Scale-free networks from a Hamiltonian dynamics

M. Baiesi[1] and S. S. Manna[1,2]
[1]*INFM, Dipartimento di Fisica, Università di Padova, I-35131 Padova, Italy*
[2]*Satyendra Nath Bose National Centre for Basic Sciences, Block-JD, Sector-III, Salt Lake, Kolkata 700098, India*
(Received 3 May 2003; revised manuscript received 9 July 2003; published 10 October 2003)

Contrary to many recent models of growing networks, we present a model with fixed number of nodes and links, where a dynamics favoring the formation of links between nodes with degree of connectivity as different as possible is introduced. By applying a local rewiring move, the network reaches equilibrium states assuming broad degree distributions, which have a power-law form in an intermediate range of the parameters used. Interestingly, in the same range we find nontrivial hierarchical clustering.

DOI: 10.1103/PhysRevE.68.047103 PACS number(s): 89.75.Hc, 05.10.−a, 05.65.+b

In their theory of random graphs (RG), Erdös and Rényi showed that these graphs, composed of N vertices (or nodes), connected probabilistically by a set of edges have several interesting properties [1]. Among them the most striking one is the slow rate of growth ($\sim \log N$) of the diameter of the giant component. This ''small world'' property is very important in connected networks represented by single component graphs, since it reflects the efficiency of the network for transport or communications [2]. Over last few years it is becoming increasingly evident that most real-world networks have indeed small world properties [3,4], e.g., electronic communication networks like Internet [5], World-Wide Web [6], social networks of acquaintances [7], and of collaborations [8].

On the other hand, some important properties distinguish real-world networks from RG, motivating the rapid growth of interest in this field. Many real-world networks have broad nodal degree distributions, $P(k)$ (the degree of a node is the number k of links meeting at that node) often characterized by a power tail $P(k) \sim k^{-\gamma}$ that indicates a scale-free (SF) character of the network [4]. Moreover, in real-world networks one observes a high degree of clustering, which measures the local correlations among the links of the network and implies that neighbors of a node are more likely to be neighbors [2] (this feature has also been associated to the term small world [2]). The clustering often scales with the degree of the relative node. This is connected to a hierarchical organization of the network [9], where clustered blocks connect to form larger units, etc.

To introduce the correlations between nodes that distinguish scale-free character of networks (SFN's) from RG's, in the past years SFN's have been extensively modeled by growing networks in which a preferential attachment (PA) rule shapes the nodes degree [3,4], (i.e., each new node is linked to an old one with a probability proportional to the degree of the old node [10]). However, biological networks, including food webs [11], metabolic networks [12,13], and protein-protein interaction networks [14,15] display the features listed above, although both the PA and the growing process are debatable in these cases. For example, for protein-protein networks they have proposed both growing network models without PA [16] and models with a dominant stationary, asymmetric PA [17]. In food webs, where links represent the prey-predator relations, the PA and growth are particularly unsuitable to describe the situation. Thus, in order to achieve a better understanding of the principles shaping a part of the real networks, it is worth to spend some effort to discover dynamics leading to nongrowing SFN's. It has been already shown that models with fixed number of nodes do not require linear PA, but SF distributions arise from an algebraic PA rule including an exponent which can vary in a wide range [18].

In this paper, we address the particular problem of whether SFN's can arise from mechanisms excluding both PA and growth. Recent works [19–21] proposed theories of networks at equilibrium. In some of these cases [20,21] a SFN can be generated simply by choosing the desired degree distribution to be SF. On the other hand, different works obtained SFN without plugging in an *a priori* degree distribution [22–25]. In the spirit of the latter strategy, we propose an example of equilibrium network with Hamiltonian that can yield hierarchical SFN's. The energy function depends on the degrees of the nodes and of their neighbors [19]. Since it favors connections between nodes with degrees as different as possible, it leads to networks with disassortative mixing [26]. Furthermore, the simulation is implemented by using a local rewiring rule. Dynamics of this kind appear natural for biological networks, which indeed are disassortative [13,26]. In particular, in food webs we expect each species to find it inconvenient to interact with similar (and competing) ones. We notice that in food webs both exponential and SFN's are found [11], as we have in our model.

