General formalism for inhomogeneous random graphs

Bo Söderberg*
Complex Systems Division, Department of Theoretical Physics, Lund University, Sölvegatan 14A, S-22362 Lund, Sweden
(Received 19 July 2002; published 18 December 2002)

We present and investigate an extension of the classical random graph to a general class of inhomogeneous random graph models, where vertices come in different types, and the probability of realizing an edge depends on the types of its terminal vertices. This approach provides a general framework for the analysis of a large class of models. The generic phase structure is derived using generating function techniques, and relations to other classes of models are pointed out.

DOI: 10.1103/PhysRevE.66.066121 PACS number(s): 02.50.−r, 64.60.−i, 89.75.Fb

I. INTRODUCTION

The concept of random graphs (RG) has recently become the target of an increasing interest, as a tool for modeling various kinds of networks, arising e.g. in physics, biology, and biophysics, as well as in social and information-technological structures.

The classical RG model [1–3] describes a homogeneous, sparse random graph of order N, where each edge is randomly and independently realized with a fixed probability $p=c/N$. For large orders, there is a critical value of $c=1$, above which almost every graph contains a single giant connected component being of order $O(N)$, with the remaining components being small compared to N. This model yields an asymptotic degree distribution that is Poissonian with the average degree given by c.

Many real-life networks, such as the internet, have been shown to possess other types of degree distribution, sometimes displaying a power law behavior over many orders of magnitude, ruling out the classical RG as the relevant model. A number of alternative RG models have been suggested in an attempt to yield random graphs with more general types of degree distribution, such as the desired power behavior. Some of these models describe dynamical random graphs, where the graphs arise as the result of a stochastic growth process, such as randomly grown networks [4,5], or scale-free networks based on preferential attachment [6]. Others focus on describing ensembles of random graphs with certain given properties, without bothering about how they came about; a particularly interesting approach of this type, possessing a high degree of generality, is based on considering random graphs of fixed order with a given arbitrary degree distribution [7–10].

In this paper, we will investigate a general class of sparse *inhomogeneous* RG models, by means of a straightforward generalization of the classical model to a situation where vertices may come in different *types*, such that the probability for an edge depends on the types of its pair of terminal vertices. While this class of models inherits certain features from the homogeneous model—such as the existence of a critical hypersurface in parameter space, beyond which asymptotically almost every graph has a giant component—it is capable of producing a wide class of asymptotic degree distributions, among these distributions with power law behavior.

This general class of models is shown to contain a number of existing models as special cases, and can be used as a general framework for the analysis of various RG models.

The structure of this paper is as follows. In Sec. II, some of the more salient features of the classical model are briefly reviewed, while our generalization is presented and analyzed in Sec. III. In Sec. IV, a number of special cases are discussed, while Sec. V contains our conclusions.

II. THE CLASSICAL MODEL

Here we will briefly review some of the more prominent properties of the classical random graph in the large N limit, to pave the ground for the subsequent analysis of its generalization.

Definition 1. Let $\mathcal{G}(N,c)$, with c being a real positive number, denote the ensemble of graphs of order $N>c$, where each edge is independently realized with probability $p=c/N$. This ensemble has a critical value of $c=1$, above which almost all graphs for large N have a single large connected component — the giant component — with a finite fraction of the vertices, while the remaining components are small.

A. Exposing connected components

The standard method to reveal the size distribution of the orders of components is to expose these components as follows. Start with a single (random) vertex, reveal its neighbors by following edges, then their neighbors, etc. Let n_k be the number of vertices exposed for the first time in step k of this process. The distribution of n_k, given the previous numbers $n_0=1,n_1,\ldots,n_{k-1}$, becomes

$$P(n_k)=\binom{N-\sum_{l=0}^{k-1} n_l}{n_k}(1-q^{n_{k-1}})^{n_k}(q^{n_{k-1}})^{N-\sum_{l=0}^{k} n_l}, \tag{1}$$

where $q=1-c/N$.

