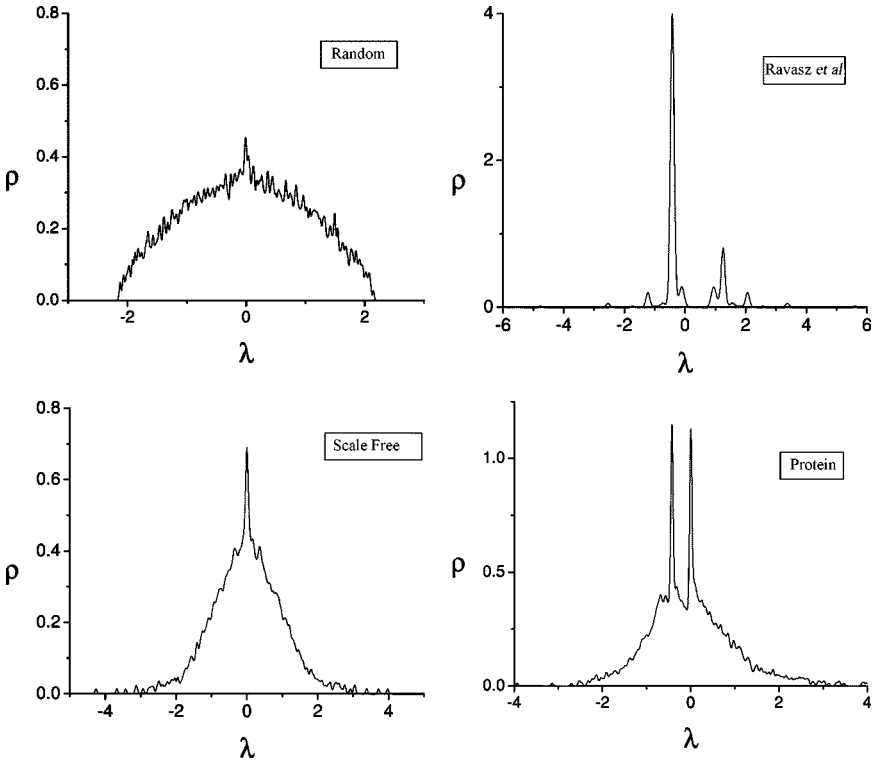

FIG. 1. Smoothed density of states for a random, scale-free, hierarchical network of Ravasz *et al.* (all with 1024 nodes) and the protein-protein interaction network (with 1297 nodes).

ric, all eigenvalues are real. $\delta_\epsilon(x)$ is a smoothed delta function that tends to the real Dirac delta as $\epsilon \to 0$. Choosing ϵ to be a few units of the mean level spacing produces a smooth level density even for small networks, which is easier to visualize than the spiked density produced by the delta functions. Following Farkas *et al.* [15] we define scaled variables ρ and λ by

$$\lambda = \overline{\lambda}/\sqrt{Np(1-p)}, \quad \rho = \overline{\rho}\sqrt{Np(1-p)}, \tag{2}$$

where $p = \overline{k}/N$ is the average number of links per node divided by the total number of nodes. For random networks the density of states can be computed analytically from random matrix theory and the result is the so-called Wigner's semicircular law. In the scaled variables it becomes simply $\rho(\lambda) = \sqrt{4-\lambda^2}/2\pi$ if $|\lambda| < 2$ and zero otherwise.

Figure 1 shows the density of states for four different networks. All networks have $N=1024$ nodes, except for the protein-protein network which has $N=1297$. Figure 1(a) shows $\rho(\lambda)$ for a random network with $p=0.0057$, following closely Wigner's semicircular law. Figure 1(b) shows a scale-free network with $p=0.0058$, exhibiting a triangular profile [15,16]. Figure 1(c), showing the results for the hierarchical network of Ravasz *et al.* [13], has a peculiar density, which we shall discuss in more detail. Finally Fig. 1(d) shows $\rho(\lambda)$ for a protein-protein interaction network [5] and also has a distinct behavior, looking more like a superposition of two independent scale-free networks.

