Uncorrelated random networks

Z. Burda[1,2] and A. Krzywicki[3]
[1]*Fakultät für Physik, Universität Bielefeld, Postfach 100131, D-33501 Bielefeld, Germany*
[2]*Institute of Physics, Jagellonian University, ul. Reymonta 4, 30-059 Kraków, Poland*
[3]*Laboratoire de Physique Théorique, Bâtiment 210, Université Paris-Sud, 91405 Orsay, France*
(Received 2 July 2002; published 25 April 2003)

We define a statistical ensemble of nondegenerate graphs, i.e., graphs without multiple-connections and self-connections between nodes. The node degree distribution is arbitrary, but the nodes are assumed to be uncorrelated. This completes our earlier publication [Phys. Rev. **64**, 046118 (2001)] where trees and degenerate graphs were considered. An efficient algorithm generating nondegenerate graphs is constructed. The corresponding computer code is available on request. Finite-size effects in scale-free graphs, i.e., those where the tail of the degree distribution falls like $n^{-\beta}$, are carefully studied. We find that in the absence of dynamical internode correlations the degree distribution is cut at a degree value scaling like N^{γ}, with $\gamma=\min[1/2,1/(\beta-1)]$, where N is the total number of nodes. The consequence is that, independently of any specific model, the internode correlations seem to be a necessary ingredient of the physics of scale-free networks observed in nature.

DOI: 10.1103/PhysRevE.67.046118 PACS number(s): 05.10.−a, 05.40.−a, 87.18.Sn

I. INTRODUCTION

This paper is a direct continuation of Ref. [1]. The importance of defining statistical ensembles of random graphs in order to understand the geometry of wide classes of networks independently of any specific model was emphasized there. Concepts borrowed from field theory were used to define the ensemble of uncorrelated graphs, and an algorithm generating such graphs was proposed. The general philosophy of our approach was illustrated by focusing on a graph subensemble, namely, on an ensemble of connected trees with a scale-free degree distribution, where a number of hopefully interesting analytic results could be presented. But it should have been obvious that the adopted framework is of much broader applicability. Actually, we have explicitly stated that our algorithm generates efficiently not only trees, but also so-called pseudographs, called degenerate graphs in Ref. [1]. However, we have also indicated that we encountered problems dealing with simple (i.e., nondegenerate) graphs. Hence the discussion of these nondegenerate graphs was postponed. We are now returning to the problem of defining the statistical ensembles of networks, which in the meantime has attracted the attention of other researchers [2–5] (at this point is it fair to mention also Ref. [6], an early paper on uncorrelated graphs).

Although some overlap with Ref. [1] is unavoidable, if this paper is to be self-contained, we would like to reduce the overlap to a minimum. The reader is invited to consult Ref. [1] when he finds the discussion of this paper too sketchy.

We shall not dwell much in introducing the subject. Let us recall that a graph is just a collection of nodes (vertices) connected by links (edges). It is a mathematical idealization representing the various networks one encounters in nature, in social life, in engineering, etc. Quite often the pattern of connections between nodes looks fairly random. The concept of a random graph emerges quite naturally.

Random graphs are interesting in themselves. There exists a classical theory of random graphs, a beautiful piece of pure mathematics [7]. It turns out, however, that most large graphs one encounters in applications are not covered by this theory. The access—relatively recent—to the corresponding data triggered a rather intense activity (see Refs. [8,9] and references therein).

Networks are also interesting considered from a broader perspective: it is useful to represent the architecture, so to say the skeleton of many complex systems by an appropriate network. Hence, graphs are in a sense a gateway to the theory of complex systems, an exciting and promising new direction of research.

The plan of this paper is as follows: In Sec. II we recall the definition of the ensemble of uncorrelated random graphs and we discuss the algorithm generating the graphs. In Sec. III we present the results of a sample of computer simulations, aimed to help understanding finite-size effects. The latter play a very important role as soon as the degree distribution has a fat tail. We explain the behavior of the data using an analytic argument. We conclude in Sec. IV. For definiteness, we consider undirected graphs only, as in Ref. [1].