In the large N limit with a fixed c, $P(n_k)$ tends to $e^{-n_{k-1}c}(n_{k-1}c)^{n_k}/n_k!$, and the process reduces to a Poisso-

*Electronic address: Bo.Soderberg@thep.lu.se

nian branching tree model $\mathcal{B}(c)$, with each vertex independently branching to a number of new vertices, where this number is a Poissonian random variable with average c. The distribution p_n over the order n of the resulting tree is conveniently analyzed in terms of the generating function $F(z)=\Sigma_n p_n z^n$, which must satisfy

$$F(z)=z\exp[c(F(z)-1)]. \tag{2}$$

This can be solved iteratively for each z, and $F(z)$ must be a stable fixed point of the corresponding iterated map; it is easy to see that this implies $|F(z)|<1/|c|$. For $|z|\leqslant 1$, there is a unique solution for $F(z)$, reachable from the starting point 0, given by $F(z)=\hat{C}(z)/c$, where $\hat{C}(z)$ is the unique solution to $xe^{-x}=zce^{-c}$ in the unit disk. Expanding the corresponding inverse of xe^{-x} yields the exact result $p_n=n^{n-1}c^{n-1}e^{-nc}/n!$.

Particularly interesting is the result for $z=1$, defining $f\equiv F(1)$, which represents the total probability and might be expected to be 1, which is an obvious solution to Eq. (2) for $z=1$. Indeed, for $c<1$, this solution is stable, but for $c>1$, it becomes unstable, and another fixed point becomes the attractor, given by $f=\hat{c}/c$, where $\hat{c}$ is the unique solution to $\hat{c}e^{-\hat{c}}=ce^{-c}$ in the interval $[0,1]$.

Thus, for $c<1$ the branching model is subcritical, and always terminates after a finite number of steps, while for $c>1$ it is supercritical — the deficit in total probability is due to a finite probability $1-f$ that the order n of the generated tree becomes infinite, i.e., that the branching process never terminates.

For a large but finite N, this corresponds to all components being small, i.e. $o(N)$, for $c<1$, while for $c>1$ there exists a single giant component of order $\sim N(1-f)$ with the remaining components being small, having an order distribution similar to that obtained for the complementary c value $\hat{c}$.

III. GENERALIZATION TO INHOMOGENEOUS RANDOM GRAPHS

The classical RG model can be generalized in a straightforward way to inhomogeneous graphs by assuming that vertices can come in different *types* $i\in\{1,\ldots,K\}$. This enables us to consider a very general class of inhomogeneous RG models, to be referred to as IRG:

Definition 2. Given a positive integer K, a K-dimensional vector $\mathbf{r}=\{r_1,\ldots,r_K\}$ of positive probabilities summing to 1, and a symmetric $K\times K$ *matrix* $\mathbf{c}$ with non-negative elements c_{ij}, let $\mathcal{G}(N,K,\mathbf{r},\mathbf{c})$ denote the ensemble of graphs G of order N, defined as follows.

(i) Each vertex is independently assigned a type $i\in\{1,\ldots,K\}$ with probability r_i.

(ii) Independently for each unordered pair of vertices, the corresponding undirected edge is realized with probability $p_{ij}=c_{ij}/N$, *where* (i,j) is the corresponding pair of vertex types.

Remark 1. An asymptotically equivalent alternative is to fix the number of vertices of each type to certain values $N_i\approx Nr_i$, and possibly also the number of edges between vertices of types i,j to fixed values $E_{ij}\approx(1-\delta_{ij}/2)c_{ij}N_iN_j/N$.