Figure 2 shows the density of states for the same networks obtained with the randomized connectivity matrices A_R. For each network we diagonalized 20 matrices with random distributions of ± 1's and calculated the average density over this ensemble. The averaged density satisfies $\rho(\lambda) = \rho(-\lambda)$. The scale-free and random networks are not sensitive to sign randomization, since their original spectra are already symmetric. The hierarchical network density of states of Ravasz *et al.*, on the other hand, changes considerably. It keeps the minimum at $\lambda = 0$, whereas all other networks have a peak there. Also, the density has sharp peaks with high intensity at certain values of $|\lambda|$, becoming very small away from the peaks. Interestingly, this type of spectrum has also been observed in the context of percolation theory for *random* three-dimensional networks near the percolation threshold [17]. The biological network also has an interesting structure, deviating from the pure scale-free case. However, in contrast to the network of Ravasz *et al.*, it has a peak at $\lambda = 0$.

The hierarchical network of Ravasz *et al.* is built from a fully connected network with four nodes. This unit is then replicated 3 times and the four identical networks are connected together. The network thus formed is then viewed as the new unit, and the replicating and connecting process is repeated [13]. Although the exact repetition of this process is artificial, one expects real modular networks to exhibit some type of self-similar structure. In what follows we shall show that networks built from such basic units have indeed a very characteristic spectrum, that can be used to identify its modular nature.

Consider first a fully connected network with N nodes. The connectivity matrix is $(A_N)_{ij} = 1 - \delta_{ij}$. The eigenvalues of A_N can be calculated immediately and we find $\lambda_1 = N-1$, $\lambda_2 = \lambda_3 = \cdots = \lambda_N = -1$. The first eigenvector $|w_1\rangle$, corresponding to the largest eigenvalue λ_1, has components $w_{1,i} = 1$. All the other eigenvectors are degenerate and satisfy $\Sigma_i w_{j,i} = 0$. It is possible to choose them so as to have very few nonzero elements. The linear update equation $x^{t+1} = A_N x^t$ decouples

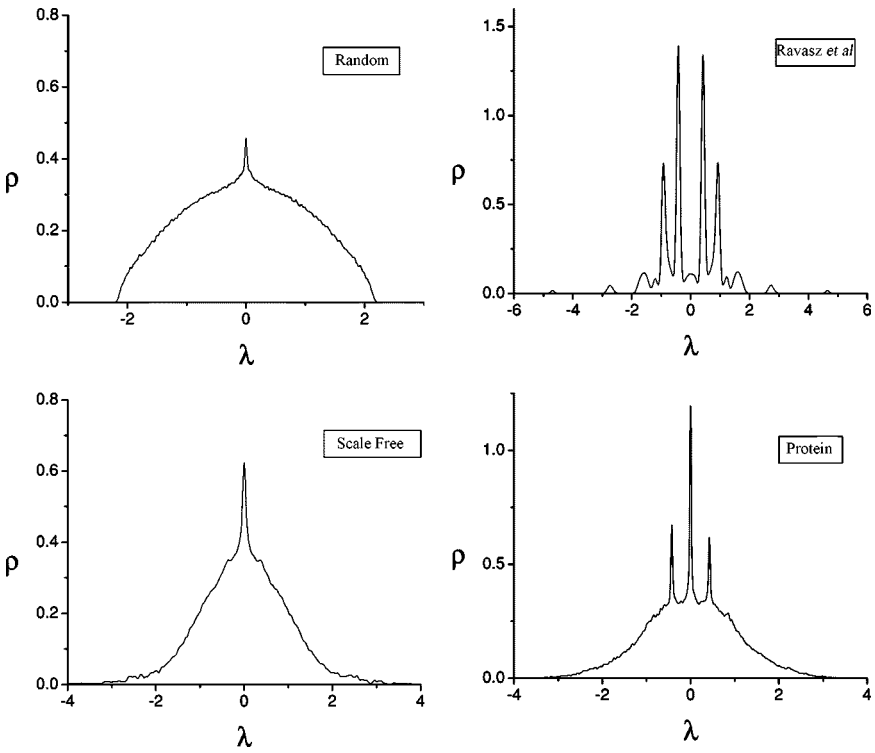

FIG. 2. Smoothed density of states for the *randomized* networks of Fig. 1.