II. DEFINING THE ENSEMBLE

Let us recall the construction proposed in Ref. [1]. The partition function for the ensemble of random graphs is written as the formal integral defining a minifield theory

$$Z=\int_{-\infty}^{+\infty} \mathrm{d}\,\phi \exp\left\{\frac{1}{\kappa}\left[-\phi^2/2\lambda+g\sum_{n=1}^{\infty} p_n\phi^n\right]\right\}. \qquad (1)$$

It will be seen that the non-negative constants p_n correspond to the degree distribution, while the auxiliary constants g, λ, and κ control the dependence of Z on the number of nodes, links, and loops, respectively. The integral does not exist, but Z can be treated formally as a generating function in the Gaussian perturbation theory. The main idea is to expand the exponential under the integral in Eq. (1) in powers of g,

$$Z=\int \mathrm{d}\,\phi \exp\left\{\frac{1}{\kappa}[-\phi^2/2\lambda]\right\}\left[1+\frac{g}{\kappa}\sum_n p_n\phi^n + \frac{1}{2!}\left(\frac{g}{\kappa}\right)^2\sum_{n,m} p_n p_m \phi^{n+m}+\cdots\right], \tag{2}$$

to get a series in g with well defined coefficients, viz. integrals with the Gaussian measure of integer powers of ϕ. Each such integral is equal to a sum of contributions, which can be represented graphically by the so-called Feynman diagrams [10]. We have explained in Ref. [1] how such a diagram emerges, using a particular example. We do not have enough space to develop the point in more details. For those readers who are not conversant with field theory techniques we list the rules for constructing and calculating the Feynman diagrams corresponding to the term of order $O(g^N)$ in Eq. (2).

Each diagram has N labeled nodes. One should draw all topologically distinct diagrams, distributing degrees among nodes in all possible manners and connecting nodes pairwise. Self-connections and multiple connections between nodes are allowed. Notice the similarity with the Molloy-Reed construction [11]. With each diagram is associated a number, called the Feynman amplitude, determined by the following rules: each node of degree n contributes a factor $(g/\kappa)p_n n!$ and each link contributes a factor $\kappa\lambda$. There is a symmetry factor 1/2 associated with every line connecting a node to itself and a symmetry factor $1/m!$ associated with every m-tuple connection between nodes [12]. There is also a factor $Z_0/N!$, the factorial being a remnant of the expansion of an exponential and Z_0 being the value of the Gaussian integral.

Finally, the series representation of the partition function reads

$$Z=Z_0\sum_{L,N}\frac{g^N}{N!}\kappa^{L-N}\lambda^L\sum_D \frac{1}{S(D)}\prod_{j=1}^N p_{n_j} n_j!, \tag{3}$$

where one sums over labeled diagrams D having a fixed number of nodes and links, respectively, N and L. $S(D)$ is the product of factors 2 and factorials associated with self-connections and multiple connections and n_j is the degree of the jth node. One can show that the analogous series for $\log Z$ receives contributions of connected diagrams only. In this case the expansion in powers of κ is a loop expansion: the leading term corresponds to $L=N+1$ and comes from tree diagrams, the next term comes from one-loop diagrams, etc.

Our idea is to identify the Feynman diagrams of the toy model defined by Eq. (1) with the graphs of a statistical ensemble. Indeed, Feynman diagrams are identical to graphs familiar to network community people, except that there are definite rules to calculate the corresponding weights. The minifield formulation enables one to summarize compactly the content of the model and has also the advantage of being a good starting point for analytical calculations, like those of Ref. [1].

In the following, we always work with graph ensembles where N and L are fixed. Hence, up to an irrelevant factor, the weight w of a labeled graph that is nondegenerate, i.e., such that nodes are neither multiply connected nor connected to themselves, is just

$$w\sim\prod_{j=1}^N p_{n_j} n_j!. \tag{4}$$

In the presence of degeneracies one has to multiply the right-hand side (rhs) of Eq. (4) by the factor S^{-1} appearing in Eq. (3).

Equation (4) gives the weight of a microstate. Notice the factorized form and therefore the absence of nontrivial, dynamical correlations. Notice also that with the choice $p_n \propto (\mathrm{const})^n/n!$ all nondegenerate graphs with the same N and L are equiprobable, because

$$\sum_j n_j = 2L. \tag{5}$$

Thus, with a Poissonian p_n one recovers the classical graph ensemble of Erdös and Rényi. The ensemble under discussion is the most conservative generalization of the classical ensemble to the case of an arbitrary degree distribution.

