Evolving networks with disadvantaged long-range connections

R. Xulvi-Brunet[1] and I. M. Sokolov[1,2]
[1]*Institut für Physik, Humboldt Universität zu Berlin, Invalidenstrasse 110, D-10115 Berlin, Germany*
[2]*Theoretische Polymerphysik, Universität Freiburg, Hermann Herder Strasse 3, D-79104 Freiburg, Germany*
(Received 7 May 2002; published 26 August 2002)

We consider a growing network, whose growth algorithm is based on the preferential attachment typical for scale-free constructions, but where the long-range bonds are disadvantaged. Thus, the probability of getting connected to a site at distance d is proportional to $d^{-\alpha}$, where α is a tunable parameter of the model. We show that the properties of the networks grown with $\alpha<1$ are close to those of the genuine scale-free construction, while for $\alpha>1$ the structure of the network is quite different. Thus, in this regime, the node degree distribution is no longer a power law, and it is well represented by a stretched exponential. On the other hand, the small-world property of the growing networks is preserved at all values of α.

DOI: 10.1103/PhysRevE.66.026118 PACS number(s): 05.50.+q, 89.75.Hc

Complex weblike structures (the small-world or scale-free networks) have recently become an object of extensive investigation, and in recent years great success in understanding the properties of these structures has been achieved (see Refs. [1,2] for a review). Apart from appealing mathematics, this recent interest is due to the fact that many natural and technological systems, such as polymer networks [3], the science collaboration network [4–6], and networks of chemical reactions in a living cell [7–9] seem to be organized according to some internal principles. Thus, the Internet [10], the network of human sexual contacts [11], and the WWW [12] possess similar structures, e.g., they are all based on the preferential attachment of the newly introduced nodes to the highly connected old ones. All these networks show the small-world property: the typical distance (in terms of the number of intermediate connections) between two nodes grows logarithmically with the web's size.

One of the prominent examples of a mathematical model of such a growing network is the scale-free (SF) construction of Albert and Barabási [1,2,13]; and one of its most interesting properties is the very specific form of the probability distribution of the degree k_i of nodes (i.e., of the number of bonds connecting any given node i with other ones in the network): $P(k)\propto k^{-3}$ [1,2,13–16]. Many models have been presented, based on the same two most important ingredients: growth and preferential attachment. Examples are models with an accelerated growth of the network [17,18], models with a nonlinear preferential attachment [16], with nodes provided with an initial attractiveness [14,19], with growth constraints as aging and cost [20,21], models that have a competitive aspect of the nodes [22], or models of networks that incorporate local events as the addition, rewiring, or removal of nodes or edges [23].

The SF construction may be a reasonable approximation for such world-spanning networks such as one of the Internet's information transmission channels or one of the formal links of the WWW. On the other hand, in many situations (as in a network of human sexual contacts) a connection means a physical contact, i.e., that the contacting individuals, representing the nodes of the network, have to occur at the same site and at the same time, thus introducing a clear geographical aspect. In what follows, we present a simple model taking into account this metrical (geographical) aspect, where the probability of connecting two nodes depends both on the number of connections that the nodes already have (as in the genuine SF construction), and on the distance between them. That is, we treat an emerging network in a metric space. In this emerging network the probability that a newly introduced node n is connected to a previously existing node i is proportional to the number k_i of the already existing connections of node i (preferential attachment prescription), but, on the other hand, the too long bonds are disadvantaged, because this probability depends on the Euclidean distance d_{in} between the nodes n and i as $d_{in}^{-\alpha}$, (clearly, a "scale-free" function), with $\alpha>0$.

