IRis

Northeastern University

Physics Faculty Publications

Department of Physics

April 01, 2003

# Topology and correlations in structured scale-free networks

A Vazquez

M Boguna

Y Moreno

R Pastor-Satorras

A Vespignani
*Northeastern University*

Recommended Citation

Vazquez, A; Boguna, M; Moreno, Y; Pastor-Satorras, R; and Vespignani, A, "Topology and correlations in structured scale-free networks" (2003). *Physics Faculty Publications.* Paper 191. http://hdl.handle.net/2047/d20002151

This work is available open access, hosted by Northeastern University.

# Topology and correlations in structured scale-free networks

Alexei Vázquez,[1] Marián Boguñá,[2] Yamir Moreno,[3] Romualdo Pastor-Satorras,[4] and Alessandro Vespignani[5]
[1]*International School for Advanced Studies and INFM, Via Beirut 4, Trieste I-34014, Italy*
[2]*Departament de Física Fonamental, Universitat de Barcelona, Av. Diagonal 647, 08028 Barcelona, Spain*
[3]*The Abdus Salam International Centre for Theoretical Physics, P.O. Box 586, Trieste I-34014, Italy*
[4]*Departament de Física i Enginyeria Nuclear, Universitat Politècnica de Catalunya, Campus Nord, 08034 Barcelona, Spain*
[5]*Laboratoire de Physique Théorique (UMR du CNRS 8627), Bâtiment 210, Université de Paris–Sud, 91405 Orsay Cedex, France*
(Received 7 September 2002; published 21 April 2003)

We study a recently introduced class of scale-free networks showing a high clustering coefficient and nontrivial connectivity correlations. We find that the connectivity probability distribution strongly depends on the fine details of the model. We solve exactly the case of low average connectivity, providing also exact expressions for the clustering and degree correlation functions. The model also exhibits a lack of small-world properties in the whole parameter range. We discuss the physical properties of these networks in the light of the present detailed analysis.

DOI: 10.1103/PhysRevE.67.046111 PACS number(s): 89.75.−k, 87.23.Ge, 05.70.Ln

## I. INTRODUCTION

Recently, a major scientific effort has been devoted to the characterization and modeling of a wide range of social and natural systems that can be described as networks [1,2]. Systems such as the Internet [3–6] or the World Wide Web [7], social communities [8], food webs [9], and biological interacting networks [10–13] can be represented as a graph [14], in which nodes represent the population individuals and links the physical interactions among them. Strikingly, many of these networks have complex topological properties and dynamical features that cannot be accounted for by classical graph modeling [15]. In particular, small-world properties [16] and scale-free degree distributions [17] (where the degree or connectivity of a node is defined as the number of other nodes to which it is attached) seem to emerge frequently as dominant features governing the topology of real-world networks. These global properties imply a large connectivity heterogeneity and a short average distance between nodes, which have considerable impact on the behavior of physical processes taking place on top of the network. For instance, scale-free (SF) networks have been shown to be resilient to random damage (absence of a percolation transition) [18–20] and prone to epidemic spreading (null epidemic threshold) [21–24].

The detailed scrutiny of the topological properties of networks has pointed out that small-world and scale-free properties come often along with nontrivial degree correlations and clustering properties. Recently, an interesting class of networks has been introduced by Klemm and Eguíluz by proposing a growing model in which nodes are progressively deactivated with a probability inversely proportional to their connectivity [25]. Analytical arguments and numerical simulations have lead to the claim that, under general conditions, the deactivation model, allowing a core of $m$ active nodes, generates a network with average degree $\langle k\rangle = 2m$ and degree probability distribution $P(k) = 2m^2k^{-3}$. Interestingly, the scale-free properties are associated to a high clustering coefficient. For this reason the deactivation model has been used to study how clustering can alter the picture obtained for the resilience to damage and epidemic spreading in SF networks [26,27].

In this paper, we revisit the analysis of the deactivation model. We find an analytical solution in the case of minimal values of active nodes $m$ (low average connectivity). In addition, large-scale numerical simulations exhibit a noticeable variability of the degree distribution with $m$. In particular, the degree exponent strongly depends on $m$ for the general case considered in Ref. [25]. The model topology is also susceptible to several details of the construction algorithm. By means of large-scale numerical simulations we study the deactivation model topology in the whole range of $m$ and for different algorithm parameters. We calculate analytically the clustering coefficient and connectivity correlation functions. Also in this case a variability with respect to the model parameters is found. Extensive numerical simulations confirm the analytical picture presented here.

In the generated networks, we also report the lack of small-world properties. In the whole parameter range, we find a network diameter increasing linearly with the number of nodes forming the network [28]. The networks' topology is similar to a chain of dense clusters locally connected. The networks are thus similar to a one-dimensional lattice in what concerns their physical properties. In particular, diffusion and spreading processes might be heavily affected by the increasing average distance among nodes that make the system similar to a one-dimensional chain. In this perspective, we discuss the properties of epidemic spreading and resilience to damage in networks generated with the deactivation model.

## II. DEACTIVATION MODEL

The deactivation model introduced by Klemm and Eguíluz [25] is defined as follows: Consider a network with directed links. Each node can be in two states, either active or inactive. The model starts from a completely connected graph of $m$ active nodes and proceeds by adding new nodes one by one. Each time a node is added, (1) it is connected to all active nodes in the network; (2) one of the active nodes is

 ©2003 The American Physical Society

selected and set inactive with probability

$$p_d(k_i^{\rm in})=\frac{\left[\sum_{j\in\mathcal{A}}(a+k_j^{\rm in})^{-1}\right]^{-1}}{a+k_i^{\rm in}}; \tag{1}$$

and (3) the new node is set active. The sum in Eq. (1) runs over the set of active nodes $\mathcal{A}$, $a$ is a model parameter, and $k_i^{\rm in}$ denotes the in-degree of the $i$th node.

As we shall show below, this model is quite sensitive to the order in which steps (2) and (3) are performed and, therefore, it is better to discriminate the following cases: Model *A*, step (2) is performed *before* step (3), and model *B*, step (2) is performed *after* step (3). For $m\to\infty$, both models can be solved analytically in the continuous $k^{\rm in}$ approximation, after introducing the probability density that an active node has in-degree $k^{\rm in}$ [25]. Moreover in this limit, the order of steps 2 and 3 is irrelevant, obtaining the same in-degree distribution $P(k^{\rm in})\sim(a+k^{\rm in})^{-\gamma}$ with

$$\gamma=2+\frac{a}{m}. \tag{2}$$

The model is usually simulated by using $a=m$. In this way the deactivation probability is inversely proportional to the total connectivity of the nodes $(m+k^{\rm in})^{-1}$ and the connectivity distribution results to be $P(k)=2m^2k^{-3}$. Interestingly, due to the deactivation mechanism, the networks show a high clustering coefficient that approaches a constant value in the infinite size limit [25].