At this point the statistical ensemble is basically defined. However, in this paper, we wish to focus on nondegenerate graphs, which are the primary objects in graph theory. They correspond to a subensemble of Feynman diagrams. In the conventional applications of field theory no specific recipe is formulated to single these diagrams out. Such a recipe is, however, needed here. Otherwise our definition of the ensemble would be too vague to be useful in applications.

Before going farther let us outline the strategy we shall follow: as stated above, our goal is now to complete the definition of the ensemble by the construction of an algorithm generating nondegenerate graphs. But we do not achieve this goal directly. First we construct, following Ref. [1], an algorithm generating graphs that are degenerate. Then we show that the ensemble of these degenerate graphs is isomorphic, as far as the degree distribution is concerned, to the known model of balls-in-boxes [14]. Using this mapping of one model on another we conclude that asymptotically the degree distribution P_n in the ensemble of degenerate graphs is just $p_n: P_n \to p_n$ for $N\to\infty$. Since we suspect that in this limit the degree distributions are the same for degenerate and nondegenerate graphs, we impose the appropriate constraint on the algorithm and perform a sample of computer experiments, to be described in Sec. III. The results might seem surprising at first sight, but a clear picture eventually emerges when we estimate analytically, in the ensemble of degenerate graphs, the likelihood that a node is neither self- nor multiply-connected.

In a growing network model the construction of graphs is recursive and mimics a real physical process. In a static model like ours one does not refer to any physical process. An ensemble is defined and the relative frequency of occurence of distinct graphs is fixed: If graphs A and B have weights $P(A)$ and $P(B)$, respectively, then they should be

generated with a relative frequency equal to $P(A)/P(B)$. Naively, one could imagine generating graphs uniformly in the space of graphs, accepting graph A, say, with probability $P(A)$. However, such a uniform sampling is in practice very difficult to insure. Furthermore, in an ensemble of very many graphs the acceptance rate of the naive algorithm would be very small, since the normalized weight of any given graph is roughly speaking of the order of the inverse of the number of graphs. A clever idea is to introduce an appropriate random walk (Markov process) in the space of graphs, $\ldots \to A_k \to A_{k+1} \to A_{k+2} \to \ldots$, which performs an importance sampling. The random walk is driven by the Markovian transition matrix $P(A \to B)$. One can easily show that if the transition matrix fulfills the detailed balance condition

$$P(A)P(A \to B) = P(B)P(B \to A) \tag{6}$$

the frequency of the configuration A_k in the Markov process is proportional to $P(A_k)$, provided one has moved away from the initial configuration. There are many $P(A \to B)$ fulfilling the detailed balance condition for a given probability measure $\{P(A), \forall A\}$. One is free to choose any one. The simplest and popular choice

$$P(A \to B) = \min\{1, P(B)/P(A)\} \tag{7}$$

is usually referred to as the Metropolis algorithm [13]. The general idea of the method is problem independent. However, the choice of the proposed new configuration B, given the current one A, is made by taking into account the particularities of the problem at hand. Usually one proposes to change only slightly a small number of parameters in the current configuration. This insures a reasonable acceptance rate and minimizes the risk of performing time consuming calculation for nothing.

The transition $A \to B$ logically involves two steps: one proposes B among all candidates and one accepts the proposal with a certain probability. One can write $P(A \to B)$ as a product of the probability P_c of choosing a particular candidate and of the probability P_a of accepting it: $P(A \to B) = P_c(A \to B)P_a(A \to B)$.

Our algorithm [1] works as follows. In the current configuration a random oriented link $\vec{i}j$, the candidate for rewiring, is chosen. This is done with the probability $1/2L$. Then we select a vertex k, with the probability $1/N$. The proposed move consists of rewiring $\vec{i}j$ into $\vec{i}k$. Thus, $P(A \to B) = 1/2LNP_a(A \to B)$, and similarly for $A \leftrightarrow B$. Inserting this into the detailed balance condition and dropping the factors $1/2LN$, identical on both sides of the equation, we obtain