Based on extensive numerical simulations of a one-dimensional situation, we show that even if the length penalties are mild, the model exhibits properties that differ strongly from those of the usual scale-free networks. Thus, the corresponding degree distribution function $P(k)$ depends strongly on α. We show, in particular, that for $\alpha<1$ the behavior of $P(k)$ is similar to the behavior of the SF model without penalties, so that asymptotically $P(k)\propto k^{-3}$ (a distribution that possesses a mean, but no dispersion, and corresponds to strong, universal fluctuations). On the other hand, for $\alpha>1$ the behavior of $P(k)$ is well described by a stretched exponential $P(k)\propto\exp(-bk^{\gamma})$, with the power γ depending on α, so that the fluctuations in k are rather weak. We discuss the reasons for such a dramatic change, being rooted in the probability of the connection between the nodes as a function of the distance, and the overall structure of the emerging network, preserving its small-world nature even at large (probably at all) α values.

We start from a one-dimensional lattice of L sites, spaced by a unit distance, and apply cyclic boundary conditions. We will let our network grow on this structure, so that each lattice site will be a possible location of a network's node. We denote by n_i the position in the lattice of a node i. The distance d_{ij} between any two nodes i and j is defined as

$$d_{ij}=\min\{|n_i-n_j|,(L-|n_i-n_j|)\}. \tag{1}$$

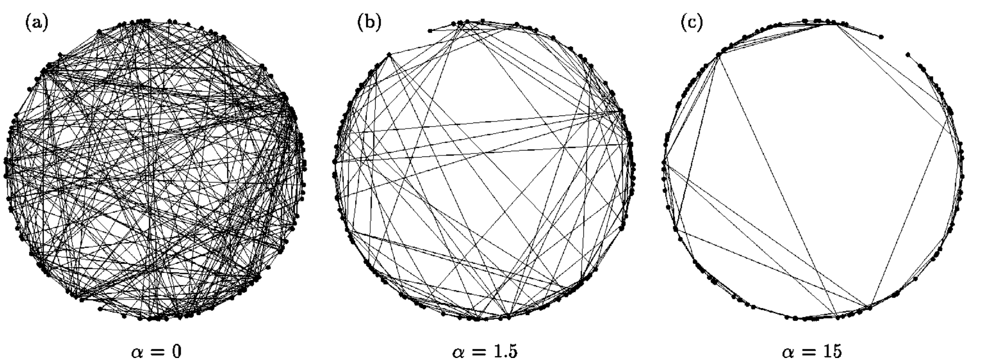

FIG. 1. Networks generated using the simulation prescription, Eq. (2), with different values of α: (a) $\alpha=0$, (b) $\alpha=1.5$, and (c) $\alpha=15$. All three examples have 300 edges, $L=10^6$, $N=105$, and $m=3$. Note the change in the appearance of the networks. The network (a) is a genuine SF construction while (c) strongly resembles the Watts and Strogatz small-world network.

Let us now construct the network. First, we choose randomly an even number m_0 of sites from the lattice and we bind them in pairs with one bond each. This will be our initial condition. That is, at $t=0$, our network will consist of m_0 nodes connected in pairs. As in the SF model, we will add at every time step a new node to our network (linear growth). We proceed according to the following rule: at every time step, we choose at random a free site of our lattice, and pose the new node there. This new node is then connected through m edges ($m \leqslant m_0$) with m different nodes already present in the network. After t time steps the algorithm results in a network with $t+m_0$ nodes and $mt+m_0/2$ edges. In contrast with the SF model, the probability Π for the new node n to be connected to an old one i will depend not only on the number of edges k_i, which i already possesses, but also on the distance d_{in} between them,

$$\Pi(k_i, d_{in}, \alpha) = \frac{k_i \cdot d_{in}^{-\alpha}}{\sum_j k_j \cdot d_{jn}^{-\alpha}}. \tag{2}$$

Here the sum in the denominator goes over all nodes in the system except the newly introduced one, and α is a real non-negative parameter describing the distance penalties. For large α, the probability of a connection between two distant nodes is very small. On the other hand, for a very small α the probability is almost independent of the distance. In the case $\alpha=0$, our model reduces to the genuine scale-free one. Note that our model is to some extent also scale-free: the connection probabilities depend only on the *relative* distances.