At lower values of $m$, it has been claimed that finite size effects set in and the connectivity distribution shows deviations from the predicted behavior. We shall see in the following section that for $a=m\leqslant 10$ the model presents a very different analytical solution that yields a connectivity distribution very far from the $m\to\infty$ limit. In addition, the deactivation model topology is very sensible to changes in the details of the growing algorithms.

## III. DEGREE DISTRIBUTION

### A. Model *A*

Let us first focus on model *A* with $m=2$, i.e., the smallest value of $m$ for which the model is nontrivial. In this case, after adding a new node we have only two nodes at the deactivation step. One of them will be set inactive and replaced by the new added node that has in-degree 0. In the worst case, the other node will have in-degree $0+1$, the 0 coming from its initial in-degree and the 1 from the connection to the newly added node, and in general, it will have in-degree larger than or equal to 2. Later on, at the next deactivation step, the in-degrees of both nodes will have increased by one resulting in one active node with in-degree 1 and another with in-degree $K\geqslant 2$, where $K$ is the in-degree of the active node with largest in-degree that coincides with the oldest node. Then, following Eq. (1), one of them will be deactivated with probability

$$p_d(K)=\frac{1+a}{1+2a+K},\quad p_d(1)=1-p_d(K). \tag{3}$$

Each time the oldest node is not deactivated, its in-degree increases by one and, therefore, the probability that the oldest node has in-degree $K$ is just the probability that it is not deactivated in $K-2$ steps, with running in-degree $2,3,\ldots,K-1$. Thus, the probability $\tilde{P}(K)$ of creating a deactivated node of in-degree $K$ is equal to the probability that the largest node is not deactivated in $K-2$ steps and is deactivated in the last step, i.e.,

$$\begin{aligned}\tilde{P}(K)&=\prod_{\ell=2}^{K-1}[1-p_d(\ell)]p_d(K)\\&=\frac{\Gamma(3+2a)}{\Gamma(1+a)}\frac{\Gamma(a+K)}{\Gamma(2+2a+K)},\end{aligned} \tag{4}$$

where $\Gamma(x)$ is the standard gamma function [29]. On the other hand, every time that the oldest node is not deactivated, the other, with in-degree 1, is deactivated. Hence, in the $K-1$ deactivation steps leading to the generation of a node with in-degree $K$, $K-2$ nodes with in-degree 1 are created. The average number of nodes with in-degree 1 created in the process is then

$$\tilde{P}_1=\sum_{K=2}^{\infty}(K-2)\tilde{P}(K)=\frac{2+a}{a}. \tag{5}$$

Therefore, the in-degree distribution will be given by

$$P(k^{\rm in})=C^{-1}\begin{cases}\tilde{P}_1, & k^{\rm in}=1\\ \tilde{P}(k^{\rm in}), & k^{\rm in}>1,\end{cases} \tag{6}$$

where $C$ is a constant, obtained from the normalization condition $\Sigma_{k^{\rm in}}P(k^{\rm in})=1$, which has the value

$$C=\tilde{P}_1+\sum_{K=2}^{\infty}\tilde{P}(K)=\frac{2a+2}{a}. \tag{7}$$

From this equation, we obtain the analytic expression for the in-degree distribution

$$P(k^{\rm in})=\begin{cases}\dfrac{2+a}{2+2a}, & k^{\rm in}=1\\[2ex] \dfrac{\Gamma(2+2a)}{\Gamma(a)}\dfrac{\Gamma(a+k^{\rm in})}{\Gamma(2+2a+k^{\rm in})}, & k^{\rm in}>1.\end{cases} \tag{8}$$

For large $k^{\rm in}$, we can expand the previous expression using Stirling's approximation to obtain that the in-degree distribution follows the asymptotic behavior:

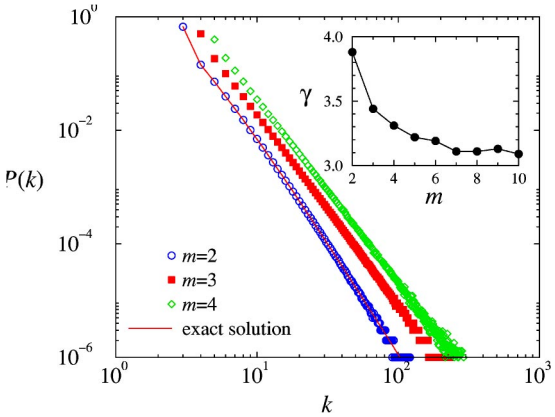


FIG. 1. Degree distribution of model $A$ for $a=m$, network size of $N=10^7$, and different values of $m$. The continuous line is the exact distribution for $m=2$ given by Eq. (10). The inset shows the value of the exponent $\gamma$ as a function of $m$ obtained from numerical simulations.

$$P(k^{\text{in}})\sim k^{\text{in}-\gamma},\quad \gamma=2+a. \tag{9}$$

Moreover, since the out-degree of all nodes is $m$, the degree $k$ of a node (in-degree plus out-degree) is $m+k^{\text{in}}$ and will follow the same distribution shifted by $m$. For the particular case $a=m=2$, the degree distribution takes the form

$$P(k)=\begin{cases}\dfrac{2}{3}, & k=3\\[2ex] \dfrac{120}{k(k+1)(k+2)(k+3)}, & k>3.\end{cases} \tag{10}$$

In Fig. 1, we plot the degree distribution obtained from numerical simulations of model $A$ for $a=m$. For $m=2$, the numerical points are in very good agreement with the exact distribution given in Eq. (10) with a power law decay with exponent $\gamma=2+a=4$. In the limiting case of large $m$, the continuous approach predicts the exponent 3 [25] [see Eq. (2)], giving us a lower bound. Hence,

$$\text{model } A \text{ with } a=m \;\Rightarrow\; 3<\gamma\leqslant 4 \tag{11}$$

and, therefore, the degree distribution has always a bounded second moment. For larger $m$ the distribution follows a power law decay but with an exponent $\gamma$ that depends on $m$. In order to show that the degree distribution approaches for each $m$ an asymptotic power law behavior with $\gamma>3$ we performed large-scale simulations of networks with $N=10^7$ nodes. In Fig. 1, we report the behavior of the exponent $\gamma$ as a function of $m$. For all values of $m<10$, the degree exponents strongly deviates from the $m\to\infty$ limit.