into $y_i^{t+1}=\lambda_i y_i^t$ and $y_i^t=\lambda_i^t y_i^0$. The dominant mode is the "center of mass" y_1, meaning that the network synchronizes and responds as a unit to the perturbation. All other modes involve fewer nodes and correspond to oscillations of fixed amplitude. The density of states for a fully connected network has only two peaks: one at $\lambda=-1$ and the other at $\lambda=N-1$, the former being $N-1$ times larger than the latter.

Next we consider *star networks*, which are characterized by a single hub—i.e., a single central node to which all other nodes are connected. Star networks emerge in systems in which preferential attachment is superlinear, meaning that the probability that a new node attaches to old nodes increases faster than expected by linear preferential attachment [18]. Starlike clusters are very common in biological networks (see, for instance, [19]) and their eigenvalues and eigenvectors can also be computed exactly. In the idealized star network the nodes connect only to the central node, which we label 1. The connectivity matrix is given by $A_{i1}=A_{1i}=1$ for $i=2,3,\ldots,N$ and $A_{ij}=0$ otherwise. The eigenvalues are $\lambda_1=\sqrt{N-1}$, $\lambda_2=\lambda_3=\cdots=\lambda_{N-1}=0$ and $\lambda_N=-\sqrt{N-1}$. Starlike structures are known sources of zero eigenvalues [16]. The first eigenvector $|w_1\rangle$ has components $w_{11}=\sqrt{N-1}$ and $w_{1i}=1$ for $i\geqslant 2$. The last eigenvector $|w_N\rangle$ is given by $w_{N1}=\sqrt{N-1}$ and $w_{Ni}=-1$ for $i\geqslant 2$. All the other degenerate eigenvectors satisfy $w_{j,1}=0$ and $\Sigma_{i=2}^N w_{j,i}=0$.

Now we consider a network whose connectivity matrix has a modular organization consisting of four main blocks, each one very similar to the others. The number 4 is chosen only for comparison with the model of Ravasz *et al.*, but could be any number. We assume that the blocks are fully connected, so that we know their eigenvectors and eigenvalues when they are decoupled. Let $|w_i^\alpha\rangle$ be the ith eigenvector of the block labeled by α. Since the blocks are all identical, the eigenvalues are degenerate: $\lambda_1^\alpha=M-1$ and $\lambda_i^\alpha=-1$ for $i\neq 1$, where M is the dimension of the blocks. The connectivity matrix can represented in block form by

$$A=\begin{pmatrix} A_M & v_{12} & v_{13} & v_{14} \\ v_{12}^T & A_M & v_{23} & v_{24} \\ v_{13}^T & v_{23}^T & A_M & v_{34} \\ v_{14}^T & v_{24}^T & v_{34}^T & A_M \end{pmatrix} \equiv A^0+V, \tag{3}$$

where A_M are fully connected $M\times M$ matrices, A^0 is the *unperturbed* matrix, with the four uncoupled A_M blocks, and V is a sparse perturbation, representing the weak connection between nodes of different blocks.

The perturbation breaks the degeneracy between the blocks. The first eigenvalue becomes $\lambda=\lambda_0+\mu$ and the corresponding eigenvector $|v_1^\alpha\rangle=\Sigma_\beta a_{\alpha\beta}|w_1^\beta\rangle+|\xi\rangle$ where the sum over β runs over the blocks and represents the linear combination between the originally degenerate vectors and the last term is the correction due to the perturbation. Writing the eigenvalue equation for $|v_1^\alpha\rangle$ and keeping only linear terms in the perturbation V leads to the condition

$$\sum_\beta a_{\alpha\beta}[\langle w_1^\alpha|V|w_1^\beta\rangle-\mu\delta_{\alpha\beta}]=0. \tag{4}$$