Our initial condition is slightly different from that of Barabási and Albert, where the initial m_0 nodes are not connected: in our case all nodes introduced at $t=0$ have exactly one edge, which allows us to use Eq. (2) from the very beginning. This simplifies the algorithm, since we do not have to distinguish between the initial and the further steps. The only difference with the genuine SF construction is that at time t one has $mt+m_0/2$ (instead of mt) edges present; hence, the asymptotic behavior of both models for $t \to \infty$ is the same.

Three examples of the evolving networks of this kind are given in Fig. 1. Here $m=3$, $L=10^6$, $N=105$, and $m_0=6$ (so that all three networks have exactly 300 edges). Three different values of α were used: $\alpha=0.0$ (scale-free model), $\alpha=1.5$ and $\alpha=15.0$. Note that increasing the value of α leads to marked changes in the topology of the network. Figure 1(a) corresponds to a genuine scale-free construction and exhibits a lot of long bonds connecting distant sites. On the other hand, only few such bonds are present in Fig. 1(c).

In our further simulations we use a lattice of $L=2\times10^7$ sites; the maximum number of the introduced nodes is $N=2\times10^5$. All simulation results are based on the average

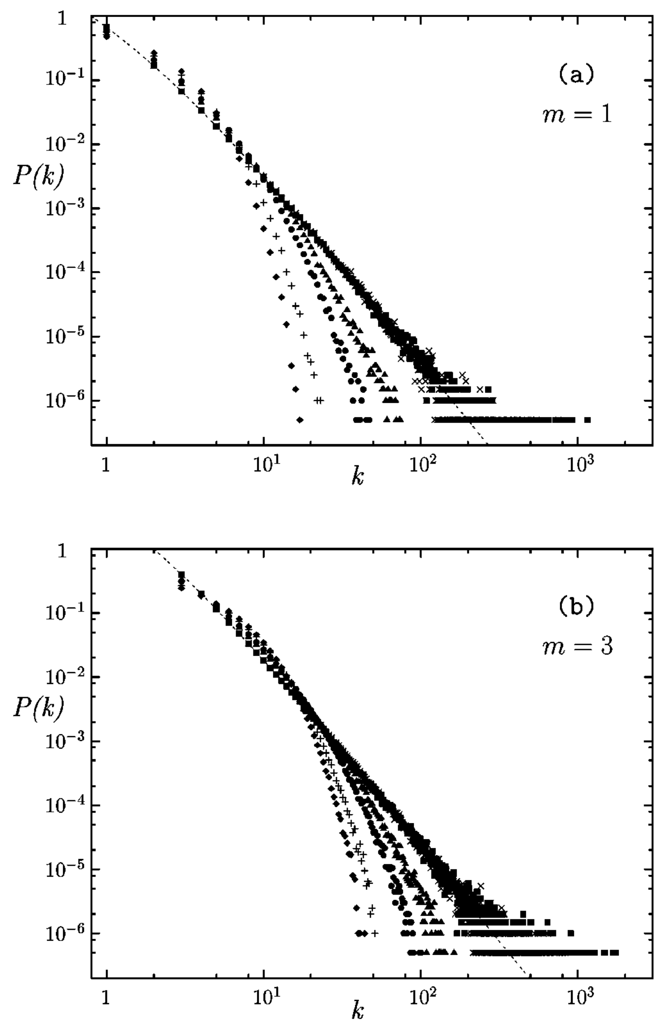

FIG. 2. The degree distribution $P(k)$ for different values of α and for $m=1$ (a) and $m=3$ (b). The values of α are $\alpha=0$ (squares), $\alpha=0.8$ (crosses), $\alpha=1.5$ (triangles), $\alpha=2$ (filled circles), $\alpha=5$ (plusses), and $\alpha=45$ (diamonds). The dashed lines correspond to the theoretical curve for the scale-free model, (Refs. [1,2,13]).