We consider a connected network, represented by a single component undirected graph, composed of N nodes connected by L undirected links (edges). The network topology is uniquely determined by its adjacency matrix $\mathbf{c}$, such that $c_{ij}=1$ if nodes i and j are linked, and 0 otherwise. The degree of a node i is indicated as $k_i=\Sigma_{j\neq i} c_{ij}$. We define an energy associated with a link between the ith and the jth nodes having the following form:

$$\epsilon_{ij} = -c_{ij}\left(1-\frac{\min\{k_i,k_j\}}{\max\{k_i,k_j\}}\right). \tag{1}$$

This equation implies that the energy of a link decreases with the difference in nodal degrees at the two ends of the link

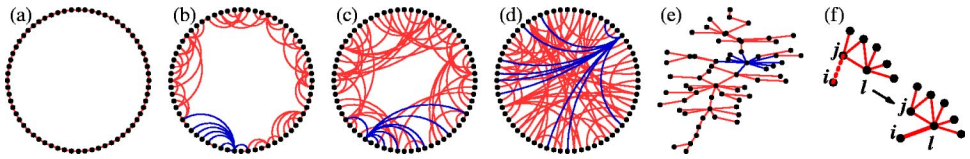

FIG. 1. (Color online) (a)–(d) Example with $N=L=64$, $\bar{k}=2$ at $\alpha=2$. Snapshots after t iterations of the LRM (f), with (a) $t=0$, (b) $t=4N$, (c) $t=12N$, and (d) $t=N^2$. Blue (darker in gray scale) links meet at the node with the highest degree. (e) Different layout of (d). (f) LRM described in the text.

and it contributes no energy when the link connects two nodes of same degrees. The Hamiltonian of a network configuration $\mathcal{G}$ is then

$$H(\mathcal{G})=\sum_{i<j}\epsilon_{ij}. \tag{2}$$

To generate the initial connected network, we first add $L_0=N$ links to form a graph with the topology of a ring [see Fig. 1(a)]. The remaining L-L_0 links are added sequentially and randomly, to connect unlinked nodes. The network is evolved by using a local rewiring move (LRM), depicted in Fig. 1(f). A set of three nodes is randomly selected. First, a node i is selected with probability $1/N$, second, the node j which is a neighbor of i is selected with probability $1/k_i$, and finally the node $l\neq i$ which is a neighbor of j is selected with probability $1/(k_j-1)$. If $c_{il}=0$, a LRM attempts to delete the i-j link of the graph $\mathcal{G}$ and introduce the i-l link to obtain the new graph $\mathcal{G}'$ with a probability

$$p=\min\{1,e^{-\alpha[H(\mathcal{G}')-H(\mathcal{G})]}\}, \tag{3}$$

where α is a tunable parameter. For $\alpha=0$ the LRM is always accepted [27]. In this case the difference between the typical graphs and RG's is reflected in the degree distribution. Using a master equation approach [28], one can show that the degree distribution $P(k)$ indeed decays exponentially, as shown numerically in Fig. 2. In the opposite limit of $\alpha\to\infty$, LRM strongly favors connecting nodes with degrees as different as possible. As a result, usually graphs have several high degree nodes (hubs) connected to many other monodegree nodes (leaves). Notice that the LRM cannot split the network in disjoint components.

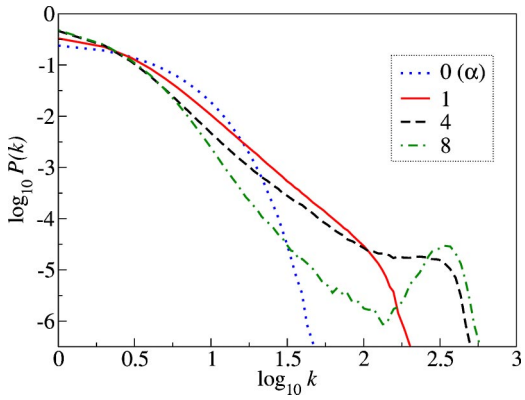

FIG. 2. (Color online) Log-log plot of the degree distributions for $N=4096$, $\bar{k}=4$, and at various α's.

