Random initial condition in small Barabasi-Albert networks and deviations from the scale-free behavior

Paulo R. Guimarães, Jr.,[1,3] Marcus A. M. de Aguiar,[2] Jordi Bascompte,[3] Pedro Jordano,[3] and Sérgio Furtado dos Reis[4]
[1]*Instituto de Biologia, Universidade Estadual de Campinas (UNICAMP), Caixa Postal 6109, 13083-970 Campinas, SP, Brazil*
[2]*Instituto de Física "Gleb Wataghin," Universidade Estadual de Campinas (UNICAMP), Caixa Postal 6165, 13083-970 Campinas, SP, Brazil*
[3]*Integrative Ecology Group, Estación Biológica de Doñana, CSIC, Apdo. 1056, E-41080 Sevilla, Spain*
[4]*Departamento de Parasitologia, Instituto de Biologia, Universidade Estadual de Campinas (UNICAMP), Caixa Postal 6109, 13083-970, Campinas, SP, Brazil*
(Received 3 November 2004; published 18 March 2005)

Barabasi-Albert networks are constructed by adding nodes via preferential attachment to an initial core of nodes. We study the topology of small scale-free networks as a function of the size and average connectivity of their initial random core. We show that these two parameters may strongly affect the tail of the degree distribution, by consistently leading to broad-scale or single-scale networks. In particular, we argue that the size of the initial network core and its density of connections may be the main responsible for the exponential truncation of the power-law behavior observed in some small scale-free networks.

DOI: 10.1103/PhysRevE.71.037101 PACS number(s): 89.75.Hc, 87.23.−n, 05.45.−a, 87.10.+e

I. INTRODUCTION

Complex networks describe a large number of social, physical, and biological systems [1–9]. The very basic organizing principles of complex networks are encoded, in some level, in network topology [1,3,9,10]. For example, their degree distribution, which is the cumulative probability distribution of the number of edges per node, captures in quantitative terms some rules that govern the connection of nodes in growing networks [1,5,11]. The Barabasi-Albert (BA) model for growing networks proposes that the two main organizing principles acting during the buildup of complex network are growth and "preferential attachment." Under this mechanism there is a nonuniform probability with which a new node connects to an existing node of the network, which increases with the number of connections of that node [11]. The BA model generates a degree distribution that decays as a power law, implying that the system does not have a particular scale (scale-free networks) [11]. Although several physical and biological systems are indeed scale free [3,11,12], there are several examples of complex networks, such as the small mutualistic networks of interactions among plants and animals [7], in which an exponential truncation of the power-law behavior predominates for large degrees [1,5]. These networks are called broad-scale networks and are more homogeneous than scale-free networks (Fig. 1). This observed truncation in power-law behavior can be explained by the small size of these networks [10,13] or by mechanisms such as the addition of links limited by aging or connection costs [1], forbidden links [7], and information filtering [1,5]. These mechanisms suggest that, in broad-scale networks, preferential attachment is constrained by node characteristics operating during the network evolution.

However, size effects and growth constraints might not be the sole responsibility for the exponential tail of the degree distribution. In this paper we argue that, for relatively small networks, the initial set of nodes over which the network evolves has strong effects on the tail of the degree distribution. To understand why this is so we recall that the BA network is constructed from a small number of nodes which we call the "initial core." Then, at each step a new node is added and connected to the already existing ones following the preferential attachment rule [11]. Although the original model does not say anything about connections between the initial nodes, later work has assumed that the initial core is totally connected [5,14], totally disconnected [15], or randomly connected [16]. In this paper we demonstrate that, for small networks, a randomly connected initial core in BA networks consistently generates a truncation of the degree distribution and that the truncation depends on the relative size of this random initial core. Moreover, we show that highly connected large initial cores generate degree distributions that markedly deviate from the power-law regime. This indicates that the tail of the degree distribution might contain information about the genesis of the network. We note that the characterization of small networks has been pointed as one of the leading questions in network research [17]. This work aims at contributing in this direction.