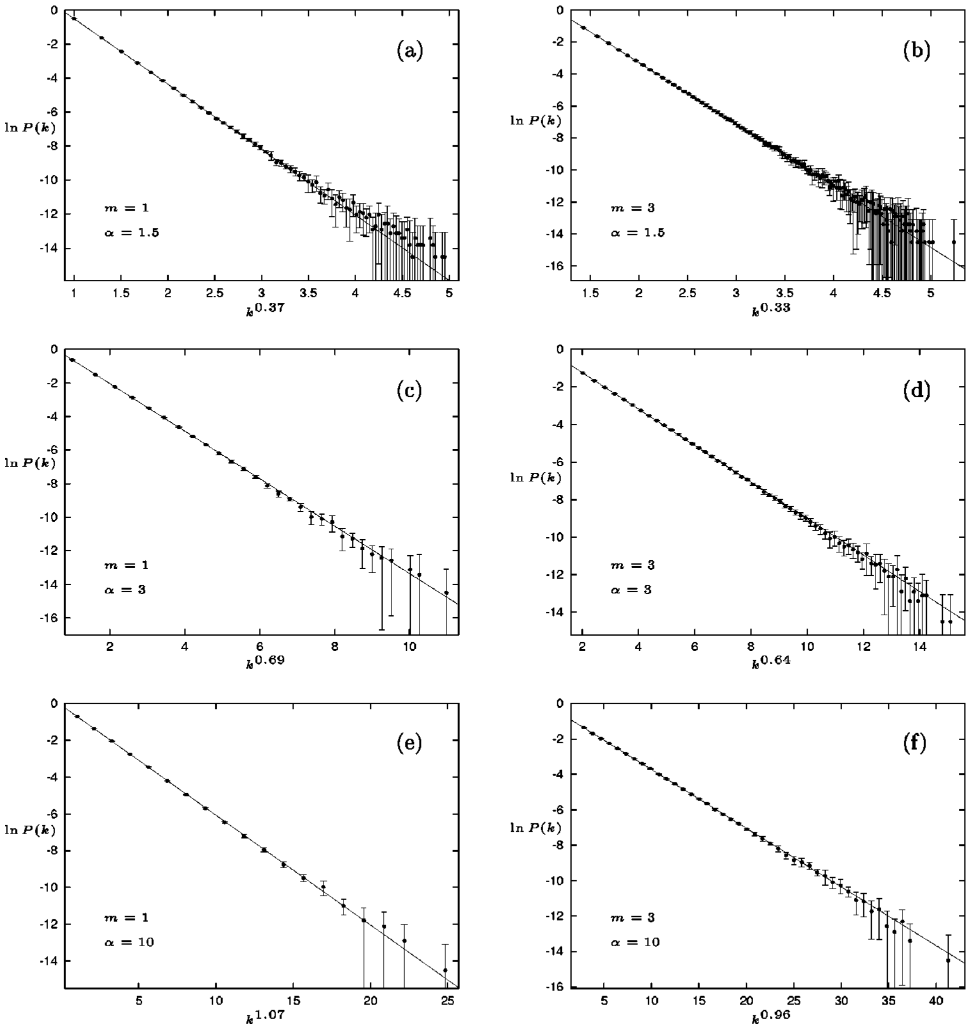

FIG. 3. Shown is $\ln P(k)$ as a function of k^{γ}, where γ is the output of the fit, Eq. (3) (see text for details). The parameters are (a) $m=1$, $\alpha=1.5$, $\gamma=0.37$; (b) $m=3$, $\alpha=1.5$, $\gamma=0.33$; (c) $m=1$, $\alpha=3$, $\gamma=0.69$; (d) $m=3$, $\alpha=3$, $\gamma=0.64$; (e) $m=1$, $\alpha=10$, $\gamma=1.07$; and (f) $m=3$, $\alpha=10$, $\gamma=0.96$.

over ten realizations of this structure. The error bars on Figs. 3–5 correspond just to this ensemble average. The simulations are done for several values of α and for two values of m, the number of the outgoing bonds: $m=1$ and $m=3$; $m_0=2m$.

