

RANDOM NETWORKS WITH CONCAVE PREFERENTIAL ATTACHMENT RULE

by Steffen Dereich (Marburg) and Peter Mörters (Bath)

Version of 15.10.10

Abstract: Many of the phenomena in the complex world in which we live have a rough description as a large network of interacting components. Random network theory tries to describe the global structure of such networks from basic local principles. One such principle is the preferential attachment paradigm which suggests that networks are built by adding nodes and links successively, in such a way that new nodes prefer to be connected to existing nodes if they have a high degree. Our research gives the first comprehensive and mathematically rigorous treatment of the case when this preference follows a *nonlinear*, or more precisely concave, rule. We survey results obtained so far and some ongoing developments.

1. NETWORKS

Although networks are in principle simple objects, they can show immense complexity when their size is very large. They can therefore often offer a meaningful rough description of complex systems in nature, society or technology. Examples of the kind of networks we have in mind are the following:

- a *cell* may be described by its metabolism as a network of chemicals with edges representing chemical reactions transforming one substance into another one;
- a *social network* has human individuals as nodes and edges representing either friendship, acts of communication or other social interactions;
- in a *collaboration graph* the nodes represent scientists which are connected by an edge if they have collaborated or if they have authored a joint paper;
- the *world-wide web* consists of web pages connected by hyperlinks;
- the *internet* has routers and computers as nodes which have physical or wireless links.

The mathematical notion of networks is easy to derive.

Definition 1. A network is defined as a finite set \mathcal{V} of nodes, or vertices, together with a set $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ of links, or edges. The degree of a vertex $v \in \mathcal{V}$ is the number of vertices $w \in \mathcal{V}$ with $(w, v) \in \mathcal{E}$ (called the outdegree) plus the number of vertices $w \in \mathcal{V}$ with $(v, w) \in \mathcal{E}$ (called the indegree). Two vertices $v, w \in \mathcal{V}$ are connected if there exist finitely many vertices $v = v_0, v_1, \dots, v_n = w$ such that for every $i \in \{1, \dots, n\}$ we have $(v_{i-1}, v_i) \in \mathcal{E}$ or $(v_i, v_{i-1}) \in \mathcal{E}$. This defines an equivalence relation on \mathcal{V} and therefore a partition of the set \mathcal{V} of vertices into connected components.

At the beginning of the current millennium scientists from a variety of backgrounds have empirically observed features that many real networks have in common and they have suggested that the same basic principles determine the structure of these networks. The most notable of these features are the following:

- *Power law degree distribution.* In many networks the proportion of vertices having degree k is of the order $k^{-\tau}$, for some exponent $\tau \geq 1$, when the number of nodes is sufficiently large relative to k . Moreover, the power-law exponent τ has been identified as a crucial characteristic of the global network topology. Networks with a power law degree distribution are sometimes also called *scale-free*.
- *Existence of hubs.* In some networks there are a small number of highly connected nodes, which are pivotal in the sense that they are member to many of the shortest paths in the network. Their function is, loosely speaking, to hold the network together. As a result the global topology of the network is on the one hand *robust* under random removal of nodes, but on the other hand *highly vulnerable* under targeted removal of nodes, i.e. removal of the hubs.
- *Clustering.* Edges in large networks are typically not independent. For example, if some node has links to two other nodes, this increases the chance that there exists a link between these two nodes. This is sometimes crudely measured in terms of the clustering coefficient, which gives the proportion of pairs of edges $(u, v), (u, w) \in \mathcal{E}$ for which $(v, w) \in \mathcal{E}$. Loosely speaking, the clustering effect becomes manifest in a clustering coefficient which is not decaying with the network size.
- *Small world phenomenon.* This means that the distance between two randomly chosen nodes, and sometimes also the overall *diameter* of the network, is small. This phenomenon was first observed in social networks and is often confirmed by anecdotal evidence. For example it is said that in Great Britain nobody is further than five handshakes away from the Queen. Mathematically this means that in a network with n vertices, the typical distances are of order $\log n$, and in some cases even of order $\log \log n$.