### B. Model *B*

Using similar arguments we can compute the degree distribution of model $B$ for $m=1$. In this case we also have two nodes at the deactivation process, the one just added and the one surviving from the previous deactivation step. The former has in-degree 0, while the latter (the oldest) has in-degree $K\geqslant 1$, and one of them is deactivated with probability

$$p_d(K)=\frac{a}{2a+K},\quad p_d(0)=1-p_d(K). \tag{12}$$

The probability that when the oldest node is deactivated it has degree $K$ is given by

$$\tilde{P}(K)=\prod_{\ell=1}^{K-1}[1-p_d(\ell)]p_d(K)=\frac{\Gamma(1+2a)}{\Gamma(a)}\frac{\Gamma(a+K)}{\Gamma(1+2a+K)}. \tag{13}$$

In the process of creating a node of in-degree $K$, $K-1$ nodes of in-degree 0 have been created. The average number of nodes with in-degree 0 created is

$$\tilde{P}_0=\sum_{K=1}^{\infty}(K-1)\tilde{P}(K)=\frac{a+1}{a-1}. \tag{14}$$

Thus, the analytic expression for the normalized in-degree distribution is given by

$$P(k^{\text{in}})=C^{-1}\begin{cases}\tilde{P}_0, & k^{\text{in}}=0\\[1ex] \tilde{P}(k^{\text{in}}), & k^{\text{in}}>0,\end{cases} \tag{15}$$

with the normalization constant

$$C=\tilde{P}_0+\sum_{k=1}^{\infty}\tilde{P}(K)=\frac{2a}{a-1}. \tag{16}$$

From here follows the expression for the degree distribution (where $k=m+k^{\text{in}}$)

$$P(k)=\begin{cases}\dfrac{1+a}{2a}, & k=1\\[2ex] \dfrac{\Gamma(2a)}{\Gamma(a-1)}\dfrac{\Gamma(a+k-1)}{\Gamma(2a+k)}, & k>1.\end{cases} \tag{17}$$

For large $k$ the degree distribution follows the asymptotic behavior:

$$P(k)\sim k^{-\gamma},\quad \gamma=1+a. \tag{18}$$

In Fig. 2, we show the degree distribution obtained from numerical simulations of model B with $a=m$. For $a=m=1$, we recover the predicted exponent $\gamma=-2$. Also in this case, we provide large-scale numerical simulations ($N=10^7$) of networks with larger values of $m$. The obtained distributions still follow a power law decay but with an exponent $\gamma$ that is a continuously increasing function of $m$. It is worth remarking that for $m<10$, the degree exponent is stable and strongly differs from the value $\gamma=3$.

It is worth noticing that for $a=m=1$, the analytic solution, Eq. (17), is singular, as can be readily seen from the $\Gamma(a-1)$ factor in the denominator. In fact, the solution in this case is $P(k)=\delta_{k,1}$, that is, in the thermodynamic limit

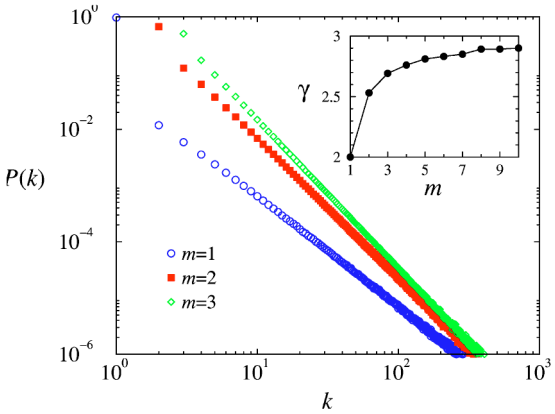


FIG. 2. Degree distribution of model $B$ for $a=m$, network size of $N=10^7$, and different values of $m$. The inset shows the value of the exponent $\gamma$ as a function of $m$ obtained from numerical simulations.

(infinitely large network), the weight of the nodes with degree 1 is overwhelming with respect to the nodes with different connectivity. This singularity is rooted in the fact that the distribution, with exponent $\gamma=-2$, lacks a finite first moment in the thermodynamic limit, while we know that, by definition, model $B$ has average connectivity $\langle k\rangle=2$. This necessarily implies that there must be an implicit dependence on the network size $N$ in the degree distribution for $a=m=1$, dependence that cannot be assessed by our analytic solution since we are already working in the infinite network limit. We can nevertheless estimate the functional form of the degree distribution for a finite network composed by $N$ nodes, which has a maximum connectivity $k_c$, such that there are no nodes with degree larger than $k_c$. Assuming that the distribution for $k>1$ follows the same functional form as Eq. (17), we have that for $a=1$,

$$P_N(k)=\begin{cases} C_1, & k=1 \\ \dfrac{C_2}{k(k+1)}, & 1<k\leqslant k_c, \end{cases} \tag{19}$$

where $C_1$ and $C_2$ are constants to be determined by the normalization conditions $\Sigma_{k=1}^{\infty}P_N(k)=1$ and $\Sigma_{k=1}^{k_c}kP_N(k)=2$ (the upper limit in the first normalization condition can be taken to be infinite, since the corrections stemming from $k_c$ are of lower order). From these two conditions we obtain, in the continuous $k$ approximation that replaces sums by integrals,

$$C_1=1-\frac{2\ln(3/2)}{\ln\left(\dfrac{1+k_c}{2}\right)}, \quad C_2=\frac{2}{\ln\left(\dfrac{1+k_c}{2}\right)}. \tag{20}$$

For finite SF networks with degree distribution $P(k)\sim k^{-\gamma}$, the maximum degree $k_c$ scales with the number of nodes as $k_c\sim N^{1/(\gamma-1)}$ [2]. In the present case, we have $k_c\sim N$, and thus, for large $N$,

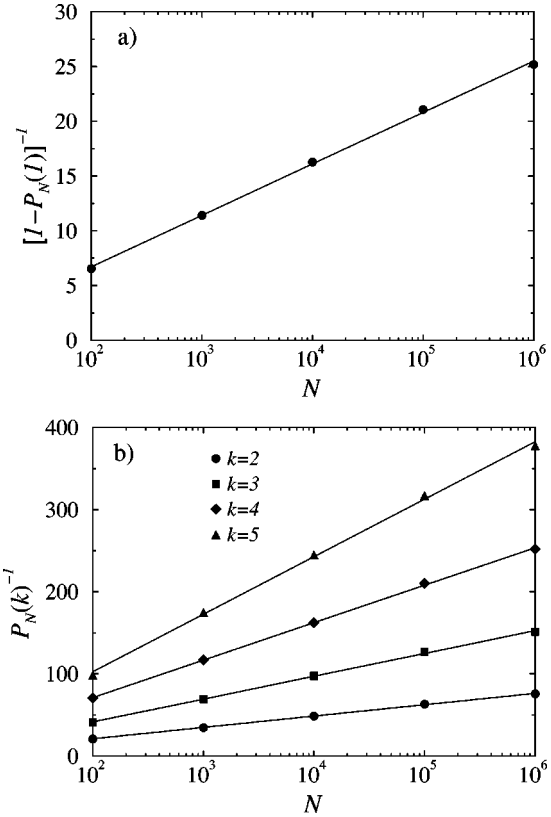


FIG. 3. Scaling of the degree distribution $P_N(k)$ for the $B$ model with $a=m=1$ at fixed $k$, as a function of the network size $N$, for (a) $k=1$ and (b) $k>1$. The solid lines are least-squares fits to the form $[1-P_N(1)]^{-1}\sim\ln N$ in (a) and $P_N(k)^{-1}\sim\ln N$ in (b), as predicted by Eq. (22).