One of the prominent features of the scale-free model is that the distribution of the degrees of the nodes decays as a power law, i.e., as $P(k)\sim k^{-\gamma}$, with $\gamma=3$. This corresponds to the fact that the mean number of connections per site exists, but its dispersion diverges. Let us discuss now how

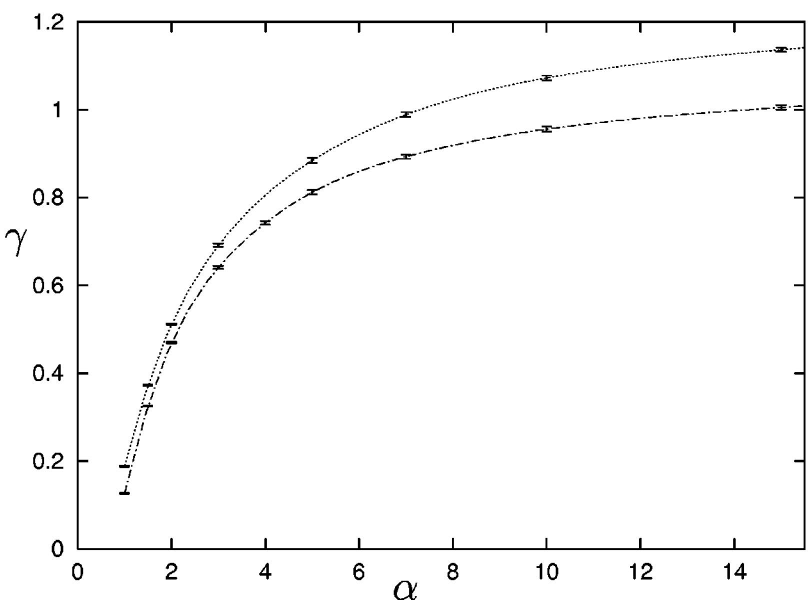

FIG. 4. The parameter γ as a function of α. The upper dependence corresponds to $m=1$, and the lower one to $m=3$. The lines are drawn as a guide for eyes.

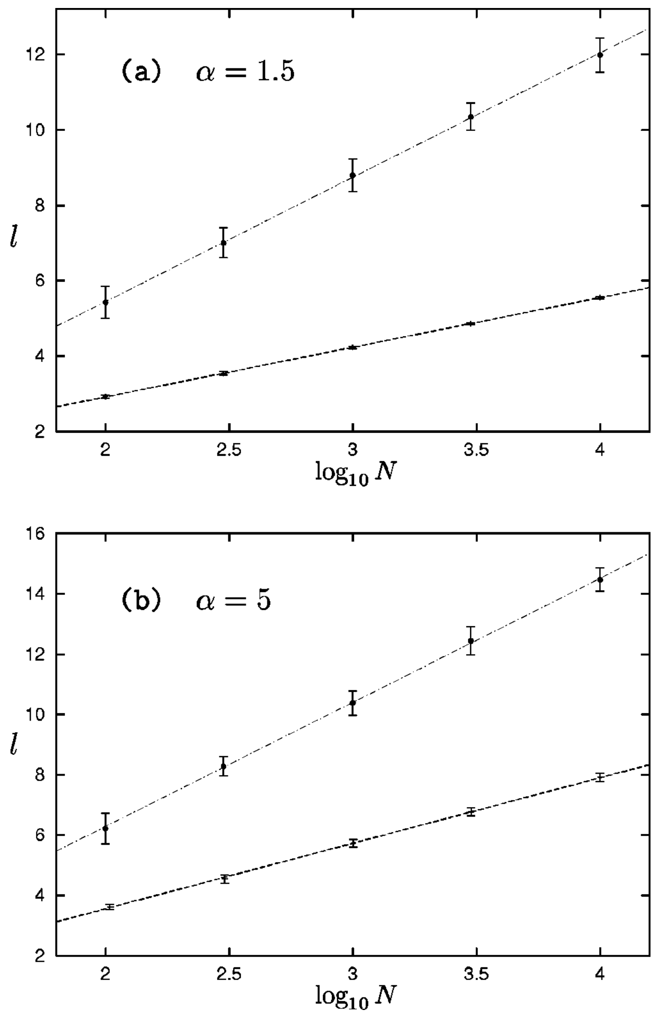

FIG. 5. The diameter of a network as a function of the number of sites N. Panel (a) corresponds to $\alpha=1.5$ and panel (b) corresponds to $\alpha=5$. The upper lines in each panel are those for $m=1$, the lower lines correspond to $m=3$. Note the logarithmic scale.