A. Revelation of a connected component

In analogy to the classical $\mathcal{G}(N,c)$ model, the model $\mathcal{G}(N,K,\mathbf{r},\mathbf{c})$ can be analyzed by recursively revealing a connected component by exploring neighbors, starting from a single vertex. Let $n_{i,k}$ be the number of new vertices of type i revealed in the kth stage of the revelation (so $n_{i,0}=\delta_{i,i_0}$, with i_0 the type of the starting vertex). Given the number of revealed vertices of different types in the previous stages, $n_{i,k}$ obeys the conditional distribution

$$P(n_{i,k})=\begin{pmatrix} N_i-\sum_{l=0}^{k-1} n_{i,l} \\ n_{i,k}\end{pmatrix}\left(1-\prod_j q_{ij}^{n_{j,k-1}}\right)^{n_{i,k}} \times\left(\prod_j q_{ij}^{n_{j,k-1}}\right)^{N_i-\sum_{l=0}^{k} n_{i,l}}, \tag{3}$$

where $q_{ij}=1-p_{ij}=1-c_{ij}/N\approx\exp(-c_{ij}/N)$. This expression can be simplified in different domains.

B. Small components and the branching process approximation

As long as the order of the revealed part is small as compared to N, we can approximate Eq. (3) by the Poisson distribution

$$P(n_{i,k})\approx e^{-\sum_j r_i c_{ij} n_{j,k-1}}\frac{\left(\sum_j r_i c_{ij} n_{j,k-1}\right)^{n_{i,k}}}{n_{i,k}!}. \tag{4}$$

This corresponds to approximate the revelation process by a Poissonian random branching process.

For the distribution $P(n)$ of the total number n of generated vertices stating from a random vertex, we can define a *generating function* $F(z)=\Sigma_n P(n)z^n$. Since the distribution will depend on the type of the initial vertex, F must be written as the weighted average of the corresponding generating functions $F_i(z)$ for the distributions $P_i(n)$ conditional on the initial type i, i.e.,

$$F(z)=\sum_i r_i F_i(z), \quad \text{with} \quad F_i(z)=\sum_n P_i(n)z^n. \tag{5}$$

The vector $\mathbf{F}$ having the different F_i as components satisfy the coupled set of equations

$$F_i(z)=z\exp\left(\sum_j c_{ij}r_j(F_j(z)-1)\right), \tag{6}$$

which is the inhomogeneous version of Eq. (2).

Remark 2. A more elaborate set of generating functions $\tilde{F}_i(z_1,\ldots,z_K)=\Sigma_{\mathbf{n}}P_i(\mathbf{n})\Pi_j z_j^{n_j}$ could be defined, using a distinct variable z_i for each type i, with n_i the total number of revealed vertices of type i. These would obey equations

obtained by replacing the "z" after the equal sign in Eq. (6) by "z_i." Here we do not care about the detailed type content, and the simpler version, $F_i(z)=\tilde{F}_i(z,\ldots,z)$, will suffice.

Interpreting Eq. (6) as a K-dimensional iterated map (replace "=" by ":="), the proper solution is the stable fixed point reached from the starting point $\mathbf{F}=\mathbf{0}$. Particularly interesting is the result for $z=1$, so let $f_i=F_i(1)$, expressing the probability that the branching process will terminate, conditional on the type of the starting vertex, and let the unconditional counterpart be denoted by $f=\Sigma_j r_j f_j=F(1)$. The f_i satisfy the coupled set of equations

$$f_i=\exp\left(\sum_j c_{ij}r_j(f_j-1)\right), \tag{7}$$

with a naive solution $\mathbf{f}=\mathbf{1}$, the stability of which can be analyzed by means of linearization of Eq. (7) around $\mathbf{f}=\mathbf{1}$, yielding $\{c_{ij}r_j\}$ as the relevant matrix. This is all we need in order to pin down the appearance of the giant, as well as its asymptotic size, and we state the result without proof (it follows by analogy to the corresponding result for the classical model):

Theorem 1. (A) The model $\mathcal{G}(N,K,\mathbf{r},\mathbf{c})$ is subcritical if the eigenvalues of the matrix $\{c_{ij}r_j\}$ are all less than one in absolute value; the graphs then a.a.s. possess no giant component. (B) When some eigenvalue is larger than one, the model is supercritical, and the graphs a.a.s. possess a giant component; its number n_i of vertices of type i asymptotically satisfies $n_i/N\sim r_i(1-f_i)$, where the f_i corresponds to a stable solution of Eq. (7).