$$1-C_1\sim\frac{1}{\ln N}, \quad C_2\sim\frac{1}{\ln N}. \tag{21}$$

Therefore, in the limit $N\to\infty$, we recover a singular degree distribution with $C_1\to 1$ and $C_2\to 0$. We can check numerically this result by noticing that, from Eqs. (19) and (21), the degree distribution at fixed $k$ should scale as

$$1-P_N(1)\sim\frac{1}{\ln N}, \quad P_N(k)\sim\frac{1}{\ln N}, \quad k>1. \tag{22}$$

We have verified this scaling form in Fig. 3. Therefore, in model $B$ with $a=m=1$ we obtain a degree distribution that decays as $k^{-2}$, but with a normalization constant for $k>1$ that decays with the network size as $1/\ln N$. Finally, it is worth mentioning that the second moment of the distribution is diverging as $\langle k^2\rangle\sim N/\ln N$. Despite this singular behavior for $a=m=1$, however, Eq. (17) remains exact for any value of $a\neq 1$.

From the results of Fig. 2, together with the upper bound $\gamma=3$ obtained from the large $m$ approximation [25], we have that

$$\text{model } B \text{ with } a=m \quad\Rightarrow\quad 2\leqslant\gamma<3, \tag{23}$$

and, therefore, the degree distribution has a divergent second moment.

The analysis made above has shown that the deactivation model is quite sensitive to the order in which steps (2) and (3) are performed, yielding degree distributions with a finite or divergent second moment, depending on the order. In addition, the exponent $\gamma$ is rather sensible to the value of $a=m$, showing a wide range of variation. This fact has not been noticed in previous works where this model has been considered [25–27], prompting that some of the conclusions obtained in those works should be reconsidered in this perspective.

## IV. CLUSTERING COEFFICIENT

We can go beyond the degree distribution and compute the clustering coefficient $c(k)$ as a function of the node degree $k$ [6,28]. For this quantity we can perform an analytic calculation for any value of $a$ and $m$ and for both models $A$ and $B$. In order to compute the clustering coefficient, we will consider the network as undirected and denote by $k_i=k_i^{\mathrm{in}}+m$ the total degree of the node $i$.

The clustering coefficient of the node $i$ is defined by [16]

$$c_i=\frac{2e_i}{k_i(k_i-1)}, \tag{24}$$

where $e_i$ is the number of edges between the neighbors of node $i$ and it is divided by its maximum possible value $k_i(k_i-1)/2$. In the deactivation model, new edges are created between the active nodes and the added node. Hence, $e_i$ remains constant for inactive nodes and increases only for the active ones. Moreover, all the active nodes are connected. Hence, each time we add a node, the degree $k_i$ of each active node, $i$ increases by one and $e_i$ increases by $m-1$, where $m-1$ are just the new links between the new neighbor of $i$ (the added node) and the remaining active nodes. Therefore, the dynamics of $e_i$ is given by

$$\frac{\partial e_i}{\partial t}=(m-1), \tag{25}$$

while the connectivity obeys the relation $k_i(t)=m+t$. Here, $t=0$ corresponds to the time at which the node $i$ was created. Besides, when the node is added it has degree $m$, thus $e_i(0)=m(m-1)/2$ and, therefore, $c_i(0)=1$. Integrating Eq. (25) with this initial condition and substituting the result in Eq. (24), taking into account that $t=k_i-m$, we obtain

$$\begin{aligned} c(k)&=\frac{m(m-1)}{k(k-1)}+\frac{2(m-1)(k-m)}{k(k-1)}\\ &=\frac{2(m-1)}{k}-\frac{(m-1)(m-2)}{k(k-1)}, \end{aligned} \tag{26}$$

where the last expression in Eq. (26) is obtained after some algebraic manipulations. Equation (26) recovers the results previously obtained in Ref. [28]. For $m=1$, the network is a tree, and therefore we obviously recover $c(k)=0$. For $m=2$, we obtain the exact behavior $c(k)=2/k$. For $m>2$, the asymptotic behavior for large $k$ is $c(k)\sim 1/k$ [25]. Interestingly, we recover in this model the same behavior of $c(k)$ found in other systems in Ref. [30].

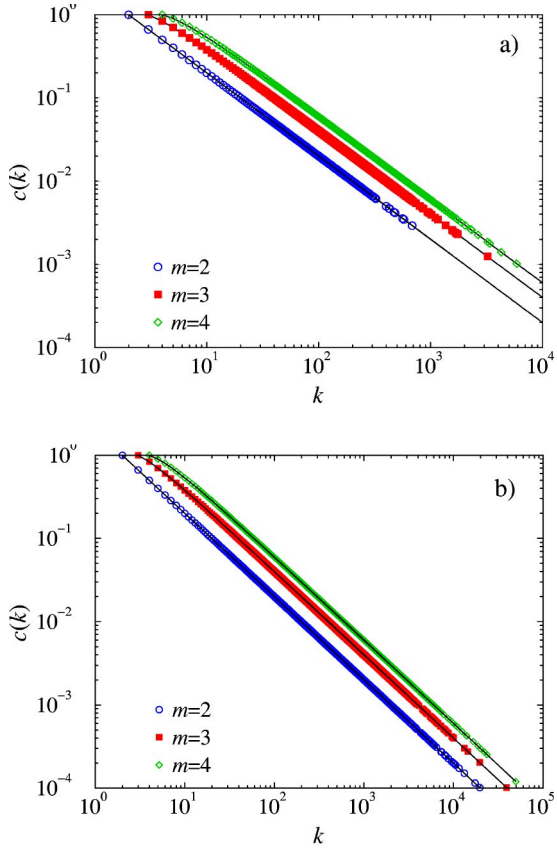


FIG. 4. Clustering coefficient as a function of the node degree for different values of $m$. The points were obtained from numerical simulations of (a) model $A$ and (b) model $B$, up to a network size $N=10^5$. The continuous lines correspond with the analytical solution given in Eq. (26)

In Fig. 4, we plot the clustering coefficient as a function of the node degree obtained for models $A$ and $B$ and different values of $m$ from numerical simulations. As it can be seen, the numerical dependency coincides with the analytical expression in Eq. (26).