A mathematical theory of such networks necessarily relies on models which are *probabilistic* in order to be able to remove unnecessary local detail from the picture, and *asymptotic* in the number of nodes, in order to capture effects due to the large size of the networks. The described features are however typically *not* present in the objects which were in the focus of twentieth century random graph theory. In particular, the much studied Erdős-Rényi graphs have empirical distributions converging in a strong sense to a Poisson distribution, which has exponentially decaying tails at infinity, have no hubs and exhibit no clustering effects.

The call for models of networks which have the desired properties was initially answered by some classes of models which were specifically constructed to possess at least one of these features. The most studied models of this type are

- *The configuration model.* Starting point for this model, which was first introduced by Bender and Canfield in [BC78], is a finite sequence d_1, \dots, d_n of nonnegative integers. The i th vertex is then given degree d_i and i ‘half-edges’ are drawn starting in i .

Half-edges are matched uniformly at random with some small modification to avoid multiple edges and unmatchable half-edges. The configuration models with power-law degree sequences are often used to investigate the dependence of quantities like average distances on the power-law exponent, see e.g. [HHZ07].

- *The random intersection graphs.* Each vertex is assigned a set of features and two vertices are connected by an edge if they have a feature in common. This model, first studied by Singer [Sin95], is designed to have the clustering features of real networks but fails to have a power-law degree distribution, see [Sta04].
- *The small-world graphs.* In this model, first introduced by Watts and Strogatz in [WS98], one starts with a regular graph consisting of n cyclically arranged vertices, which are all connected by an edge to their k nearest neighbours. This graph is then successively modified, replacing each edge (v, w) with a small probability p by an edge linking v to a vertex uniformly chosen at random. Watts and Strogatz argue that, for intermediate values of p , the resulting graph is highly clustered and has the small world property.

Neither of these models offer a convincing *explanation* how the specific features of large networks arise. This however, in our opinion, should be one of the main goals of random network theory. Probably the best available explanation, and certainly the most studied one, is the *preferential attachment paradigm*, which we introduce in the next section.

2. THE PREFERENTIAL ATTACHMENT PARADIGM

The ground-breaking idea of Barabási and Albert [BA99] was to base a network model on two simple principles: (i) networks are built dynamically by adding vertices and edges successively, and (ii) new vertices prefer to be attached to existing vertices which already have a high degree in the existing network. Krapivsky, Redner and Leyvraz [KRL00] suggested that the strength of this preference can be modulated by a function $f: \{0, 1, 2, \dots\} \rightarrow (0, \infty)$ in the sense that the probability of a new vertex being linked to an existing vertex is proportional to the function f applied to its degree. This allows some freedom in the actual definition of the model, for example by varying the initial graph, the number of edges established when a new vertex is created or the dependence between those edges. It is however often argued that the qualitative features of the resulting network should *not* depend on the details of the definition, but only on the nature of the attachment function f .

Barabási and Albert [BA99] and their followers argue that, by using a *linear* function f of the current degree, we obtain networks whose degree distribution follows a power law. When f is *superlinear* the behaviour is more extreme and it is shown by Oliveira and Spencer [OS05] that a dominant vertex emerges, which attracts a positive proportion of all future edges. After n steps, this vertex has degree of order n , while the degrees of all other vertices remain bounded. In the most extreme cases eventually all vertices attach to the dominant vertex. The behaviour of the model for superlinear attachment function f is, of course, not in line with the structure of real networks. Therefore the focus of our project is on *sublinear*, or more precisely *concave*, attachment functions f . This allows us to study the transitions between

the case of linear functions f , corresponding to the original suggestion of Barabási and Albert on the one hand, and constant functions f , corresponding to a complete switch-off of the preferential treatment of high degrees and hence behaviour analogous to the classical Erdős–Rényi graphs on the other hand. The main aim of our research is to identify the boundaries between the different types of behaviour in this class of models.

We choose a variant of the model which attaches every new node to a random number of existing nodes in a way which allows maximal independence between edges.

Definition 2. For any $0 \leq \gamma < 1$ we call a concave function $f: \{0, 1, 2, \dots\} \rightarrow (0, \infty)$ with $f(0) \leq 1$ and

$$\Delta f(k) := f(k+1) - f(k) \leq \gamma \quad \text{for all } k \geq 0,$$

a γ -attachment rule, or simply attachment rule. Observe that any f satisfying these conditions is increasing with $f(k) \leq k+1$ for all $k \geq 0$.