Here, a.a.s. stands for *asymptotically almost surely*, i.e., with probability $\to 1$ as $N\to\infty$.

Remark 3. It appears natural to require in addition that $\mathbf{c}$ cannot be block diagonalized; otherwise ergodicity would be broken, and the graph would trivially decompose into distinct subgraphs, which could be treated separately.

In the supercritical case, the generating functions $F_i(z)$ can be *renormalized* with f_i, to yield generating functions for the finite (nongiant) component part. Let $\hat{F}_i(z)=F_i(z)/f_i$. Then $\hat{\mathbf{F}}$ is a stable solution of

$$\hat{F}_i(z)=z\exp\left(\sum_j c_{ij}r_j f_j(\hat{F}_j-1)\right), \tag{8}$$

with $\hat{\mathbf{F}}(1)=\mathbf{1}$. This describes a subcritical branching process with renormalized parameters $\hat{r}_i=r_i f_i/\mathbf{r}\cdot\mathbf{f}$ and $\hat{c}_{ij}=c_{ij}\mathbf{r}\cdot\mathbf{f}$. For a finite N, we must have $\hat{N}\sim N\mathbf{r}\cdot\mathbf{f}$, and we see that this conserves $\hat{p}_{ij}=\hat{c}_{ij}/\hat{N}=p_{ij}$; thus, the renormalized model is simply the naive restriction of the original one to the subset of vertices outside the giant component.

C. Large components and the deterministic approximation

When the giant component is revealed, another approximation can be made to Eq. (3). Once the number of revealed vertices become of $O(N)$, the distribution of $n_{i,k}$ becomes sharply peaked around its average, due to a factor of N appearing in the exponent. As a result, the fluctuations become negligible, yielding a deterministic iterative equation for the consecutive revealed numbers. In terms of the fraction $g_{i,k}=1-\Sigma_{l=0}^{k}n_{i,l}/(Nr_i)$ of all vertices of type i not yet revealed after step k, this yields

$$g_{i,k}=g_{i,k-1}\exp\left(\sum_j c_{ij}r_j(g_{j,k-1}-g_{j,k-2})\right), \tag{9}$$

revealing the conserved quantities

$$\mu_i\equiv g_{i,k}\exp\left(\sum_j c_{ij}r_j(1-g_{j,k-1})\right). \tag{10}$$

The values of μ_i must be ~ 1, since their values can only change in an earlier stage when the number of revealed vertices is still small, but then $g\sim 1$; thus, in the large N limit we can safely assume $\mu_i=1$. The two-step recursion (9) reduces to a one-step recursion, taking the form $g_{i,k}=e^{\Sigma_j c_{ij}r_j(g_{j,k-1}-1)}$, which can be seen as iterating the map

$$g_i\to\exp\left(\sum_j c_{ij}r_j(g_j-1)\right) \tag{11}$$

until a stable fixed point is reached. If the model is subcritical, this is given by the trivial fixed point $g_i=1$, whereas for a supercritical model a nontrivial fixed point with $g_i<1$ results, signalling the existence of a giant component containing a fraction $1-g_i$ of the vertices of type i.

Eq. (11) is identical to Eq. (7), which was derived in the limit of small numbers of revealed vertices; thus, we have established the same set of equations in two different limits.

Remark 4. A third, heuristic way of estimating the size of the giant component is as follows. Suppose the giant contains a fraction n_i of the vertices of type i. Then we can estimate its neighborhood, i.e., the set of vertices connected to at least one vertex in the giant (which of course must be the giant itself), as follows, based on the rather bold assumption that the edge probabilities do not depend on whether any or both of its terminal vertices are in the giant: The total number of vertices of type i is Nr_i. For each of these, the probability of not being connected to any of the vertices in the giant is $\exp[-\Sigma_j c_{ij}n_j/N)$. Thus we can expect a number $Nr_i(1-\exp(-\Sigma_j c_{ij}n_j/N)]$ of vertices of type i in the neighborhood, i.e., in the giant. Writing n_i as $Nr_i(1-g_i)$, we recover Eq. (11), in spite of the bold assumptions involved.