$$P(A)P_a(A \to B) = P(B)P_a(B \to A), \tag{8}$$

which has the Metropolis solution for $P_a(A \to B)$. Now, we use the fact that, according to Eq. (4), $P(A)$ is a product of the node weights $p_n n!$. Furthermore, we observe that the rewiring changes $n_k \to n_k + 1$ and $n_j \to n_j - 1$ only, leaving the degrees of other nodes intact, to get

$$P_a(A \to B) = \min\left\{1, \frac{(n_k+1)!\,p_{n_k+1}(n_j-1)!\,p_{n_j-1}}{n_k!\,p_{n_k} n_j!\,p_{n_j-1}}\right\} = \min\{1, (n_k+1)R(n_k+1)/n_j R(n_j)\}, \tag{9}$$

where $R(n) = p_n/p_{n-1}$. When $n_j = 1$, the attempt is rejected, so that nodes with zero degree are never created. Notice, that we directly sample links to be rewired. The graphs produced by this algorithm are in general degenerate and multiply connected. It turns out that the detailed balance condition and the way of sampling links insure that the symmetry factors in the weights of degenerate graphs come out correctly.

The presence of the factor $(n_k+1)/n_j$, on the rhs of Eq. (9) means that the rewired nodes are sampled independently of their degree [15]. Furthermore, the rewiring depends on the node degrees only and is insensitive to the rest of the underlying graph structure. Hence, as far as the distribution of node degrees is concerned, the model is isomorphic to the well known balls-in-boxes model [14], defined by the partition function

$$z \propto \sum_{\{n_j\}} p_{n_1} \cdots p_{n_N} \delta\left(M - \sum_{j=1}^{N} n_j\right), \tag{10}$$

and describing M balls distributed with probability p_n among N boxes (in our case $M = 2L$). The constraint represented by the Kronecker delta on the rhs of Eq. (10) is satisfied "for free" when $N \to \infty$ by virtue of Khintchin's law of large numbers, provided $\langle n \rangle = \Sigma_n n p_n / \Sigma_n p_n = M/N$. The finiteness of $\langle n \rangle$ is always tacitly assumed in this paper. Hence, when the last condition is met the occupation number distribution of a single box $P_n \to p_n$ for $N \to \infty$.

Consequently, in the statistical ensemble, including degenerate graphs, the degree distribution becomes p_n asymptotically when the number of links is set to

$$L = \frac{1}{2} N \langle n \rangle. \tag{11}$$

When this condition is not met, the asymptotic degree distribution differs from p_n, which is, in a sense, renormalized. In particular, when L is smaller, this distribution is p_n times an exponentially falling factor. When L is larger the situation depends on the shape of p_n. When the latter is scale free, $p_n \propto n^{-\beta}$ for large n, the distribution is $\propto p_n$, except that an extra singular node with degree of order $O(N)$ shows up. These phenomena were discussed at length in the context of the balls-in-boxes model [14], and also in our preceding work [1].

So far, only an algorithm generating degenerate graphs has been constructed. It is trivial to convert it into an algorithm producing nondegenerate graphs. It suffices for that to add before the Metropolis test a few lines of code checking that the nodes i and k are neither identical nor linked. However, this check introduces a bias and it is not obvious what will be the degree distribution at finite N.

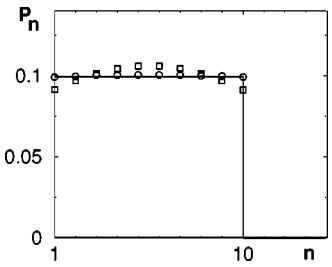

FIG. 1. The approach towards the limiting rectangular shape of the calculated connectivity distribution P_n: $N=100$ (squares) and 1000 (circles). The figure illustrates the claim that finite size corrections fade away rapidly when the input connectivity distribution has a finite support.

A priori, the Metropolis test should insure that the number of nodes of degree n is close to Np_n, provided the last number is large enough. And for fixed n it can be made arbitrarily large with a proper choice of N. Hence, a possible deviation of the degree distribution from p_n should be a finite size effect disappearing in the limit $N\to\infty$ when the couplings p_n are defined on a finite support. However, one has a problem when p_n has a fat tail.

Let the degree distribution fall like $n^{-\beta}$, $\beta>2$. For finite N it cannot fall like that indefinitely, there is a natural cutoff scaling as

$$n_{\max}\propto N^{1/(\beta-1)}. \tag{12}$$

The argument is well known: the expected number of nodes with $n>n_{\max}$ is less than unity. The presence of this cutoff was used by Dorogovtsev *et al.* [16] to explain why the observed scale-free networks are always characterized by a relatively small β.