Given an attachment rule f , we define a growing sequence $(\mathcal{G}_n)_{n \in \mathbb{N}}$ of random networks by the following iterative scheme:

- The network \mathcal{G}_1 consists of a single vertex (labelled 1) without edges,
- at each time $n \geq 1$, given the network \mathcal{G}_n , we add a new vertex (labelled $n+1$) and
- insert for each old vertex $m \leq n$ independently a directed edge $(m, n+1)$ with probability

$$\frac{f(\text{indegree of } m \text{ at time } n)}{n},$$

to obtain the network \mathcal{G}_{n+1} .

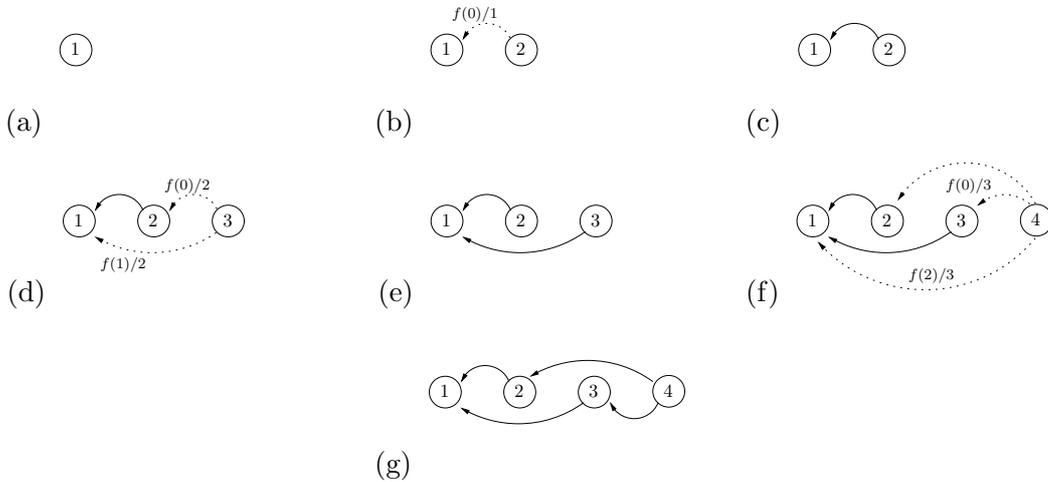


FIGURE 1. (a) We start with a single vertex labelled 1, (b) then the vertex labelled 2 is introduced and linked to 1 with probability $f(0)/1$. (c) Suppose this link is established, then (d) a vertex labelled 3 is introduced and attached to vertex 1 with probability $f(1)/2$ and to vertex 2 with probability $f(0)/2$. (e) Suppose only the former link is established, then (f) a vertex labelled 4 is introduced and attached to vertex 1 with probability $f(2)/3$ and to vertex 2,3 with probability $f(0)/3$. Figure (g) shows the network after four steps.

Remark 1. Edges in the random network \mathcal{G}_n are dependent if they point towards the same vertex and independent otherwise. Formally we are dealing with directed networks, but indeed, by construction, all edges are pointing from the younger to the older vertex, so that the directions can trivially be recreated from the undirected (labelled) graph. All the notions of connectedness, which we discuss in this paper, are based on the *undirected* networks. \diamond

In the following sections we will survey the results obtained in [DM09, DM10, DMM10] on this model. We will give brief sketches of the proofs, which are intended to give a flavour of the breadth of the methods used, and can be skipped by more casual readers.

3. EMPIRICAL DEGREE DISTRIBUTIONS

We start by testing when our model has a power law degree distribution. To this end we denote by $\mathcal{Z}[m, n]$, for $m, n \in \mathbb{N}$, $m \leq n$, the indegree of the m th vertex after the insertion of the n th vertex, and by $X_k(n)$ the proportion of nodes of indegree k at time n , that is

$$X_k(n) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}\{\mathcal{Z}[i, n] = k\}.$$

The random probability vector $X(n) = (X_k(n) : k \in \mathbb{N} \cup \{0\})$ is called the *empirical indegree distribution*. Our first theorem establishes convergence.