## V. DEGREE CORRELATION FUNCTION

Degree correlations can be characterized by analyzing the nearest neighbor average degree introduced in Refs. [5,6], defined as

$$k_{nn,i}=\frac{D_i}{k_i}, \tag{27}$$

where $D_i$ is the sum of the degrees of the neighbors of node $i$. In uncorrelated networks, the quantity $k_{nn,i}$ does not show any dependence on the degree of the node $i$. This is not the case when degree correlations are present. In this case, $k_{nn,i}$ is a function of the degree of the node whose nearest neighbors are analyzed. In particular, we can face two possible

kinds of correlation. In the first situation, nodes with high connectivity will connect more preferably to highly connected nodes; a property referred to as "assortative mixing." On the opposite side, it is possible to have "dissortative mixing"; i.e., highly connected nodes are preferably connected to nodes with low connectivity [31].

In the deactivation model, when the node is added it has degree $m$ and $D_i = m\langle k\rangle_{\mathcal{A}}$, where $\langle k\rangle_{\mathcal{A}}$ is the average degree among active neighbors. Then, if the node $i$ is active, it is, by construction, neighbor of the $m-1$ remaining active nodes. Thus, every time a new node is added, $k_i$ increases by one and $D_i$ increases by $(m-1)+m$, $m-1$ because the degree of the remaining $m-1$ neighbors have also increased by one and $m$ because the new neighbor has degree $m$. Hence,

$$\frac{\partial D_i}{\partial t} = (2m-1) \tag{28}$$

for each active node $i$. Integrating this equation, taking into account the initial condition $D_i(0) = m\langle k\rangle_{\mathcal{A}}$ and the relation $t = k_i - m$, we obtain that

$$D_i' = (2m-1)(k_i - m) + m\langle k\rangle_{\mathcal{A}} \tag{29}$$

when the node $i$ is deactivated. Now, when an active node becomes inactive, its degree remains fixed but the degree of its active neighbor nodes will still increase until they get deactivated. Therefore, in the infinite time limit, we have

$$D_i = D_i' + \Delta D_i, \tag{30}$$

where $\Delta D_i$ is the increase of $D_i$, since node $i$ was set inactive until all its neighbors are set inactive.

Hence, from Eqs. (27), (29), and (30), it follows that

$$k_{nn,i} = 2m - 1 + \frac{m\langle k\rangle_{\mathcal{A}} + \Delta D_i - m(2m-1)}{k_i}. \tag{31}$$

It remains now the task to assess the possible dependence of $\Delta D_i$ on the connectivity $k_i$ (it is clear that the long time average of $\langle k\rangle_{\mathcal{A}}$ must be independent of the connectivity of any deactivated node). For the minimum $m$ ($m=2$ for model $A$ and $m=1$ for model $B$), the degree of an active node set inactive is not correlated with the degree of the remaining active nodes, since those remaining nodes have always degrees 2 and 3 in model $A$ with $m=2$, and degree 1 in model $B$ with $m=1$, independent of the degree of the last deactivated node. Therefore, in this case $\Delta D_i$ cannot depend on $k_i$. This lack of correlations is also clear for $m \gg 1$, where the sum $\Sigma_{j\in\mathcal{A}}(a+k_j)^{-1}$ in Eq. (1) is a constant [25] and, therefore, the degree of the active nodes in not correlated with the degree of the inactive nodes. For intermediate values of $m$, however, the degree of the active nodes may be correlated in such a way that $\Delta D_i$ depends on $k_i$.

In Fig. 5, we plot the dependency of the average nearest neighbors degree $\overline{k_{nn}}(k)$ as a function of the degree $k$ for models $A$ and $B$ and different values of $m$. In the case of model $A$, $\overline{k_{nn}}(k) - (2m-1) \sim 1/k$ even for $m \neq 2$, in agreement with Eq. (31). In the case of model $B$, $\overline{k_{nn}}(k) - (2m-1) \sim 1/k$ for $m=1$ and $m=10$ but decays faster for intermediate values of $m$, a behavior that we are not able to explain. Thus, in this case the correlations between the active node degrees introduce stronger deviations for intermediate values of $m$. In all cases, however, we find that correlations in the deactivation model are of "disassortative" nature; i.e., highly connected nodes are preferably connected with poorly connected nodes. It is also worth stressing that the results for model B with $m=1$ must be taken with a grain of salt, given the singular nature of the model exposed in Sec. III B.

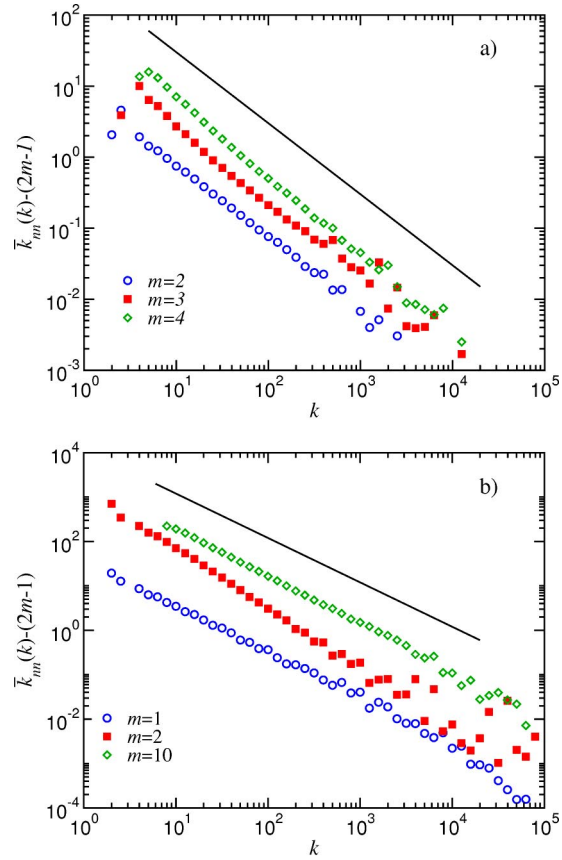


FIG. 5. Average nearest neighbor degree as a function of the degree $k$ for different values of $m$. The points were obtained from numerical simulations of (a) model $A$ and (b) model $B$, up to a network size $N=10^5$, averaging over 1000 realizations. The continuous lines correspond with the analytical dependency $\overline{k_{nn}}(k) - (2m-1) \sim 1/k$.