Hence, coming back to the algorithm, there is always a range of n where fluctuations in the number of nodes are very large. Increasing N does not help. Now, if certain fluctuations are systematically favored by the constraint excluding degeneracies, then the resulting degree distribution can strongly deviate from the input weights p_n. We dedicate a separate section to the discussion of this problem.

III. FINITE-SIZE GRAPHS: DEGREE DISTRIBUTION

Let us first consider a case where the support of p_n is finite, in order to check that in this case the problem is indeed under control. We perform a numerical experiment, setting for definiteness $p_n=1/10$ for $n\leqslant 10$ and $p_n=0$ otherwise, while $L=2.75N$ as dictated by Eq. (11). The result is shown in Fig. 1: as expected, the convergence of the degree distribution towards the input one is very fast.

Our next experiment is with scale-free graphs. As in Ref. [1] we set [17]

$$p_n=(\beta-1)\frac{\Gamma(2\beta-3)\Gamma(n+\beta-3)}{\Gamma(\beta-2)\Gamma(n+2\beta-3)}\propto n^{-\beta}, \tag{13}$$

but the generated graphs are now nondegenerate with loops, instead of trees (the graphs are also, in general, not connected). Since $\langle n\rangle=2$, we also set $L=N$. We have chosen this example to illustrate a behavior which, as we shall argue in a moment, is generic. Since the same choice was made in Ref. [1], the reader can directly compare the results obtained in the two ensembles studied.

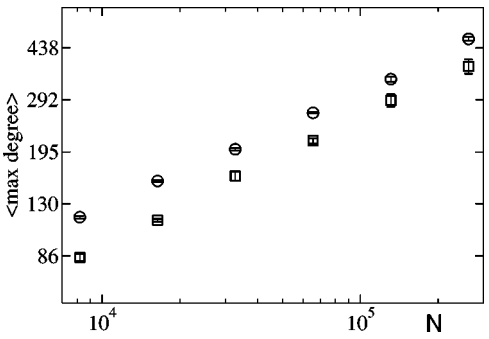

FIG. 2. The evolution with N of the average maximum degree of a graph for $\beta=2.1$ (circles) and 3.0 (squares) and $N=2^{13}$–2^{18}. This log-log plot illustrates the observation that for scale-free graphs with the exponent β approaching 2, the maximum node degree evolves with N at a rate close to that observed for $\beta=3$, in variance with Eq. (12). An analytic argument explaining the observed behavior is given in the text.

The maximum node degree $n_{\max}$ has been measured in long runs. The results, for $\beta=2.1$ and 3.0, are shown in Fig. 2. The important point is that the rate of increase of $n_{\max}$ with N is almost the same for the two values of β, contrary to what one can read from the rhs of Eq. (12). The exponent is slightly below the value 1/2, expected asymptotically (see below). The autocorrelation time increases roughly at the same rate, from about 860 (1810) to 5460 (10300) sweeps for $\beta=2.1$ (3.0); in a sweep one attempts to rewire all the L links of the graph. The fraction of nodes and links belonging to the giant component decreases very slowly from about 0.57 and 0.78 (0.68 and 0.81) at $N=2^{14}$ to 0.55 and 0.77 (0.68 and 0.81) at $N=2^{18}$.

The shape and the evolution with N of the degree distribution is shown in Figs. 3 and 4. It is manifest that the distribution does approach the expected limit. However, the approach is very slow and nonuniform, especially for $\beta=2.1$. The results of the computer experiment can be understood using the following heuristic arguments.

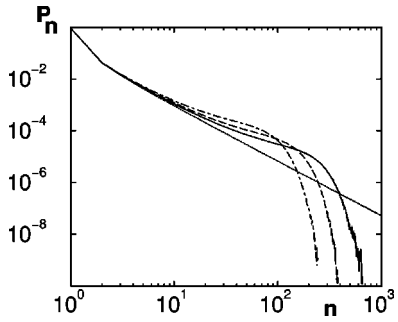

FIG. 3. The degree distribution for $\beta=2.1$ and $N=2^{14}$ (dot-dashed), 2^{16} (dashed), and 2^{18} (solid). The dotted line is the shape at infinite N. The cutoff scales like for $\beta=3$ (see the next figure). Since the distribution is normalized to unity, this results in a deviation of its shape from the asymptotic one.