For all the other eigenvectors, whose degeneracy is much bigger, we write $|v_n^\alpha\rangle=\Sigma_{\beta m}a_{\alpha\beta}^{nm}|w_m^\beta\rangle+|\xi\rangle$ where the sum now runs over β and m, with $n,\ m\neq 1$. The eigenvalue equation for this case is

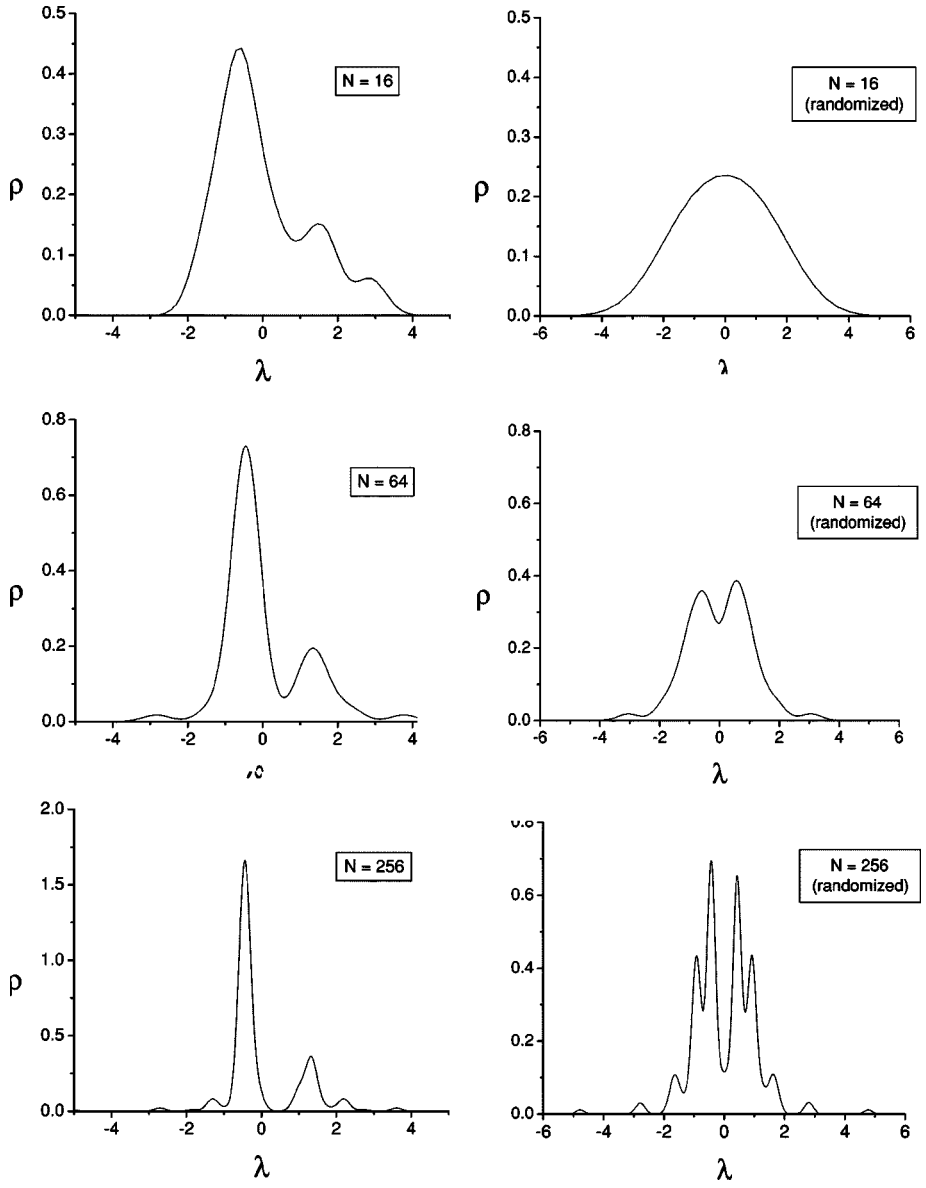

FIG. 3. Smoothed density of states for the hierarchical network of Ravasz *et al.* for the original and the randomized connectivity matrix with $N=16$, 64, and 256 nodes. The smoothing parameter is 5 times the average level spacing, except for $N=16$, where it is 2 times the average level spacing.