In the deactivation model, either $A$ or $B$, for a fixed network size $N$ and assuming that $\Delta D_i$ does not grow faster than $k_i$, we have that in the limit $k_i \to \infty$, $k_{nn,i} \to 2m-1$. That is, the average nearest neighbor degree of the hubs (nodes with largest $k_i$) equals $\langle k\rangle - 1$, as previously pointed out in Ref. [26]. However, this fact does not necessarily imply that $\Delta D_i$ is independent of $N$. One way to check this point is to compute the average of $k_{nn,i}$ over all nodes, $\langle \overline{k_{nn}}\rangle_N = \Sigma_k P(k)\overline{k_{nn}}(k)$. Let us assume that $\overline{k_{nn}} \sim \langle k\rangle - 1 + \alpha/k$, where $\alpha$ is depending on $\Delta D_i$. If $\Delta D_i$ is approaching a constant value, we should obtain $\langle \overline{k_{nn}}\rangle_N \sim$ const, independently of $N$. In Fig. 6, we show how $\langle \overline{k_{nn}}\rangle_N$ behaves with increasing $N$ for $a=m$. For model $A$, where $3 < \gamma \leq 4$, it

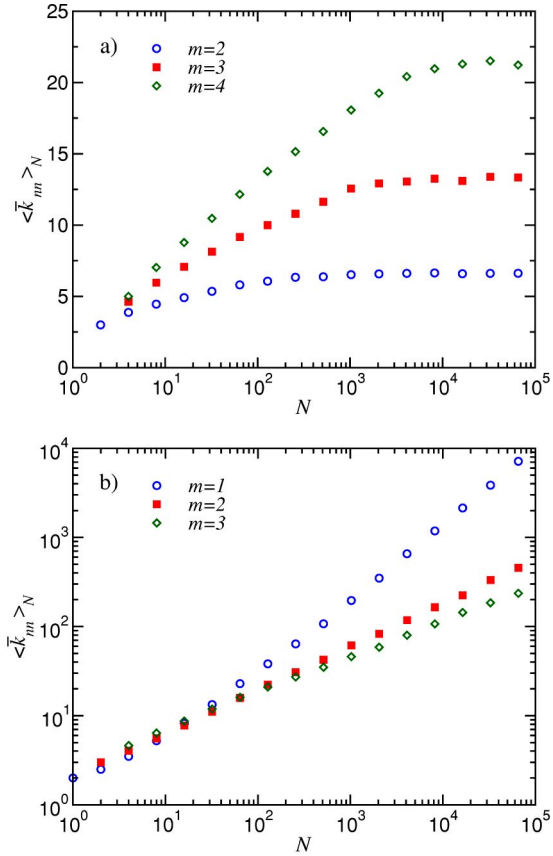


FIG. 6. Average nearest neighbor degree as a function of the network size $N$ for different values of $m$. The points were obtained from numerical simulations of (a) model $A$ and (b) model $B$, up to a network size $N=10^5$, averaging over 1000 realizations.

approaches a stationary value for $N \gg 1$. Moreover, the asymptotic limit of $\langle \overline{k_{nn}} \rangle_N$ increases with increasing $m$. In fact, with increasing $m$ the exponent $\gamma$ decreases approaching the limit $\gamma=3$ for $m \gg 1$, where $\langle k^{nn} \rangle$ diverges logarithmically with $N$. On the contrary, for model $B$, where $2 \leqslant \gamma < 3$, $\langle \overline{k_{nn}} \rangle_N$ is growing with $N$ following a power law. This implies that $\Delta D_i$ is a diverging function of $N$ and that in the thermodynamic limit (in which we perform first the limit $N \to \infty$), the average nearest neighbor connectivity curve is progressively shifting to larger and larger values. This finally points out that the average nearest neighbor connectivity of hubs is not a well-defined quantity since the $k_i \to \infty$ limit must be performed only after the $N \to \infty$ limit. The divergence of $\langle \overline{k_{nn}} \rangle_N$ with $N$ is related to a general property of SF networks with diverging connectivity fluctuations and it is dictated by the detailed balance of connectivity [32,33].

## VI. DIAMETER AND SHORTEST PATH LENGTH

Another fundamental topological feature of complex networks is identified by the scaling of the average path length among nodes and the network's diameter. The minimum path between two nodes is given by the minimum number of intermediate nodes that must be traversed to go from node to node. The average minimum path length $\langle d \rangle$ is thus defined

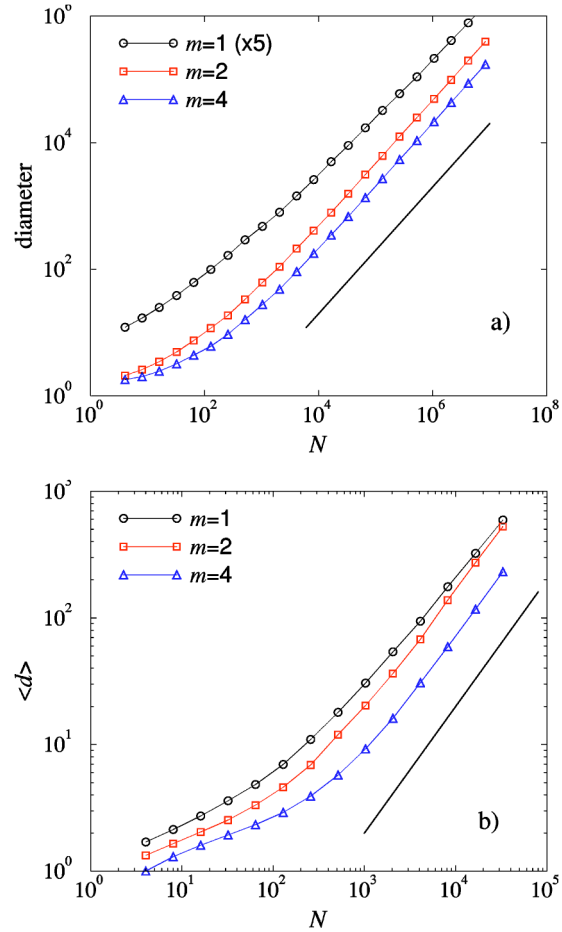


FIG. 7. Scaling of the diameter (a) and the average shortest path $\langle d \rangle$ (b) in the model $B$ for different values of $m$. The reference lines have slope 1. For the sake of clarity, the curve for $m=1$ in (a) has been shifted by a factor 5.

as the minimum path distance averaged over all the possible pairs of nodes in the network. Similarly, the network diameter is defined as the largest among the shortest paths between any two nodes in the network.

While regular networks (for instance hypercubic lattices) have a diameter scaling with the size $N$ as the inverse of the Euclidean dimension, many complex networks show striking small-world properties; i.e., in an average one can go from one node to any other node in the system by passing through a very small number of intermediate nodes [16]. In this case the graph diameter grows logarithmically, or even slower, with the system's number of nodes $N$.