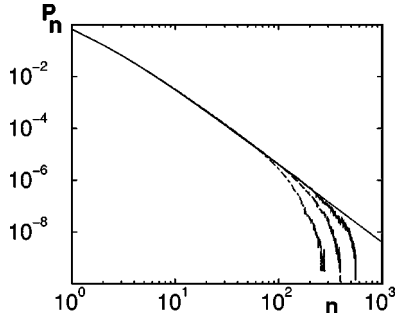

FIG. 4. The degree distribution for $\beta=3.0$ and $N=2^{14}$ (dot-dashed), 2^{16} (dashed), and 2^{18} (solid). The dotted line is the shape at infinite N.

Let $n \gg 1$ be the degree of a node in the tail, so that there is no more than one node of that degree. We consider the ensemble of degenerate graphs and we estimate the fraction of graphs where the node in question has no self-connections or multiple connection. The counting is easily performed by considering the symmetric adjacency matrix $C_{ij}=\#\text{links } \vec{ij}$ (every undirected link contributes to two elements of C). Obviously, C_{ij} is 0 or 1 for nondegenerate graphs, but can take any integer value for degenerate ones. Let m be the label of our hub. We count the adjacency matrices satisfying $\Sigma_j C_{jm}=n$.

We take the limit $N,L\to\infty$, L/N fixed and $n/N\to 0$ [notice that it is not assumed that n is kept fixed when N increases, it is only less than the cutoff given by Eq. (12)]. In this limit, the number of degenerate graphs we are interested in is proportional to the number of ways to place n elements in N cells, possibly putting several elements in the same cell. The standard fomula of combinatorics does not apply because one has to take care of the contribution of the symmetry factors to graph weights. The number of graphs corrected by the appropriate weights is

$$W_{\text{deg}}(n)=\sum_{\{k_j\}} \frac{1}{2^{k_m}}\prod_j \frac{1}{k_j!}\,\delta\left(\sum_{i=1}^{N} k_i - n\right) \tag{14}$$

$$=\frac{1}{2\pi}\int d\alpha\, e^{i\alpha} e^{-in\alpha+(N-1/2)e^{i\alpha}}, \tag{15}$$

where $k_j \equiv C_{jm}$. There is a self-connection when $k_m>0$ and a multiple connection when, for $j\neq m$, $k_j>1$. The symmetry factors are those mentioned at the place where we summarize the Feynman rules. We have replaced the δ function by its Fourier transform and performed the independent summations over k_j. Saddle-point integration yields

$$W_{\text{deg}}(n)\propto \frac{N^n}{n!}. \tag{16}$$

A similar counting of graphs where our node is neither self- nor multiply-connected yields

$$W_{\text{non-deg}}(n)=\binom{N-1}{n}\propto \frac{N^n}{n!}\exp\left(-\frac{n^2}{2N}\right). \tag{17}$$

The ratio is [18]

$$W_{\text{non-deg}}/W_{\text{deg}}\propto \exp\left(-\frac{n^2}{2N}\right). \tag{18}$$

Hence, the entropy of the nondegenerate graphs is dramatically reduced compared to that of graphs with degeneracies and this reduction depends strongly on n. This is the origin of the bias mentioned in Sec. II.

Suppose we have generated a very large sample of scale-free graphs without any care for degeneracies. For each such graph the histogram of degrees contains a tail of sparsely located columns of unit height. However, we know, because of the mapping on the balls-in-boxes model, that the histogram for the whole sample has the shape of p_n for $n<n_{max}$. Now, what happens when we exclude degeneracies, specifically when we check whether a node in the tail of the histogram does have the forbidden connections? Equation (18) tells us that the rejection rate is nonuniform in n and that nearly all candidate graphs will be rejected when n is large compared to $N^{1/2}$.