Theorem 1 (Asymptotic empirical indegree distribution, see Theorem 1.1(a) in [DM09]).
Let

$$\mu_k = \frac{1}{1 + f(k)} \prod_{l=0}^{k-1} \frac{f(l)}{1 + f(l)} \quad \text{for } k \in \mathbb{N} \cup \{0\},$$

which is a sequence of probability weights. Then, almost surely,

$$\lim_{n \rightarrow \infty} X(n) = \mu$$

in total variation norm.

Remark 2. Rudas et al. in [RTV07] discuss a preferential attachment model in which every new vertex is connected to *exactly one* existing vertex chosen with a probability proportional to a function $w : \mathbb{N} \cup \{0\} \rightarrow (0, \infty)$ of its indegree. The asymptotic indegree distribution they obtain coincides with ours if f is chosen as an appropriate multiple of w . \diamond

Before discussing the implications of this result further we have a look at the outdegrees. Note first that, while the indegree of a vertex is dynamically changing as the network grows, the outdegree is fixed at the moment the vertex is introduced into the graph. The actual distribution of the outdegree of a vertex given the network at the time of its introduction is a rather complex function of this graph. Fortunately, this random distribution converges almost surely to a simple deterministic probability distribution.

Theorem 2 (Asymptotic outdegree distribution, see Theorem 1.1(b) in [DM09]). *The conditional distribution of the outdegree of the $(n+1)$ st incoming node given the graph at time n converges almost surely in the total variation norm to the Poisson distribution with parameter*

$$\lambda := \int f d\mu < \infty,$$

where μ is the asymptotic indegree distribution defined in Theorem 1.

Remark 3. As the outdegree distribution of vertices is asymptotically Poissonian and therefore very light-tailed, it is the asymptotic empirical *indegree* distribution that determines the nature of our network. For example, the question whether a network has power-law behaviour can be decided from the asymptotic indegree distribution μ . \diamond

Example 1. Suppose $f(k) = \gamma k + \beta$ for fixed $\gamma, \beta \in (0, 1]$ and for all $k \in \mathbb{N} \cup \{0\}$. Then the asymptotic empirical indegree distribution can be expressed in terms of the Γ -function,

$$\mu_k = \frac{1}{\gamma} \frac{\Gamma(k + \frac{\beta}{\gamma}) \Gamma(\frac{\beta+1}{\gamma})}{\Gamma(k + \frac{1+\beta+\gamma}{\gamma}) \Gamma(\frac{\beta}{\gamma})}$$

By Stirling's formula, $\Gamma(t+a)/\Gamma(t) \sim t^a$ as t tends to infinity. Hence,

$$\mu_k \sim \frac{\Gamma(\frac{\beta+1}{\gamma})}{\gamma \Gamma(\frac{\beta}{\gamma})} k^{-(1+\frac{1}{\gamma})}, \quad \text{as } k \rightarrow \infty. \quad (3.1)$$

This is in line with earlier results for linear attachment functions, see Katona and Móri [KM06]. For a class of models where new vertices connect to a fixed number m of old ones chosen with a probability proportional to their degree plus a constant $a > -m$, Bollobás et al. [BRST01], Móri [Mór02] and van der Hofstad [Hof09] obtain analogues of Theorem 1 with degree sequences (μ_k) of order $k^{-(3+a/m)}$. The tail behaviour of these models thus coincides if

$$\gamma = \frac{1}{2 + a/m}.$$

Example 2. Suppose $f(k) \sim \gamma k^\alpha$, for $0 < \alpha < 1$ and $\gamma > 0$, then a straightforward analysis yields that

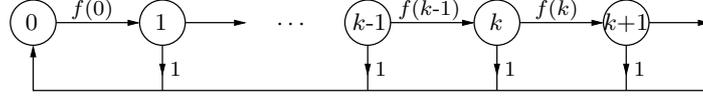
$$\log \mu_k \sim - \sum_{l=1}^{k+1} \log(1 + (\gamma l^\alpha)^{-1}) \sim -\frac{1}{\gamma} \frac{1}{1-\alpha} k^{1-\alpha}.$$

Hence the asymptotic degree distribution has stretched exponential tails. In some cases of real networks this type of tails, rather than power-law behaviour, can be justified.