In the probability (3), the introduction of the Hamiltonian (2) and the LRM have been chosen for their simplicity and for their analogy with usual rules of equilibrium statistical mechanics, but they do not yield an equilibrium distribution of Eq. (2) at an inverse temperature α. However, one can see that the above implementation gives the canonical distribution of configurations with weights $\Pi_{i=1}^{N}(k_i-1)!e^{-\alpha H(\mathcal{G})}$, hence allowing a description of the system in terms of ensemble of connected networks at the equilibrium given by the Hamiltonian $H_{\rm eq}(\mathcal{G})\equiv-\alpha^{-1}\Sigma_{i=1}^{N}\ln(k_i-1)!+H(\mathcal{G})$, at temperature α^{-1}. Thus, $H(\mathcal{G})$ can be thought as an interaction term from which we expect the arising of complex correlations in the network. Due to $H(\mathcal{G})$, the fraction of second neighbors of the node i represented by the first neighbors of node j and of node l [in the LRM, Fig. 1(f)] belongs to the subset of nodes contributing to the energy balance in a LRM. The nontrivial build up of the correlations necessary to obtain SFN's should require $\gamma<3$, meaning an average number of second neighbors $\sim N$. On the contrary, $\gamma>3$ would forbid the LRM to "feel" the global structure of the network, hardly giving a fine self-tuning of the network correlations. As shown below, indeed we find SFN's with $2<\gamma<3$.

We apply the described dynamical rules to networks composed by N up to 8192 nodes and $L=N\bar{k}/2$, where the average degree $\bar{k}$ is fixed. Thus, we consider sparse networks in the limit of large N, where the number of links L is much smaller than the maximum number $L_N^{\max}=N(N-1)/2$ of possible links in a N-clique. This is motivated by the case of most of the real networks, where typically a link between two nodes is an expensive or rare object. In our case, we have mainly used $\bar{k}=4$, such that $L=2N$.

After a large number $\tau\approx 2N^2$ of LRM attempts on the initial configurations with $\bar{k}=4$, we observe a significant reorganization of the entire network structure. For $t\gtrsim\tau$ iterations of the LRM, equilibrium is reached, as indicated by the stable shape of the distribution $P_{\alpha,N}(k)$ of the degrees (see Fig. 2). The most interesting region is $0.8\lesssim\alpha\lesssim 4.0$, where $P_{\alpha,N}(k)$ appear as power laws. By increasing α, their slopes decrease and their shape changes at large k's, where a shoulder grows, indicating the enhanced tendency in the network to form hubs, as expected. For $\alpha\gtrsim 4.0$ the fraction of hubs is even more consistent and the shoulder at high degrees in the $P_{\alpha,N}(k)$ is substituted by a bump (see, for instance, the curve for $\alpha=8$ in Fig. 2). Contrary to other models [29], an analysis of the mean degree of the largest hub indicates that here it does not attract a finite fraction of links, for $N\to\infty$.

We now focus on the range $0.8\lesssim\alpha\lesssim 4.0$, where the degree distributions are broad and power-law like. To support

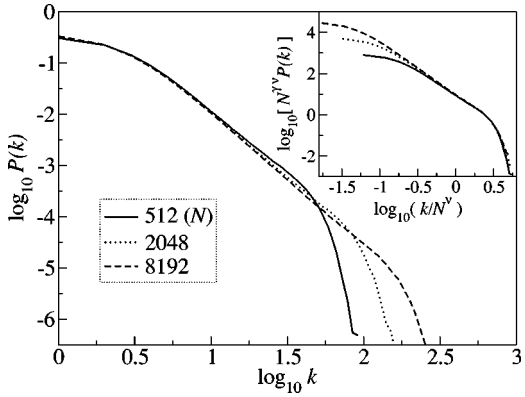

FIG. 3. Degree distributions for $\alpha=1$. Inset: distributions rescaled (4) with the γ and ν values quoted in Table I.

this SF picture we plot Fig. 3, in which a remarkable feature of this model is evident, namely, the self-similar shape of the $P_{\alpha,N}(k)$ for fixed α. This is consistent with a finite size form

$$P_{\alpha,N}(k) \simeq k^{-\gamma(\alpha)} f(k/N^{\nu(\alpha)}), \tag{4}$$

where $f(x)$ is a scaling function giving a cutoff for sufficiently large values of x, and $f(x) \sim$ const for $x \to 0$. In order to quantify the SF nature of the networks and to support the proposed scaling (4), we perform a finite size analysis, first extrapolating the value of the slope distributions at fixed α and for $N \to \infty$. We show the results in Table I. For each α, the value of γ so obtained is the starting point for attempting a data collapse, by using a rescaling of the form $N^{\gamma\nu}P(k)$ vs k/N^{ν}. The values of γ that give better collapses (see Table I) turn out to be close to the extrapolated values, such that the whole picture is consistent. In addition, we note that the data collapses get worse close to the boundaries of the SF region $0.8 \lesssim \alpha \lesssim 4.0$, while outside this range we cannot make good rescalings. This supports our first, subjective delimitation of the SF region (we stress that the case $\bar{k}=4$ is treated).