Thus, to the extent one can neglect the fluctuations of graph weights, the degree distribution in nondegenerate graphs is expected to be cut by an additional factor, roughly behaving like $\exp(-n^2/2N)$. In the absence of dynamical correlations and for $\beta\leqslant 3$ the cutoff in nondegenerate graphs is expected to scale like $N^{1/2}$ and not like the rhs of Eq. (12), i.e., independently of β. Imposing nondegeneracy one generates, at finite N, parasite correlations whose manifestation is the violation of Eq. (12). Notice that this entropy argument does not apply to trees, which within the full graph ensemble have a nearly vanishing entropy and which, as shown in [1], have their own mapping on balls-in-boxes.

The conclusion of this section is that our algorithm is efficient in generating all classes of graphs for any given degree distribution. In the specific, but most interesting case, of the maximally random nondegenerate scale-free graphs whose degree distribution falls like $n^{-\beta}$, with $\beta<3$, the tail of the degree distribution is cut at $n_{\max}\sim N^{1/2}$. This effect prevents the finite networks from fully developing the large part of the tail of their *a priori* expected degree distribution. This is a feature of the model, a result of the absence of dynamical correlations, and not a failure of the algorithm [19].

It is known that nodes are correlated in some growing network models with a scale-free behavior [17]. Here we find, independently of any specific model, that dynamical correlations seem to be needed for some networks with $\beta\approx 2$ to be observed in the real world. We also observe that the thermodynamic limit $N\to\infty$ of the maximum entropy graph model can be rather nonuniform and therefore somewhat tricky.

IV. SUMMARY AND CONCLUSION

There is much activity in producing new growing network models. Such models are invaluable for illustrating some ba-

sic dynamical mechanisms, like the preferential attachment rule [21]. However, in order to understand the generic geometries of wide classes of networks it is perhaps worthwhile to adopt a complementary approach, consistently defining the corresponding statistical ensembles and working with these static ensembles using the standard tools of equilibrium statistical mechanics.

With this motivation in Ref. [1] we have defined a statistical ensemble for an arbitrary node degree distribution. This ensemble is the most random possible, but clearly not the most general: we assumed that node degrees are independent, to the extent that this is possible when the number of nodes and links is fixed. In Ref. [1] we have rapidly focused on trees, indicating, however, that the approach extends to more complicated graphs. The discussion of the ensemble of nondegenerate graphs, those without multiple-connections and self-connections, has been postponed. The present paper completes Ref. [1].

In particular, we have constructed here an efficient algorithm, easily implemented on a computer, which enables one to generate nondegenerate random graphs. We are ready to share our computer code with interested people, on request.

Another simple algorithm has been proposed in Refs. [6,11] and used subsequently: one first generates from a given distribution p_n a set of node degrees $\{n_j\}$, and uses these numbers to construct auxiliary graphs, each with one node and n_j half-links. In the second step one connects the half-links at random. The resulting graph is usually degenerate [22]. Imposing the absence of degeneracies is in this approach very tedious, especially that a nondegenerate graph may just not exist for a given set $\{n_j\}$.

We have studied in detail the behavior of the degree distribution of finite-size graphs. In the absence of dynamical inter-node correlations, for generic scale-free nondegenerate graphs (but not trees) this distribution is cut at

$$n_{\max} \propto N^{\gamma}, \quad \gamma = \min[1/2, 1/(\beta - 1)] \tag{19}$$

at asymptotically large N. It appears that a fat tail with β rather close to 2 observed in some data could hardly show up if dynamical internode correlations were absent. And indeed, nontrivial correlations are present in models of growing networks using the preferential attachment recipe [17].

It is certainly very important to develop a theory of correlated networks. Other authors [3,4] have very recently made interesting explicit proposals in that direction. Our approach can also be rather easily generalized to include dynamical correlations. This does not mean that the *physics* of correlated networks is transparent to us at the present time. We shall hopefully return to the problem of correlations in another publication.

ACKNOWLEDGMENTS

We wish to acknowledge our friend and collaborator the late Joao D. Correia. This work was partially supported by the EC IHP Grant No. HPRN-CT-1999-000161 and by Project 2 P03B 096 22 of the Polish Research Foundation (KBN) for 2002-2004. Z.B. thanks the Alexander von Humboldt Foundation. Laboratoire de Physique Théorique is Unité Mixte du CNRS UMR 8627.

[1] Z. Burda, J. D. Correia, and A. Krzywicki, Phys. Rev. E **64**, 046118 (2001).