D. Extended type spaces

While we have assumed a finite number of types K, defining the *type space* $\mathcal{T}=\mathbb{Z}_K$, the above results should be more or less directly extendable to models where the type space $\mathcal{T}$ is a denumerable infinite set, or even a continuous manifold, under some general conditions yet to be precisely determined.

Definition 3. For a given type space $\mathcal{T}$, with a normalized measure r on $\mathcal{T}$, and a given non-negative symmetric function c(x,y) on $\mathcal{T}^2$, define $\mathcal{G}(N,\mathcal{T},\mathbf{r},\mathbf{c})$ as the ensemble of RGs of order N, where each vertex is independently assigned a

type $x \in \mathcal{T}$ according to $r(x)$, and for each vertex pair the corresponding edge is independently chosen with probability $c(x,y)/N$, with (x,y) the corresponding pair of types.

For the *denumerable* case, $\mathcal{T}=\mathbb{Z}_+$, it appears natural to require the asymptotic degree averages c_{ij}, or at least the total averages $C_i=\Sigma_{j=1}^{\infty} c_{ij} r_j$, to be uniformly bounded. For cases where the elements of **c** are unbounded, an alternative is to regularize p_{ij} for finite N by using $p_{ij}=1-\exp(-c_{ij}/N)$ instead of the unbounded c_{ij}/N.

Also, reasonable care may have to be taken that **c** is sufficiently ergodic. Let t_{ij} be 0 if $c_{ij}=0$, 1 otherwise. The matrix **t** then describes a graph in type space, with $t_{ij}=1$ corresponding to the existence of the edge (i,j). Then, sufficient ergodicity could e.g. mean that this graph should have a finite diameter, i.e. a uniformly bounded distance between vertex pairs.

For the case of a *continuous* type space $\mathcal{T}$, similar care must be taken. In addition, some kind of continuity constraint seems appropriate, both on **c** and **r**.

Note that a continuous $\mathcal{T}$ allows for a continuous *reparametrization invariance*. Thus, for the case of $\mathcal{T}=\mathbb{R}$, assume f to be a strictly increasing, continuously differentiable mapping of $\mathbb{R}$ to itself. Then the model defined by $\hat{c}(x,y)=c(f(x),f(y))$ and $\hat{r}(x)=r(f(x))f'(x)$ is completely equivalent to that with $c(x,y)$ and $r(x)$. Thus, $r(x)$ could be transformed to any desirable normalized distribution on $\mathbb{R}$. In particular, it could be transformed to the uniform distribution on the unit interval, yielding a kind of standard representation of the model. For higher-dimensional manifolds, things are more complicated, and it appears difficult to devise a universal standardization procedure.

A precise determination of feasibility conditions for extended type spaces will be the subject of future work.

E. Degree distributions

Many properties (but not all!) of a graph ensemble are reflected in its asymptotic degree distribution. In IRG, the asymptotic degree distribution p_m is determined by **r** and **c**, and given simply as the weighted average of the type-specific degree distributions $p_{m|i}$, being Poissonian with an average C_i defined by $C_i=\Sigma_j c_{ij} r_j$. The result is

$$p_m=\sum_i r_i \exp(-C_i)\frac{C_i^m}{m!}, \tag{12}$$

with the associated generating function

$$H(z)\equiv\sum_m p_m z^m=\sum_i r_i \exp[C_i(z-1)]. \tag{13}$$

This puts a limitation on the possible degree distributions that can be obtained within IRG: It must be possible to write the distribution as a positive linear combination of Poissonians, i.e.,