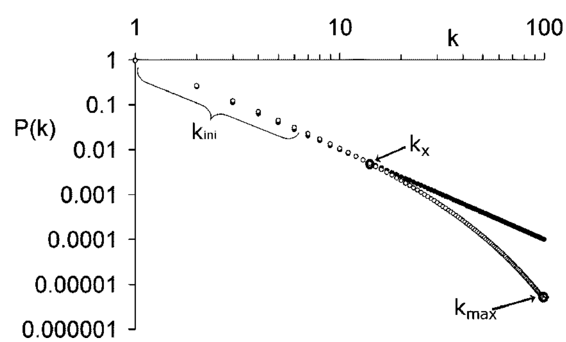

FIG. 1. Illustration of the exponential truncation of the power-law behavior of degree distributions. Open circles are recorded values for $P(k)$ for a hypothetical degree distribution. Solid circles are values of $P(k)$ predicted by a power-law distribution, computed using the first k_{ini} values of k. k_x indicates the lowest value where the "observed" $P(k)$ departs from the predicted power-law behavior. k_{max} is maximum recorded degree.

II. BA MODEL WITH A RANDOM INITIAL CORE

The network's random initial core is defined as follows: at time $t=0$, one creates m_0 nodes and connects each pair of nodes with constant probability p. Thus, this initial core of nodes is an Erdos-Renyi (ER) random graph [3]. Then, at each time step, a new node with $m \leqslant m_0$ edges is added to the network. To incorporate preferential attachment, we assume that the probability Π that a new node will be connected to node i depends on the degree k of that node, so that

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j}. \tag{1}$$

To test numerically the effect of the size of the ER initial core on the degree distribution, we simulate the evolution of different networks with fixed final size $S=100$ nodes and $m=5$ links, but with different initial connection probabilities p ($p=0.1$ and $p=0.8$) and different sizes of ER initial cores. To reduce fluctuations in the degree distribution related to the small network size [1,2], we calculated the cumulative degree distribution $P(k)$.

Preliminary simulations suggest that, even for small values of m_0, the cumulative degree distribution $P(k)$ shows an exponential truncation for large values of k. To reduce the effects of this truncation on the estimate of the exponent of power-law behavior for the degree distribution, we only used the first five recorded values of k to compute the power law exponent of $P(k)$, as illustrated in Fig. 1.

For the less connected ER core ($p=0.1$), we used the following measures to characterize the effects of m_0 on the $P(k)$ (see Fig. 1): (1) the cutoff degree k_x, in which the observed $P(k)$ departs from the predicted power-law behavior, decreasing exponentially; (2) the maximum recorded degree k_{max}; (3) the proportion of nodes with $k_i > k_x$; and (4) the strength of truncation t, which describes the rate of decrease of $P(k)$ with k, following e^{-tk} for $k > k_x$. For the highly connected ER core ($p=0.8$) we only show how the degree distribution departs from the predicted power-law behavior.

III. RESULTS

Our results show that the scale-free nature of small BA networks is strongly affected by the size of ER initial core. For $p=0.1$, the strength of the exponential truncation of the power-law behavior of the degree distribution is enhanced by ER initial core size, as shown in Fig. 2. Figure 2(a) shows that increasing the ER initial core results in a linear decrease in k_x. Moreover, the exponent t increases linearly with the size of the ER initial core, for $k \geqslant k_x$, as shown in Fig. 2(b). The earlier truncation of the power-law behavior and the increase of t with m_0 imply that the homogeneity of the BA networks (i.e., similarity of the degree k between different nodes) increases with ER core size. In fact, the increase in ER core size generates a linear reduction in the maximum recorded degree, Fig. 2(c), and a logarithmic increase in the proportion of nodes in which $k_i > k_x$, Fig. 2(d).

The degree distribution in BA network with less-connected ER initial cores, including 15%–30% of all nodes of the network, behaves as a power law for $k_i < k_x$ and as an exponential for $k_i > k_x$. Thus, these networks are broad-scale networks. For larger initial cores ($m_0=30$), the degree distribution $P(k)$ departs from the expected by the power law even earlier ($k_x < 10$). In fact, these networks cannot be character-