this distribution changes if the long-range connections are penalized. In Fig. 2 we plot the probability distribution of k for different values of α on double logarithmic scales. One readily infers that for all $0<\alpha<1$ no important differences with the scale free model ($\alpha=0$) can be detected: in any case the asymptotic behavior of $P(k)$ is well described by $P(k)\sim k^{-3}$. The distributions seem to be almost identical; however, small, but statistically significant, deviations can be detected for small k values. At $\alpha\simeq 1$ the degree distribution shows a pronounced change in its behavior and ceases being a power law; now the behavior of the model with distance penalties is quite different.

Let us concentrate on the case $\alpha>1$ and try to describe the shape of the degree distribution under such conditions. The analysis of the simulations suggests that the corresponding mathematical expression could be a stretched-exponential function of the form

$$P(k)=a\exp(-bk^{\gamma}), \tag{3}$$

where the parameters a, b, and γ depend on α and m. To obtain the values of these parameters and to analyze the goodness of this fitting function, we have fitted the data to Eq. (3) using the nonlinear least-squares Levenberg-Marquardt algorithm [24], taking into consideration the error bars as coming out of ten realizations of each situation. The data is replotted together with the outcomes of the fits in Fig. 3 on the scales in which the fitting function, Eq. (3), is represented by a straight line. One takes k^{γ} as the abscissa and $\ln P(k)$ as the ordinate of the graph. Figure 3 shows that such a fit (straight line) is surprisingly good.

The values of the exponent γ are shown as a function of α ($\alpha>1$) in Fig. 4 for the two different situations corresponding to $m=1$ and $m=3$. We see that γ monotonically grows with α, and that the dependences for $m=1$ and $m=3$ differ, i.e., the $\gamma(\alpha)$ dependence is nonuniversal.

We note that in related models of growing networks another form of degree distribution appears: an exponentially damped power law [25],

$$P(k)=ak^{\gamma}\exp(-bk). \tag{4}$$

We also tested this fit function and found out that it gives a good fit for larger α values, but is definitely inferior to our fit, Eq. (3), for $1<\alpha<3$.

A growing network with disadvantaged long bonds is a very interesting hierarchical construction. Thus, for large α, the strong correlation between the age of the connection and its length exists. The old connections, made when the nodes were sparse, are typically long, while the younger connections get shorter and shorter, since more sites in the immediate vicinity of a newly introduced site can be found. The simulations show that for a large value of α, the nodes are almost surely connected to their nearest neighbors. On the other hand, the old, long-range connections are of great importance for the overall topology of the lattice, since they guarantee that for any α the network is a small-world one.

In Fig. 5 we plot the mean number of connections between each two nodes of the network for two different values of α ($\alpha=1.5$ and $\alpha=5$) and for the two values $m=1$ and $m=3$ as a function of the network size N. The algorithm here is trivial: starting from a node (labeled 0) we pass to all nodes connected to it (nodes of the first generation, labeled 1), then to nodes of the second generation (labeled 2), etc., untill all nodes are labeled. The mean distance between this node (labeled 0) and any other given node of the network is then the sum of all values of these labels divided by $N-1$. This procedure is repeated for each node, and the overall mean value, the so-called path diameter of the network (l), is evaluated. The error bars of the figure correspond to the average of the mean diameters over ten realizations of the network. Figure 5 shows that the mean diameter of the network grows linearly in $\ln N$, i.e., it shows the typical small-world behavior. This behavior is preserved for all tested values of α; the largest value tested was $\alpha=45$, which, for $m=1$, corresponds to a practically sure connection of a newly introduced node to its nearest neighbor. The high-α networks closely resemble the simple small-world constructions [26].