$$p_m=\frac{1}{m!}\int_0^{\infty} c^m e^{-c} p(c)\,dc, \tag{14}$$

where $p(c)\geqslant 0$ describes an, *a priori* arbitrary, distribution of type-specific Poissonian degree averages $c=C_i$, assuming the possibility of a continuum of types. This is a kind of smoothness constraint. In particular, it implies

$$p_m^2 \leqslant \frac{m+1}{m} p_{m+1} p_{m-1}, \tag{15}$$

for each $m>0$. While this excludes, e.g., random regular graphs where the degree is fixed, it does allow for a wide class of degree distributions, such as distributions with a power tail, $p_m \propto m^{-\alpha}$ for large m, by letting $p(c)$ having a similar power tail, $p(c)\propto c^{-\alpha}$ for large c.

Note that a particular model in IRG is *not* determined solely by the degree sequence, which depends on c_{ij} only through the average $C_i=\Sigma_j c_{ij} r_j$. This is in contrast to a class of recently considered models [7–10]; such models define a particular subclass of IRG, however, as will be shown below.

IV. SPECIAL CASES OF INTEREST

For $K=1$, of course the known properties of the classical RG model is recovered. Below we will consider a few less trivial examples.

A. Random bipartite graph

Assuming two distinct vertex types, i.e. $K=2$, a simple ensemble of random bipartite graphs results from the choice

$$\mathbf{c}=\begin{pmatrix} 0 & a \\ a & 0 \end{pmatrix}. \tag{16}$$

With an arbitrary choice of type distribution $\mathbf{r}=(r_1,r_2)$, this yields for the asymptotic generating function $\mathbf{F}(z)=(F_1,F_2)$ the equations

$$F_1(z)=z\exp\{a r_2[F_2(z)-1]\}, \tag{17a}$$

$$F_2(z)=z\exp\{a r_1[F_1(z)-1]\}. \tag{17b}$$

For $z=1$, this yields

$$f_1=\exp[a r_2(f_2-1)],\quad f_2=\exp[a r_1(f_1-1)], \tag{18}$$

yielding the critical value of a as $a_c=1/\sqrt{r_1 r_2}$. In the symmetric case of $r_1=r_2=1/2$, we have $a_c=2$, and $f_1=f_2=f$ satisfying $f=\exp[a/2(f-1)]$.

In a similar way, ensembles of random K-partite graphs can be defined, which can be seen as generated by the complete graph $\mathcal{K}_K$. Similarly, ensembles of random graphs based on an arbitrary generating graph can be defined, with **c** proportional to the incidence matrix for the generating graph. A nice twist results from using a random graph as a generator.

B. Rank-1 c matrix

A particularly interesting special case results when **c** has the factorized form $c_{ij}=C_iC_j/\overline{C}$, where $C_i>0$ can be interpreted as a connection tendency for vertices of type i, while $\overline{C}=\Sigma_i r_i C_i$.

Writing the asymptotic generating function as $F(z)=\Sigma_i r_i F_i(z)$, we get for $F_i(z)$ in this case

$$F_i(z)=z\exp\left(\frac{C_i\sum_j r_jC_j(F_j(z)-1)}{\overline{C}}\right), \tag{19}$$

which can be reduced to a single equation for the function $G(z)=\Sigma_i r_i C_i F_i(z)/\overline{C}$, reading

$$G(z)=z\sum_i r_iC_i\exp[C_i(G(z)-1)]/\overline{C}. \tag{20}$$

In terms of the generating function $H(z)$ for the asymptotic degree distribution, Eq. (13), this can be written as

$$G(z)=z\frac{H'(G(z))}{H'(1)}, \tag{21}$$

and in terms of $G(z)$ we have

$$F(z)=zH(G(z)). \tag{22}$$

For $z=1$ in particular, we get for $g=G(1)$ the equation

$$g=\frac{H'(g)}{H'(1)}, \tag{23}$$

and linearization around the trivial solution $g=1$ yields stability for $H''(1)/H'(1)<1$, corresponding to the model being *subcritical* for $\langle C^2\rangle<\langle C\rangle$, which is equivalent to $\langle m^2\rangle<\langle 2m\rangle$ in terms of moments of the degree distribution. With **c** restricted to have rank 1, the resulting models are asymptotically equivalent to models from a superficially very different class of random graphs that has recently attracted some attention [7–10]. There, a random graph ensemble based on an arbitrary asymptotic degree distribution p_m is defined for a finite order N by randomly selecting a member from the set of graphs with a given degree sequence, such that the number of vertices with degree m is approximately Np_m.