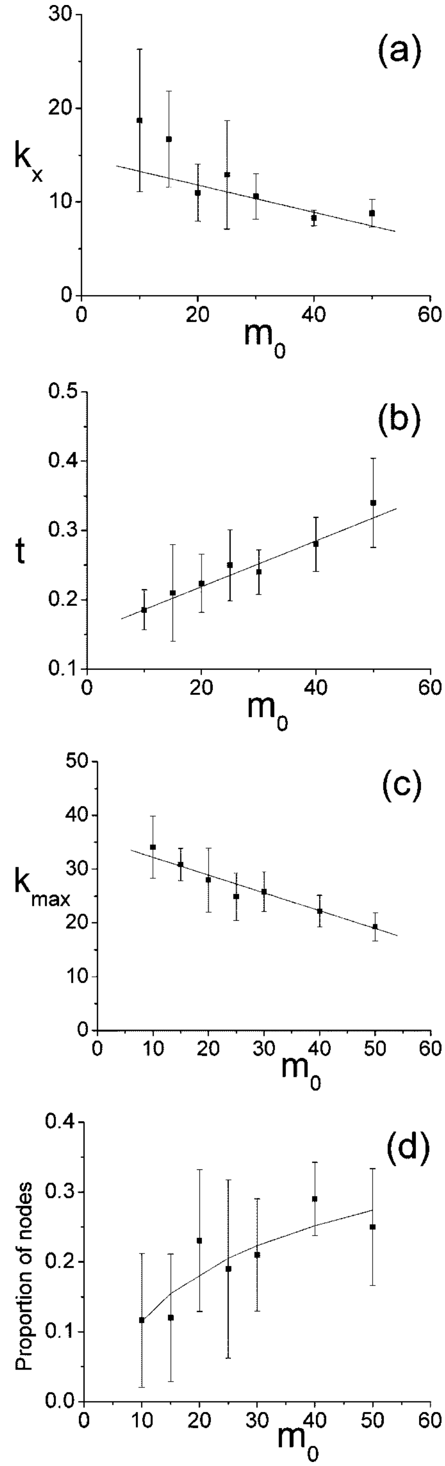

FIG. 2. Effects of the ER initial core size m_0 for $p=0.1$ on (a) the cutoff degree; (b) the strength of exponential truncation, t (see text for further details); (c) the maximum recorded degree, k_{max}; and (d) the proportion of nodes of the network with $k_i > k_x$. The network size has been fixed to 100 nodes.

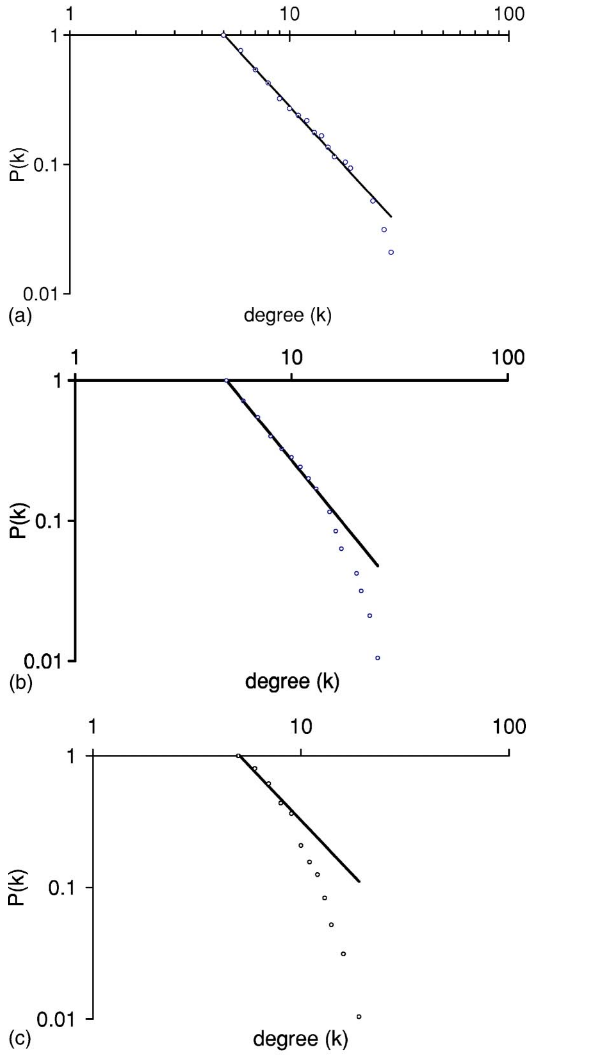

FIG. 3. Different classes of complex networks are generated by varying the size of the ER initial core m_0 for $p=0.1$. The network size has been fixed to 100 nodes. (a) scale-free networks ($m_0=15$), (b) broad-scale networks ($m_0=25$) and (c) single-scale networks ($m_0=50$).

ized as scale-free or broad-scale networks, being essentially exponential or single-scale networks [1]. Therefore, by simply changing the ER core size we are able to reproduce the main classes of complex networks [5], as shown in Fig. 3.

Figure 4 displays the degree distribution for $p=0.8$, showing that it markedly departs from the expected by power-law behavior. A gap in the range of k values appears and linearly increases with ER core size. These networks can therefore be divided into two subnetworks: (1) before the gap, a group of nodes that attached preferentially to nodes in the ER core, generating a power-law degree distribution, and (2) after the gap, a highly connected group of nodes, the ER core, in which the degree distribution is exponential.