$$\sum_{\beta m} a_{\alpha\beta}^{nm}[\langle w_n^\alpha|V|w_m^\beta\rangle - \mu\delta_{\alpha\beta}\delta_{nm}] = 0. \tag{5}$$

$$\langle w_1^\alpha|V|w_1^\beta\rangle = \sum_{k,l} w_{1,k}^\alpha V_{kl} w_{1,l}^\beta \equiv K_{\alpha\beta}, \tag{6}$$

However, each matrix element $\langle w_n^\alpha|V|w_m^\beta\rangle$ is obtained by adding elements of the matrix V with coefficients that add up to zero. Since V is sparse, we expect most of these elements to be zero and, when they are not zero, there will likely be cancellations. Therefore, the corrections to the eigenvalues are going to be small, and the density of states of A should still have a large peak around $\lambda=-1$.

On the other hand, the elements of $|w_1^\beta\rangle$ are all 1 inside the β block and zero outside:

where $K_{\alpha\beta}$ is the number of 1's in the block $v_{\alpha\beta}$. At this point we have to distinguish between random and scale-free networks:

Random coupling. We can assume that all the coupling blocks $v_{\alpha\beta}$ are similar, so we write $K_{\alpha\beta}=a$ where a is the average number of 1's in each of the v blocks. The 4×4 matrix to be diagonalized in Eq. (4) is identical to the connectivity matrix of a completely connected network of four nodes. Therefore, the four uncoupled eigenvalues $M-1$ unfold into one eigenvalue $M-1+3a$ and three eigenvalues

$M-1-a$. For random coupling we expect three main peaks in the density of states: a large peak at $\lambda=-1$, a smaller one at $M-1-a$ and an even smaller one at $M-1+3a$.

Scale-free coupling. In this case the blocks are themselves not connected randomly, they attach preferentially to, say, the first block. The 4×4 matrix to be diagonalized has the form

$$\begin{pmatrix} 0 & a & a & a \\ a & 0 & b & b \\ a & b & 0 & b \\ a & b & b & 0 \end{pmatrix}, \tag{7}$$

where $a \gg b$. In a first approximation we neglect b and the resulting matrix is that of a 4×4 star network. Therefore, the eigenvalues become $M-1-\sqrt{3}a$, $M-1$ (doubly degenerate) and $M-1+\sqrt{3}a$. Together they contribute a single symmetric peak around $M-1$ with half width $\sqrt{3}a$. Therefore, for scale-free modular matrices we expect only two main peaks in the density of states: a large one at $\lambda=-1$ and a smaller one at $\lambda=M-1$.

Figure 3 shows the density of states for the hierarchical network of Ravasz *et al.* with 16, 64, and 256 nodes. The two peaks structure is clear and consistent with our analysis of a modular scale-free network. The protein network shown in Fig. 1 is certainly not completely modular. But it is also not generically scale-free either. The two peaks at zero and -1 (in nonscaled units) suggest the existence of many starlike structures (where the eigenvalue 0 abounds) and many fully connected modules (where the eigenvalue -1 abounds).

Randomized connectivity matrices. A similar analysis can made for the case of the randomized connectivity matrices. For example, starting from a single fully connected unit of four nodes, the eigenvalue equation can be seen to be $\lambda^4-6\lambda^2-2\lambda(a_{23}a_{24}a_{34}+a_{12}a_{24}a_{14}+a_{12}a_{13}a_{23}+a_{13}a_{14}a_{34})-2(a_{12}a_{13}a_{24}a_{34}+a_{12}a_{14}a_{23}a_{34}+a_{14}a_{13}a_{24}a_{23})+3=0$. For random a_{ij}'s, the term multiplying λ averages to zero, whereas the constant term in parentheses averages to either -1 or $+1$. The averaged equation is $\lambda^4-6\lambda^2+1=0$ or $\lambda^4-6\lambda^2+5=0$. The result is a spectrum with two pairs of symmetric eigenvalues. When a modular network is constructed out of these random units, we obtain a density of states with four symmetric peaks. This can be seen in Fig. 3 for the network of Ravasz *et al.* with 16, 64, and 256 nodes.