Also for such a model, the recursive exposition of a connected component asymptotically yields a well-defined branching process, apparently very different from the Poissonian ones obtained for IRG. Here, the inital vertex is assigned a random degree m according to p_m, and subsequently branches to m daughter vertices. Each new vertex is independently assigned a degree $n>0$, distributed according to $np_n/\Sigma_m mp_m$ (consistent with the assumption that the asymptotic probability of connecting to a particular vertex is proportional to its degree), and then branches to $n-1$ daughters (since one of its edges is already used).

The asymptotic generating function $F(z)=\Sigma_k P_k z^k$ for the resulting order distribution P_k then satisfies the equation

$$F(z)=z\sum_m p_mG(z)^m, \tag{24}$$

expressing the choice of the initial degree m. Here, $G(z)$ is the *edge generating function*, which satisfies

$$G(z)=z\frac{\sum_m mp_mG(z)^{m-1}}{\sum_m mp_m}, \tag{25}$$

expressing the choice of the daughter's degree m, and its branching to $m-1$ edges.

These are nothing but Eqs. (22) and (21) in disguise, showing the complete asymptotic equivalence of the two models, despite the superficial differences; indeed, the criteria $\langle m(m-2)\rangle<0$ for subcriticality derived above are in complete accordance with the results of Ref. [8].

C. Dynamical random graph with finite memory

The last example is given by a recently proposed class of dynamical random graphs [5] with memory, where a graph is produced starting from a single node according to the combination of three random processes in continuous time, all Poissonian:

(1) For each existing vertex, new, initially isolated, vertices are added at a rate γ.

(2) For each existing vertex, new random edges are added at a rate λ, connecting it to random existing vertices.

(3) Each existing edge is deleted at a rate μ.

It is easy to see that the expected order of the graph grows with time t as $e^{\gamma t}$, and after an initial transient, vertices are only distinguished by their age, and we are led to consider a inhomogeneous model with a continuum of vertex types, $\mathcal{T}=[0,\infty[$, given by vertex age.

1. Asymptotic properties

The probability density for ages x is asymptotically given by

$$r(x)=\gamma e^{-\gamma x}. \tag{26}$$

For each pair of vertices, the probability of a connection is independent of the existence of other connections, and depends on the age x of the youngest vertex involved, and amounts to, at time t, $p(x)=(2\lambda/\gamma-\mu)(e^{(\gamma-\mu)x}-1)e^{-\gamma t}$. We obtain

$$c(x,y)=\frac{2\lambda}{\gamma-\mu}(e^{(\gamma-\mu)\min(x,y)}-1), \tag{27}$$

which seems feasible enough: c is ergodic, continuous and although $c(x,y)$ is not uniformly bounded, the average

$C(x) \equiv \int c(x,y) r(y) dy = (2\lambda/\mu)(1 - e^{-\mu x})$ is (at least for $\mu > 0$). Thus, we are lead to consider the spectrum of the integral kernel

$$G(x,y) = \frac{2\gamma\lambda}{\mu - \gamma}(1 - e^{-(\mu-\gamma)\min(x,y)})e^{-\gamma y}, \tag{28}$$

which is recognized as being proportional to the Green's function (i.e., a kernel representation of the formal operator inverse) for a particular differential operator $\mathcal{L}$ on $\mathbb{R}_+$, given by

$$\mathcal{L} = -\frac{1}{2\gamma\lambda} e^{\mu x} \frac{\partial}{\partial x}\left(\frac{\partial}{\partial x} + \mu - \gamma\right), \tag{29}$$

with boundary conditions $f(0) = 0$, and $f(x)e^{(\mu-\gamma)x/2}$ growing at most as a power of x as $x \to \infty$. Criticality results when the ground state of $\mathcal{L}$ has eigenvalue 1.