IV. SUMMARY AND DISCUSSION

Complex networks, both biotic and abiotic, often show exponential truncation in scale-invariant topology [1,5,7].

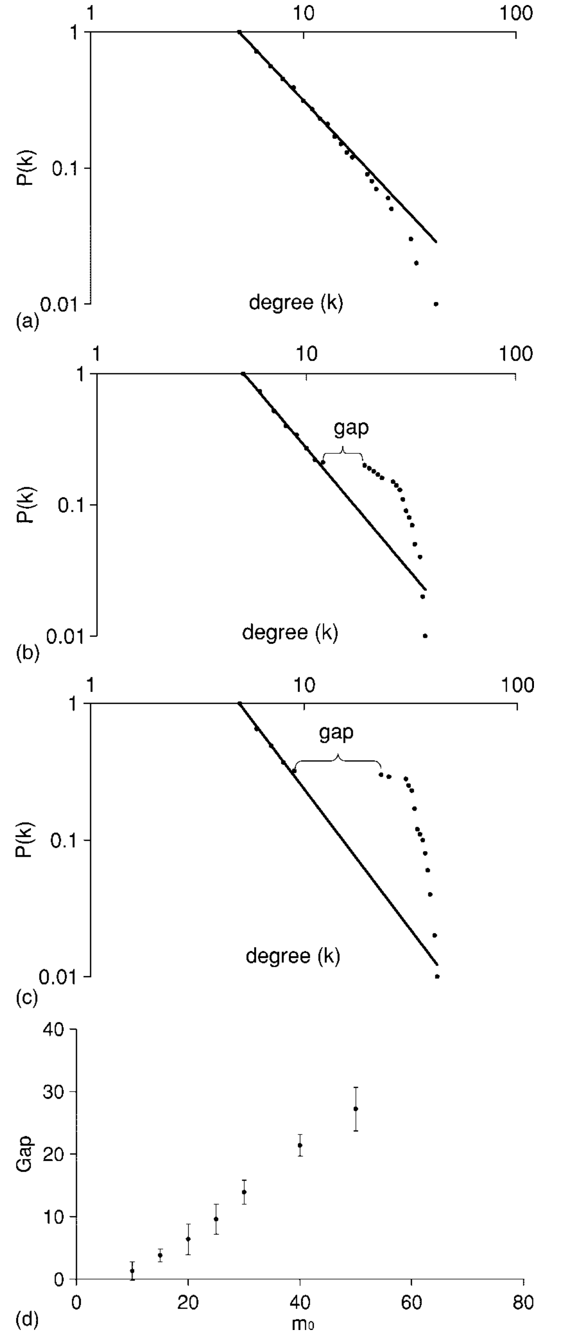

FIG. 4. Effects of the size of the ER initial core m_0 on the degree distribution for highly connected cores ($p=0.8$). The network size has been fixed to 100 nodes. (a) $m_0=10$, (b) $m_0=20$, and (c) $m_0=30$. Panel (d) shows that the gap indicated in parts (b) and (c) increases linearly with m_0 (see text for further details).

Previous works suggest that this truncation may be a result of constraints on addition of links. Some different classes of such constraints have been proposed, such as information filtering [5], aging, or connection costs [1]. In ecology, the truncation of power-law behavior was observed in food webs [19] and in coevolutionary bipartite networks of plant-animal interactions [7]. Jordano *et al.* [7] suggested that these truncations are generated by biological constraints that limit the possible links formed when species add up to the net, a phenomenon called forbidden interactions. Our results enlarge the catalog of mechanisms that could give rise to broad-scale networks, by adding the ER initial core as a new candidate. It is important to realize, however, that this new mechanism implies in a qualitatively new scenario. Assuming that there is a randomly connected initial core of nodes before preferential attachment starts to act in the network, there is no need to resort to additional constraints operating during network evolution; the truncation of power-law behavior is solely, or largely, a consequence of the system's initial condition. We expect that ER core hypothesis will be especially useful for systems in which it is possible to recognize the nodes that participated at the birth of the network from the nodes that appeared after a certain time period (e.g., species invading a food web or those occurring along a seasonal sequence).