It is interesting to examine the other stylized features of networks. In Fig. 4 we plot the mean diameter $\mathcal{D}(N)$ (the maximum of the shortest paths between any nodes of a network) as a function of $\log_{10} N$ for the some representative α values. The curves are consistent with a scaling $\mathcal{D}(N) \sim A + B \log_{10} N$ (it seems to be sublogarithmic [30] for high α). Thus, not surprisingly the small world picture is recovered also in our model.

Given a node i connected with k_i neighbors, if m_i is the number of links between these neighbors, one can quantify the (local) degree of clustering by the clustering coefficient

TABLE I. Numerical evaluation of the exponents of Eq. (4).

α	0.8	1	2	3	4
γ [a]	2.9(2)	2.8(2)	2.4(1)	2.3(2)	2.1(3)
γ [b]	3.0	2.8	2.3	2.2	2.2
ν [b]	0.4	0.45	0.55	0.58	0.58

[a]Values extrapolated for $N \to \infty$.
[b]Values that give the best data collapse.

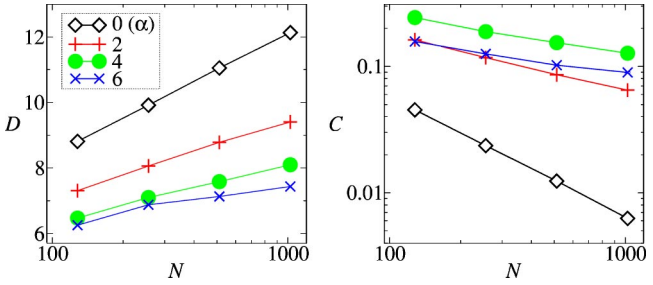

FIG. 4. (Color online) Mean diameter $\mathcal{D}$ and mean clustering coefficient C as a function of N (linear-log and log-log plot, respectively) for the same set of α values.

$C_i \equiv m_i / L_{k_i}^{\max}$. The mean clustering coefficient C is then the average of the C_i over all the nodes of all graph realizations at a given α. In Fig. 4 we also plot C as a function of N, in a log-log scale. The plots are compatible with $C \sim N^{-\sigma(\alpha)}$ for $N \to \infty$, with σ ranging from ≈ 1 to ≈ 0.25 for $0 \leqslant \alpha \leqslant 6$. For the studied N's, we notice that the highest clustering is found close to $\alpha=4.0$, while it is smaller for high and low α.

Our model also shows the power-law dependence

$$C(k) \sim k^{-\beta}, \tag{5}$$

indicating that nodes with few links are typically well clustered while hubs hardly are related to high clustering. The concept of modularity was introduced to account for the hierarchical clustering found in many networks [9,31]. In these contexts, networks are built of rather well identifiable clusters, which are themselves composed by clustered subunits, and so on. While in Fig. 5 we see that $\beta=0$ for $\alpha=0$, as for RG, for increasing α a scaling (5) takes place in a nonnegligible interval of k, with a β raising with α up to values close to $\beta=1$. Hence, in a present issue on whether $\beta=1$ is universal [9], our result is in favor of the nonuniversal character of this exponent.

In summary, we have described a static network model with a dynamics favoring networks with disassortative mixing and high clustering, where at least three phases can be identified by increasing the parameter α, respectively: exponential, scale-free, and hub-leaves regimes. The scale-free regime has a signature of modularity (the clustering coeffi-

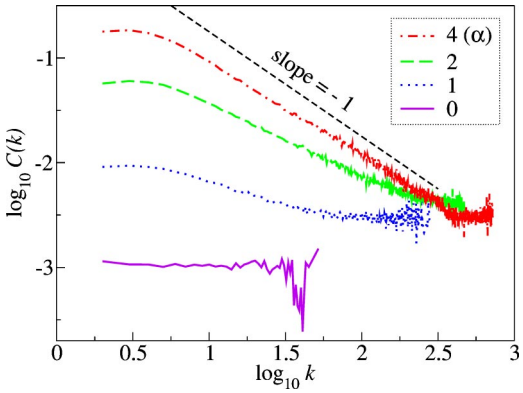

FIG. 5. (Color online) Clustering coefficient as a function of the degree, for four values of α and $N=8192$.

cient scales as a power law of the degree with a nontrivial exponent), and the exponent γ is comprised in the interesting range $2<\gamma<3$. Thus, many of the characteristics displayed by real networks, usually associated with growing networks with preferential attachment, can be obtained as well in networks with a fixed number of nodes by using a random rewiring that does not require preferential attachment.