With a *finite memory*, $\mu > 0$, the eigenvalue equation for $\mathcal{L}$ is a disguised version of Bessel's equation of order $\gamma/\mu - 1$ in the variable $y = \sqrt{8\lambda\gamma}e^{-\mu x/2}/\mu$, and criticality results when the first positive zero $X_{\gamma/\mu-1}$ of $J_{\gamma/\mu-1}$ is given by $\sqrt{8\lambda\gamma}/\mu$, i.e., for $\lambda = \mu^2 X^2_{\gamma/\mu-1}/8\gamma$.

In the special case of *infinite memory*, $\mu = 0$, the model reduces to a *randomly grown network* [4], and yields

$$\mathcal{L} = -\frac{1}{2\gamma\lambda} \frac{\partial}{\partial x}\left(\frac{\partial}{\partial x} - \gamma\right), \tag{30}$$

with eigenfunctions of the form $e^{\gamma x/2}\sin(\omega x)$ with eigenvalue $(\gamma^2 + 4\omega^2)/8\gamma\lambda$, yielding the ground state value $\gamma/8\lambda$ for $\omega = 0$, and criticality for $\lambda = \gamma/8$.

The above results are all consistent with those obtained in Refs. [4,5] on the phase structure for these models.

V. CONCLUSION

We have investigated a generalization of the classical homogeneous model of sparse random graphs, obtained by imposing a type structure on the vertices. This yields a very general class of inhomogeneous random graph models, and the asymptotic degree distributions are not restricted to Poissonians, but allow for various types of behavior, within certain limitations. Thus, e.g., power behavior is possible, while a fixed degree (regular graph) is ruled out.

The models in this class are *not* determined by the degree distribution alone, but contains an infinity of models for each possible distribution, in contrast to a recently considered class of models based on a given degree distribution. Interestingly enough, a relation does exist, since such models are shown to result in a special case of the present approach.

In other special cases it describes the asymptotic static properties of certain models of evolving random graphs, such as randomly grown networks, and dynamical graphs with memory.

Only certain aspects of the approach have been covered in this paper, and a more detailed analysis, e.g. of the feasibility conditions for extended type spaces, will be the subject of forthcoming work, as will be the investigations on further extensions of the approach e.g. to directed graphs.

ACKNOWLEDGMENTS

The author wishes to thank Tatyana Turova for stimulating discussions — and indeed for an introduction to the subject of random graphs. This work was in part supported by the Swedish Foundation for Strategic Research.

[1] P. Erdős and A. Rényi, Publ. Math. Inst. Hungar. Acad. Sci. **5**, 17 (1960).
[2] B. Bollobàs, *Random Graphs* (Academic Press, London, 1985).
[3] S. Janson, T. Łuczak, and A. Ruciński, *Random Graphs* (Wiley, New York, 2000).
[4] D.S. Callaway, J.E. Hopcroft, J.M. Kleinberg, M.E.J. Newman, and S.H. Strogatz, Phys. Rev. E **64**, 041902 (2001).
[5] T.S. Turova, Phys. Rev. E **65**, 066102 (2002).
[6] R. Albert and A.-L. Barabási, Phys. Rev. Lett. **85**, 5234 (2000).
[7] T. Łuczak, in *Random Graphs*, edited by A.M. Frieze and T. Łuczak (Wiley, New York, 1992), Vol. 2, pp. 165–182.
[8] M. Molloy and B. Reed, Random Struct. Algorithms **6**, 161 (1995).
[9] M. Molloy and B. Reed, Combinatorics, Probab. Comput. **7**, 295 (1998).
[10] M.E.J. Newman, S.H. Strogatz, and D.J. Watts, Phys. Rev. E **64**, 026118 (2001).