The hypothesis of the ER initial core can be used to explain the buildup of all three main classes of networks [1]. As the size of the initial core is negatively related with the value of the cutoff in which the degree distribution departs from the predicted by the power-law behavior, it is possible to generate networks that are essentially scale free (small ER cores), broad-scale (intermediate ER cores) and essentially single-scale (large ER cores). The similarity of the pattern generated by two distinct, highly different mechanisms—namely, initial conditions and growth constraints—implies that alternative measures are necessary before arguing that constraints are limiting the network evolution in physical or biological systems best described as small networks. A simple way to distinguish between the two mechanisms is to measure the preferential attachment probability during network growth [18]. Mechanisms of constraints predict a reduction of preferential attachment during network growth while the ER core hypothesis predicts that, once preferential attachment starts, it is no longer affected by network growth. In several natural systems, however, the addition of new nodes (e.g., species in food webs) occurs over long time periods and it becomes impossible to measure the attachment probability. As a consequence, the development of new topological measures is central to allow an adequate distinction between the effects of constraints and those of the ER initial condition on real networks.

Finally we notice that the effects of the ER core on network topology will vary depending on how densely connected it is. While a less-connected ER core generates a size-dependent exponential truncation of the degree distribution, highly connected cores will affect differently the degree distribution: as the ER core size increases the network will be more clearly structured in two sets of nodes. In the first set of nodes, the ER core, the $\langle k \rangle$ is large and the degree distribution follows an exponential decay. In the second set, the nodes in which preferential attachment is operating, the degree distribution follows a power law. The distance in the degree between the two sets, measured by the gap in k values, increases linearly with ER core size. Recently, Melián and Bascompte [20] described food webs in which there is one or more cohesive, central subnets with the remaining nodes connected to them. The initial highly dense ER core is a possible mechanism to explain this pattern, by generating a central, densely connected core of nodes.

ACKNOWLEDGMENTS

Financial support was provided by the Brazilian agencies FAPESP and CNPq (MAMA, PRG, and SFR). P.J. and J.B. were funded by Grant Nos. BOS2000-1366-C02 and REN2003-00273 from the Spanish Ministerio de Ciencia y tecnologa (MCyT), the Spanish ministerio de educación y Ciencia (Grant No. REN2003-04774 to J.B.), and Grant No. RNM-305 (Junta de Andalucía, Spain), as well as a CNPq-CSIC bilateral agreement.

[1] L. A. N. Amaral, A. Scala, M. Barthelemy, and H. E. Stanley, Proc. Natl. Acad. Sci. U.S.A. **97**, 11149 (2000).
[2] S. H. Strogatz, Nature (London) **410**, 268 (2001).
[3] R. Albert and A. L. Barabasi, Rev. Mod. Phys. **74**, 47 (2002).
[4] S. N. Dorogovtsev and J. F. F. Mendes, Adv. Phys. **51**, 1079 (2002).
[5] S. Mossa, M. Barthelemy, H. E. Stanley, and L. A. N. Amaral, Phys. Rev. Lett. **88**, 138701 (2002).
[6] M. E. J. Newman, D. J. Watts, and S. H. Strogatz, Proc. Natl. Acad. Sci. U.S.A. **99**, 2566 (2002).
[7] P. Jordano, J. Bascompte, and J. M. Olesen, Ecol. Lett. **6**, 69 (2003).
[8] J. M. Montoya and R. V. Sole, Oikos **102**, 614 (2003).
[9] M. E. J. Newman, SIAM Rev. **45**, 167 (2003).
[10] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, Phys. Rev. E **63**, 062101 (2001).
[11] A. L. Barabasi and R. Albert, Science **286**, 509 (1999).
[12] E. Ravasz and A. L. Barabasi, Phys. Rev. E **67**, 026112 (2003).
[13] T. H. Keitt and H. E. Stanley, Nature (London) **393**, 257 (1998).
[14] C. J. Melian and J. Bascompte, Ecol. Lett. **5**, 705 (2002).
[15] R. Albert and A. L. Barabasi, Phys. Rev. Lett. **85**, 5234 (2000).
[16] http://vlado.fmf.uni-lj.si/pub/networks/pajek/
[17] Summary of the public debate on the *Growing Networks and Graphs in Statistical Physics, Finance, Biology and Social Systems*, [Eur. Phys. J. B **38**, 143 (2004)].
[18] H. Jeong, Z. Neda, and A. L. Barabasi, Europhys. Lett. **61**, 567 (2003).
[19] J. A. Dunne, R. J. Williams, and N. D. Martinez, Proc. Natl. Acad. Sci. U.S.A. **99**, 12 917 (2002).
[20] C. J. Melian and J. Bascompte, J. Ecol. **85**, 352 (2004).