To summarize, we have applied linear algebra and perturbation analysis to the spectral analysis of modular networks. We have shown that the density of states contains crucial information not only about the topology of the network but also about its response to external perturbations. By comparing $\rho(\lambda)$ for a random, a scale-free and the hierarchical network of Ravasz *et al.*, we have shown that it exhibits clear fingerprints of the networks they represent. More importantly, we have shown that neither of these model networks can describe the density of states of a real protein-protein interaction network, showing that better network models are necessary to understand biological systems. In particular, the behavior of $\rho(0)$, which indicates that the real biological network has a robust homeostatic response, is not reproduced by the hierarchical model of Ravasz *et al.* Our analysis also indicates the presence of several starlike and fully connected modules in the biological network, suggesting that these structures might have to be incorporated explicitly in more realistic models.

[1] J. Cohen, F. Briand, and C. Newman, *Community Food Webs: Data and Theory* (Springer, Berlin, 1990); N. D. Martinez, Science **260**, 242 (1993); R. J. Williams and N. D. Martinez, Nature (London) **404**, 180 (2000); Jose M. Montoya and Ricard V. Sole, J. Theor. Biol. **214**, 405 (2002).

[2] C. Koch and G. Laurent, Science **284**, 79 (1999).

[3] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A.-L. Barabási, Nature (London) **407**, 651 (2000).

[4] A. Scala, L. A. N. Amaral, and M. Barthélémy, Europhys. Lett. **55**, 594 (2000).

[5] I. Xenarios *et al.*, Nucleic Acids Res. **29**, 239 (2001).

[6] D. J. de S. Price, Science **149**, 510 (1965); M. E. J. Newman, Proc. Natl. Acad. Sci. U.S.A. **98**, 404 (2001); Phys. Rev. E **64**, 016131 (2001); **64**, 016132 (2001); Ramon Ferrer and Ricard V. Sole, *ibid.* **69**, 051915 (2004); D. Braha and Y. Bar-Yam, *ibid.* **69**, 016113 (2004).

[7] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).

[8] R. Albert, H. Jeong, and A.-L. Barabasi, Nature (London) **406**, 378 (2000).

[9] Y. Bar-Yam and I. Epstein, Proc. Natl. Acad. Sci. U.S.A. **101**, 4341 (2004).

[10] H. A. Simon, *The Sciences of the Artificial*, 3rd ed. (MIT Press, Cambridge, MA, 1998), Chap. 8.

[11] Y. Bar-Yam, *Dynamics of Complex Systems* (Perseus Press, Cambridge, MA, 1997), Chap. 2.

[12] L. H. Hartwell, J. J. Hopfield, S. Leibler, and A. W. Murray, Nature (London) **402**, C47 (1999).

[13] E. Ravasz, A. L. Somera, D. A. Mongru, Z. N. Oltvai, and A.-L. Barabási, Science **297**, 1551 (2002).

[14] A.-L. Barabási and Z. N. Oltvai, Nat. Rev. Genet. **5**, 101 (2004); A. Fronczak, P. Fronczak, and J. A. Holyst, Phys. Rev. E **68**, 046126 (2003); J. D. Noh, *ibid.* **67**, 045103 (2003).

[15] I. J. Farkas, I. Derényi, A.-L. Barabási, and T. Vicsek, Phys. Rev. E **64**, 026704 (2001).

[16] S. N. Dorogovtsev, A. V. Goltsev, J. F. F. Mendes, and A. N. Samukhin, Phys. Rev. E **68**, 046109 (2003), cond-mat/0306340.

[17] R. Berkovits and Yshai Avishai, Phys. Rev. B **53**, R16 125 (1996).

[18] P. L. Krapivsky, S. Redner, and F. Leyvraz, Phys. Rev. Lett. **85**, 4629 (2000).

[19] S. Valverde, R. F. Cancho, and R. V. Solé, Europhys. Lett. **60**, 512 (2002).