In Ref. [28], it has been noticed that for large $m$ values, $\langle d \rangle$ scales linearly with the network size $N$. In the deactivation model ($A$ and $B$), we measured both the diameter and the average minimum path distance $\langle d \rangle$ as a function of $N$ for values of $a=m$ ranging from 1 to 4. In all cases we find that after a small size transient, both metrics approach a linear scaling with $N$. In Fig. 7, we report the results obtained in the case of the deactivation model with rule $B$. This evidence implies that the topology of the generated networks is approaching those of a one-dimensional lattice. In other words,

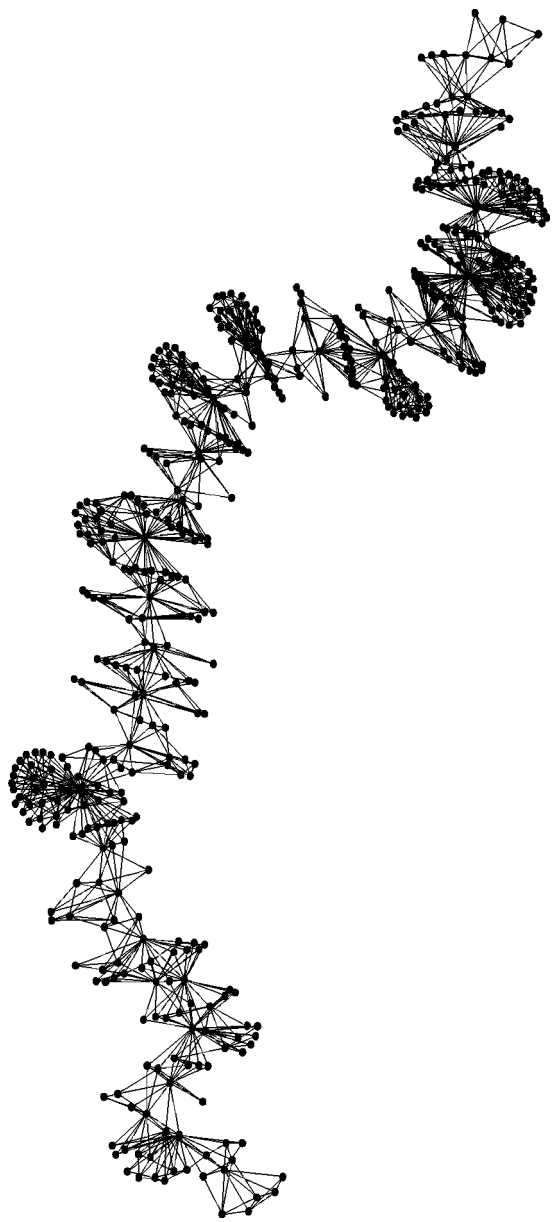

FIG. 8. Illustration of a typical network generated with the deactivation model $B$ with $a=m=3$ (the size is $N=10^3$). The linear topology with some local highly connected clusters forming a chain is evident.

the deactivation model does not exhibit small-world properties.

In order to provide a visual representation of the deactivation model topology, we report in Fig. 8 the illustration of a network generated with model $B$ and $a=m=3$. The linear topology of the network with some local highly connected clusters forming a chain is evident. The linear structure is made up of groups of nodes connected to a node which has been active for longer times and has had the possibility to develop a high number of connections. Once these hubs are deactivated, they do not receive any further connection. The network grows by adding bridge nodes that are rapidly deactivated until a new dense cluster is developed by a node that is active long enough. The growth mechanism, however, does not allow the formation of shortcuts between the deactivated region of the network and the new active nodes, hindering the development of small-world properties. The linear chain is therefore reflecting the time evolution of the structure: recently added nodes are separated from the original core of active nodes by a sequence of deactivated nodes that increases proportionally to the network size. By inspecting networks with larger $m$, we find very similar structures, with an increasing size of the dense clusters forming the linear chain. As we shall discuss in the following section, the absence of small-world properties might have a relevant effect in many physical properties of the network.

## VII. DISCUSSION AND CONCLUSIONS

In the present work, we have provided a detailed analysis of the deactivation model introduced in Ref. [25]. The model shows a rich behavior, being very sensible to the value of the parameters used in the model and slight variations of the growing algorithm. The most striking result is that the degree distribution is depending on the value of the number of simultaneously active nodes $m$ and also in the case in which $a=m$; i.e., when the deactivation probability is related to the nodes' total degree. The degree exponent is asymptotically approaching the value $\gamma=3$ only for $m\rightarrow\infty$, and the SF properties of networks suffer large variations in the range $1\leqslant m\leqslant 10$. Along with the high clustering observed in previous works, we find that the model exhibits interesting degree correlation properties. In particular, we find marked disassortative mixing properties; i.e., highly connected nodes link preferably to poorly connected nodes. The analytical expression for the degree correlation is obtained and recovered by numerical simulations. Strikingly, the SF and correlation properties are not associated with small-world properties. The numerical analysis shows that for all values of $m$, the network diameter is increasing linearly with the number of nodes. The network thus approaches a linear structure, lacking long-range shortcuts.

One of the most interesting issues related to SF networks is the effects of their complex topological features on the dynamics of spreading phenomena [21,22,24,34] and the onset of percolation transitions [18–20]. In the case of random SF networks, where degree correlations are absent, it has been found that the epidemic threshold is proportional to $\langle k\rangle/\langle k^2\rangle$ [21,22]. Uncorrelated SF networks allow the onset of large epidemics whatever the spreading rate of the infection. This is a noticeable result that has a large impact in immunization as well as control and design policies in real networks [35,36]. On the other hand, most real networks show nontrivial degree correlations and clustering properties as it is the case in the present deactivation model. Similarly, the random removal of nodes does not destroy the connectivity of SF networks with $\gamma\leqslant 3$. In other words, the percolation transition is absent, and the networks are extremely robust to random damages [18–20]. A natural question is to know whether or not the clustering properties of SF networks plus their correlations alter the general results obtained for uncorrelated networks. For this reason, several recent works have addressed the effect of such correlations in the epi-

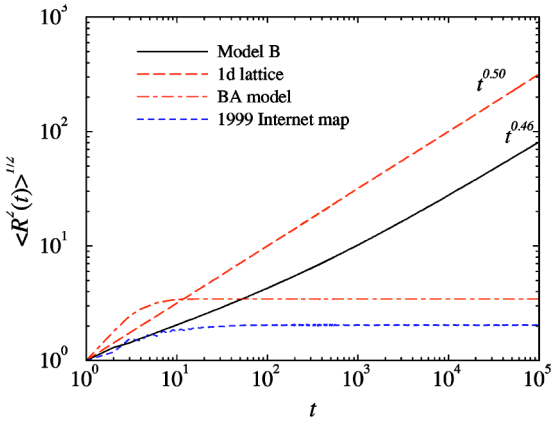


FIG. 9. Mean-square displacement of a random walker on the deactivation model with $m=3$, a one-dimensional lattice, and the Barabási-Albert model with $N=10^5$ nodes, as well as an Internet snapshot map from 1999 with 6301 nodes.

demic spreading occurring on these networks [26,27,32,37]. In particular, in Ref. [26] the existence of an epidemic threshold in the case of the deactivation model for rule *B* has been claimed.