Let us summarize our findings. We considered a growing network, whose growth algorithm is based, as in the SF construction, on a preferential attachment of the newly introduced nodes to the highly connected old ones. However, here the too long connections are disadvantaged by introducing penalties. Thus, the probability to connect two nodes separated by a distance d is proportional to $d^{-\alpha}$, where α is a variable parameter. We found out that for $\alpha<1$ the degree distribution $P(k)$ decays, as in the SF model, as $P(k)\sim k^{-3}$, whereas for $\alpha>1$ a stretched exponential form $P(k)=a\exp(-bk^{\gamma})$ gives an extremely good description of this distribution. On the other hand, the small-world property is preserved at all checked values of α.

We note that a similar model was considered by Manna and Sen [27]. However, more attention was payed to two-dimensional systems and to other properties of the network than to those discussed here.

The authors are grateful to F. Jasch for fruitful discussions and to M. Obermayr for the valuable technical assistance. Financial support by the Fonds der Chemischen Industrie is gratefully acknowledged. We are grateful to the authors of [27] for bringing this work to our attention prior to publication.

[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] S. Jespersen, I.M. Sokolov, and A. Blumen, J. Chem. Phys. **113**, 7652 (2000).
[4] M.E.J. Newman, Proc. Natl. Acad. Sci. U.S.A. **98**, 404 (2001).
[5] M.E.J. Newman, Phys. Rev. E **64**, 016131 (2001).
[6] M.E.J. Newman, Phys. Rev. E **64**, 016132 (2001).
[7] H. Jeong, B. Tombor, R. Albert, Z.N. Oltvai, and A.-L. Barabási, Nature (London) **407**, 651 (2000).
[8] D.A. Fell and A. Wagner, Nat. Biotechnol. **18**, 1121 (2000).
[9] H. Jeong, S.P. Mason, A.-L. Barabási, and Z.N. Oltvai, Nature (London) **411**, 41 (2001).
[10] M. Faloutsos, P. Faloutsos, and C. Faloutsos, Comput. Commun. Rev. **29**, 251 (1999).
[11] F. Liljeros, C.R. Edling, L.A. Nunes Amaral, H. Stanley, and Y. Aberg, Nature (London) **411**, 907 (2001).
[12] R. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **401**, 130 (1999).
[13] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[14] S.N. Dorogovtsev, J.F.F. Mendes, and A.N. Samuhkin, Phys. Rev. Lett. **85**, 4633 (2000).
[15] L. Kullmann and J. Kertész, Phys. Rev. E **63**, 051112 (2001).
[16] P.L. Krapivsky, S. Redner, and F. Leyvraz, Phys. Rev. Lett. **85**, 4629 (2000).
[17] S.N. Dorogovtsev and J.F.F. Mendes, Phys. Rev. E **63**, 025101 (2001).
[18] A.-L. Barabási, H. Jeong, E. Ravasz, Z. Néda, A. Schubert, and T. Vicsek, e-print cond-mat/0104162.
[19] H. Jeong, Z. Néda, and A.-L. Barabási, e-print cond-mat/0104131.
[20] L.A.N. Amaral, A. Scala, M. Barthélémy, and H.E. Stanley, Proc. Natl. Acad. Sci. U.S.A. **97**, 11149 (2000).
[21] S.N. Dorogovtsev and J.F.F. Mendes, Phys. Rev. E **62**, 1842 (2000).
[22] G. Bianconi and A.-L. Barabási, Europhys. Lett. **54**, 436 (2001).
[23] R. Albert and A.-L. Barabási, Phys. Rev. Lett. **85**, 5234 (2000).
[24] B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling, *Numerical Recipes: The Art of Scientific Computing* (Cambridge University Press, Cambridge 1985), p. 675.
[25] M.E.J. Newman, e-print cond-mat/0201433.
[26] D.J. Watts and S.H. Strogatz, Nature (London) **393**, 440 (1998).
[27] S.S. Manna and P. Sen, e-print cond-mat/0203216.