[2] S. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, preprint cond-mat/0204111.

[3] S. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, preprint cond-mat/0206131.

[4] J. Berg and M. Lässig, preprint cond-mat/0205589.

[5] M. Bauer and D. Bernard, preprint cond-mat/0206150.

[6] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. E **64**, 026118 (2001).

[7] B. Bollobás, *Random Graphs* (Academic Press, London, 1985).

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

[9] S. N. Dorogovtsev and J. F. F. Mendes, Adv. Phys. **51**, 1079 (2002).

[10] See any field theory textbook. A short introduction can be found in D. Bessis, C. Itzykson, and J. B. Zuber, Adv. Appl. Math. **1**, 109 (1980).

[11] M. Molloy and B. Reed, Random Struct. Algorithms **6**, 161 (1995); Combinatorics, Probab. Comput. **7**, 295 (1998).

[12] The factor 1/2 reflects the symmetry between the two orientations of a link forming a tadpole and the factor $1/m!$ reflects the symmetry of link permutations in an m-tuple connection between nodes.

[13] N. Metropolis *et al.*, J. Chem. Phys. **21**, 1087 (1953); K. Binder, *Monte Carlo Methods in Statistical Physics* (Springer, Berlin, 1986); a pedagogical introduction can be found in M. Creutz, L. Jacobs, and C. Rebbi, Phys. Rep. **95**, 201 (1983).

[14] P. Bialas, Z. Burda, and D. Johnston, Nucl. Phys. B **493**, 505 (1997); **542**, 413 (1999).

[15] Sampling links priviledges nodes with large degree. The factor $(n_k+1)/n_j$ eliminates the bias. One can drop this factor in the Metropolis test by selecting the link to be rewired as follows: select at random, in the register of nodes, the node j. Then, pick at random one of the links converging to j. With this method the mapping on the balls-in-boxes model is even more obvious, since one samples directly the *nodes* to be updated. In the model of Ref. [14] the node orders are the only degrees of freedom.

[16] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, Phys. Rev. E **63**, 062101 (2001).

[17] P. L. Krapivsky and S. Redner, Phys. Rev. E **63**, 066123 (2001).

[18] Strictly speaking, the exponent on the right-hand side is $-n^2/2N[1+O(n/N)]$. The ratio falls faster than a Gaussian at large n. The estimate given in Eq. (18) is crude. One should take into account the fact that approximately $p_1 N$ nodes have degree 1, approximately $p_2 N$ nodes have degree 2, etc. The

corresponding modification of the right-hand side of Eq. (15) is straightforward, but the formulae are somewhat cumbersome. We skip this calculation, mentioning only that once it is performed, the constant 1/2 in the exponent of Eq. (18) is replaced by another constant depending on the shape of p_n. This does not affect the scaling properties of the cut-off.

[19] Our first guess was that, at finite N, one should use a modified set of input couplings [20]. We realized that it is a wrong strategy, when we understood the entropic origin of the distorsion of the degree distribution. Another comment: It is interesting to compute the number L_r of ''redundant'' links, i.e., those producing self-connections and multiple connections, in the ensemble of scale-free graphs where the degeneracies have not been excluded. When the connectivity distribution has a finite support this number, in a generic graph, is expected to be small, although most graphs are degenerate. This is also the behavior we observe for $\beta=3$. We find $L_r=5.3(3)$, 10.6(5), 15.5(8), and 16.9(8) at $N=100$, 1000, 10 000, and 100 000, respectively. As one moves β towards 2 the situation changes dramatically. For $\beta=2.1$ we find, respectively, for the same values of N: $L_r=20.4(2)$, 166(6), 1273(15), and 9374(71). This behavior can be inferred from the analytic argument of Sec. III. We do not count redundant links there, but we estimate the weight of degenerate graphs, which involves the same information.

[20] A. Krzywicki, preprint cond-mat/0110574.

[21] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).

[22] This construction is described in detail in Ref. [11], but not as a method of constructing nondegenerate graphs, the actual object of the study of the authors. Molloy and Reed find it useful to map the ensemble of graphs they are interested in on the less constrained ensemble of ''random constructions,'' i.e., degenerate graphs (this approach was earlier used by other authors; see Ref. [7]). We are indebted to M. Bauer and D. Bernard for calling our attention to this point.