We gratefully acknowledge useful discussions with A.L. Stella. M.B. acknowledges the support of MIUR-COFIN01 and INFM-PAIS02.

[1] P. Erdös and A. Rényi, Publ. Math. (Debrecen) **6**, 290 (1959).
[2] D.J. Watts and S.H. Strogatz, Nature (London) **393**, 440 (1998); D J. Watts, *Small Worlds: The Dynamics of Networks between Order and Randomness* (Princeton, New York, 1999).
[3] A.-L. Barabási, *Linked: The New Science of Networks* (Perseus, Cambridge, 2002); S.N. Dorogovtsev and J.F.F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW* (Oxford University Press, Oxford, 2003).
[4] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002); S.N. Dorogovtsev and J.F.F. Mendes, Adv. Phys. **51**, 1079 (2002); M.E.J. Newman, e-print cond-mat/0303516.
[5] M. Faloutsos, P. Faloutsos, and C. Faloutsos, Comput. Commun. Rev. **29**, 251 (1999).
[6] S. Lawrence and C.L. Giles, Science **280**, 98 (1998); Nature (London) **400**, 107 (1999); R. Albert, H. Jeong, and A.-L. Barabási, *ibid.* **401**, 130 (1999).
[7] *The Small World*, edited by M. Kochen (Ablex, Norwood, 1989).
[8] M.E.J. Newman, Proc. Natl. Acad. Sci. U.S.A. **98**, 404 (2001).
[9] E. Ravasz and A.-L. Barabási, Phys. Rev. E **67**, 026112 (2003).
[10] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[11] J.A. Dunne, R.J. Williams, and N.D. Martinez, Proc. Natl. Acad. Sci. U.S.A. **99**, 12 917 (2002).
[12] E. Ravasz *et al.*, Nature (London) **297**, 1551 (2002).
[13] S. Maslov and K. Sneppen, Science **296**, 910 (2002).
[14] Uetz *et al.*, Nature (London) **423**, 623 (2000); I. Xenarios *et al.*, Nucleic Acid Res. **28**, 289 (2000); T. Tashiro *et al.*, Proc. Natl. Acad. Sci. U.S.A. **99**, 1143 (2000).
[15] A.-L. Barabási *et al.*, in Sitges Proceedings on Complex Networks, 2004 (unpublished).
[16] R.V. Solé, R. Pastor-Satorras, E.D. Smith, and T. Kepler, Adv. Complex Syst. **5**, 43 (2002); A. Vasquez, A. Flammini, A. Maritan, and A. Vespignani, ComPlexUs **1**, 38 (2003).
[17] J. Berg, M. Lässig, and A. Wagner, e-print cond-mat/0207711.
[18] G. Mukherjee and S.S. Manna, Phys. Rev. E **67**, 012101 (2003).
[19] J. Berg and M. Lässig, Phys. Rev. Lett. **89**, 228701 (2002).
[20] M. Bauer and D. Bernard, e-print cond-mat/0206150.
[21] Z. Burda and A. Krzywicki, e-print cond-mat/0207020.
[22] S. Valverde, R.F. Cancho, and R.V. Solé, Europhys. Lett. **60**, 512 (2002).
[23] J.P.K. Doye, Phys. Rev. Lett. **88**, 238701 (2002).
[24] G. Caldarelli, A. Capocci, P. De Los Rios, and M.A. Munoz, Phys. Rev. Lett. **89**, 258702 (2002).
[25] M. Boguna and R. Pastor-Satorras, e-print cond-mat/0306072.
[26] M.E.J. Newman, Phys. Rev. Lett. **89**, 208701 (2002).
[27] For a similar model, see Ph. Blanchard, A. Ruschhaupt, and T. Krueger, e-print cond-mat/0304563.
[28] M. Baiesi (unpublished).
[29] G. Bianconi and A.-L. Barabási, Phys. Rev. Lett. **86**, 5632 (2001).
[30] R. Cohen and S. Havlin, Phys. Rev. Lett. **90**, 058701 (2003).
[31] G. Szabó, M. Alava, and J. Kertész, e-print cond-mat/0208551.