The presence of a finite threshold in the deactivation model has been traced back to the high clustering coefficient and the finite limit of the average nearest neighbor connectivity of the largest hubs [26]. On the other hand, we have shown here that the average nearest neighbor connectivity in the system is diverging with the system size. What appears as more fundamental for the properties of spreading in the deactivation model is its linear structure with a diameter that increases with *N*. In a coarse grained picture, the epidemic spreading is dominated by the diffusion of the disease on a linear chain. In order to check this point, we have simulated a standard random walk in the *B* model with $m=3$. In Fig. 9, we plot the mean-square displacement of the random walker, $\langle R^2(t)\rangle^{1/2}=\langle[r(t)-r(0)]^2\rangle^{1/2}$, where the brackets denote an average over 250 realizations of the random walk on 250 different networks. For a purely diffusive system, as would be the case of a one-dimensional lattice, we would expect a scaling $\langle R^2(t)\rangle^{1/2}\sim t^{1/2}$. For the deactivation model we observe a slightly subdiffusive behavior with a mean-square displacement scaling as $\langle R^2(t)\rangle^{1/2}\sim t^{0.46}$. We thus conclude that dynamics on the deactivation model, is almost purely diffusive, as expected from its non-small-world character. The analysis of spreading and percolation properties in this network cannot therefore be performed at the mean-field level [21,22], but must include diffusion and most probably fluctuations, leading to a much more complex formalism based on a field theory [38]. For the sake of comparison, we have also plotted in Fig. 9 the mean-square displacement of a random walker on a Barabási-Albert network [17] and on a Internet snapshot map from 1999, collected by the National Laboratory for Applied Network Research [39]. As we observe, in these last two networks, $\langle R^2(t)\rangle^{1/2}$ saturates very quickly to a constant value, proportional to the network's diameter, indicating the presence of a strong small-world component. The essential difference of the diffusive properties between the Internet and the deactivation model does not allow to extend the conclusions obtained from the model to the spreading in the real system.

The same applies to percolation properties that naturally exhibit a finite threshold in this case. The fact that spreading and percolation properties on the deactivation model are similar to those of regular lattices because of the absence of small-world features is corroborated by the analysis of Ref. [27] that shows how the introduction of a small amount of shortcuts restores the usual absence of a percolation threshold. In this perspective, it would be extremely interesting to have a detailed study of the epidemic spreading properties in the case of the deactivation model with random rewiring [28], in order to assess the effect of clustering and degree correlations in spreading processes in SF networks with small-world properties.

## ACKNOWLEDGMENTS

This work has been partially supported by the European Commission FET Open project COSIN under Project No. IST-2001-33555. R.P.-S. acknowledges financial support from the Ministerio de Ciencia y Tecnología (Spain).

[1] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
[2] S.N. Dorogovtsev and J.F.F. Mendes, Adv. Phys. **51**, 1079 (2002).
[3] M. Faloutsos, P. Faloutsos, and C. Faloutsos, Comput. Commun. **29**, 251 (1999).
[4] G. Caldarelli, R. Marchetti, and L. Pietronero, Europhys. Lett. **52**, 386 (2000).
[5] R. Pastor-Satorras, A. Vázquez, and A. Vespignani, Phys. Rev. Lett. **87**, 258701 (2001).
[6] A. Vázquez, R. Pastor-Satorras, and A. Vespignani, Phys. Rev. E **65**, 066130 (2002).
[7] R. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **401**, 130 (1999).
[8] S.H. Strogatz, Nature (London) **410**, 268 (2001).
[9] J.M. Montoya and R.V. Solé, J. Theor. Biol. **214**, 405 (2002).
[10] A. Wagner, Mol. Biol. Evol. **18**, 1283 (2001).
[11] H. Jeong, S. Mason, A.L. Barabási, and Z.N. Oltvai, Nature (London) **411**, 41 (2001).
[12] R.V. Solé, R. Pastor-Satorras, E. Smith, and T. Kepler, Adv. Complex Syst. **5**, 43 (2002).
[13] A. Vázquez, A. Flammini, A. Maritan, and A. Vespignani, ComPlexUs **1**, 38 (2003).
[14] G. Chartrand and L. Lesniak, *Graphs & Digraphs* (Wadsworth, Menlo Park, 1986).
[15] P. Erdös and P. Rényi, Publ. Math. Inst. Hung. Acad. Sci. **5**, 17 (1960).
[16] D.J. Watts and S.H. Strogatz, Nature (London) **393**, 440 (1998).
[17] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[18] R.A. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **406**, 378 (2000).

[19] D.S. Callaway, M.E.J. Newman, S.H. Strogatz, and D.J. Watts, Phys. Rev. Lett. **85**, 5468 (2000).
[20] R. Cohen, K. Erez, D. ben Avraham, and S. Havlin, Phys. Rev. Lett. **86**, 3682 (2001).
[21] R. Pastor-Satorras and A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001).
[22] R. Pastor-Satorras and A. Vespignani, Phys. Rev. E **63**, 066117 (2001).
[23] R.M. May and A.L. Lloyd, Phys. Rev. E **64**, 066112 (2001).
[24] Y. Moreno, R. Pastor-Satorras, and A. Vespignani, Eur. Phys. J. B **26**, 521 (2002).
[25] K. Klemm and V.M. Eguiluz, Phys. Rev. E **65**, 036123 (2002).
[26] V.M. Eguiluz and K. Klemm, Phys. Rev. Lett. **89**, 108701 (2002).
[27] P. Crucitti, V. Latora, M. Marchiori, and A. Rapisarda, e-print cond-mat/0205601.
[28] K. Klemm and V.M. Eguiluz, Phys. Rev. E **65**, 057102 (2002).
[29] M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1972).
[30] E. Ravasz and A.-L. Barabási, e-print cond-mat/0206130.
[31] M.E.J. Newman, Phys. Rev. Lett. **89**, 208701 (2002).
[32] M. Boguñá and R. Pastor-Satorras, Phys. Rev. E **66**, 047104 (2002).
[33] M. Boguñá, R. Pastor-Satorras, and A. Vespignani, Phys. Rev. Lett. **90**, 028701 (2003).
[34] A.L. Lloyd and R.M. May, Science **292**, 1316 (2001).
[35] R. Pastor-Satorras and A. Vespignani, Phys. Rev. E **65**, 036104 (2001).
[36] Z. Dezsö and A.-L. Barabási, Phys. Rev. E **65**, 055103(R) (2002).
[37] C.P. Warren, L.M. Sander, and I.M. Sokolov, Phys. Rev. E **66**, 056105 (2002).
[38] J. Marro and R. Dickman, *Nonequilibrium Phase Transitions in Lattice Models* (Cambridge University Press, Cambridge, 1999).
[39] The National Laboratory for Applied Network Research (NLANR), sponsored by the National Science Foundation, provides Internet routing related information based on border gateway protocol data (see http://moat.nlanr.net/).