Sketch of proof. Theorems 1 and 2 are not hard to prove. We can construct an inhomogeneous Markov chain such that at time n the state of the chain is the indegree of a uniformly chosen vertex of \mathcal{G}_n . Indeed, this chain starts with $X_1 = 0$, and if $X_n = k$ we pick

$$X_{n+1} = \begin{cases} k+1 & \text{with probability } \frac{f(k)}{n} \frac{n}{n+1}, \\ k & \text{with probability } (1 - \frac{f(k)}{n}) \frac{n}{n+1}, \\ 0 & \text{with probability } \frac{1}{n+1}. \end{cases}$$

For large times the chain can be approximated by the homogeneous continuous-time Markov chain given by the following diagram:



It is easily checked that the probability distribution $(\mu_k : k \geq 0)$ is the invariant distribution of this continuous-time Markov chain and the *expected* empirical indegree distribution converges to $(\mu_k : k \geq 0)$. An application of Chernoff's inequality shows that this convergence extends to the empirical indegree distributions themselves, proving Theorem 1.

To understand Theorem 2 we observe that, given \mathcal{G}_n , the outdegree of vertex n is the sum of n independent Bernoulli variables with individual means going to zero uniformly and overall mean $\int f dX(n)$. Hence it converges in distribution to a Poisson random variable with mean $\int f dX(n)$, if this limit exists. It remains to use a (slightly technical) majorisation argument to show that $\lim \int f dX(n) = \int f d\mu$, where $\mu = \lim X(n)$ by Theorem 1.

4. EXISTENCE OF A PERPETUAL HUB

In this section we ask whether the network has a hub in the following sense.

Definition 3. *A vertex m is a perpetual hub if there is a time $N \geq m$ such that m is the vertex of maximal indegree in the network \mathcal{G}_n for all times $n \geq N$.*

The existence of a such a perpetual hub shows that there is a single dominant vertex in the network, though this vertex is a hub in a much weaker sense than that described in the introduction. This problem allows us to identify an interesting phase transition in the behaviour of preferential attachment networks.

Theorem 3 (Existence of a perpetual hub, see Theorem 1.5 in [DM09]).

A perpetual hub exists almost surely, if and only if

$$\sum_{k=0}^{\infty} \frac{1}{f(k)^2} < \infty.$$

Otherwise, the time of birth (or, in other words, the label) of any vertex of maximal indegree in \mathcal{G}_n tends to infinity in probability, as n goes to infinity.

Example 3. For special choices of f a much finer analysis of the behaviour of the vertex $m(n)$ of maximal indegree at time n is possible using large deviations theory. Let us assume, for example, that

$$f(k) \sim c k^\alpha \quad \text{for some } \alpha < \frac{1}{2},$$

which by Theorem 3 implies that no perpetual hub exists. Then

$$\log m(n) \sim \frac{1}{2} \frac{(1-\alpha)^{\frac{1-2\alpha}{1-\alpha}}}{(1-2\alpha)c^{\frac{1}{1-\alpha}}} \times (\log n)^{\frac{1-2\alpha}{1-\alpha}},$$

and the maximal indegree $M(n)$ of the graph at time n satisfies

$$M(n) = (c(1 - \alpha)\log n)^{\frac{1}{1-\alpha}} + (1 + o(1)) \frac{1}{2} \frac{1 - \alpha}{1 - 2\alpha} \log n.$$

Here the first summand is the typical indegree of the first vertex at time n , and the second summand a positive correction due to the choice of the vertex with maximal indegree. This example is derived from Proposition 1.18 in [DM09].

Sketch of proof. A major advantage of the setup in our model is that the evolution of the indegrees of the vertices are independent and relatively simple. In particular, when we rescale time as

$$\Psi(n) = \sum_{j=1}^{n-1} \frac{1}{j} \sim \log n \quad \text{for } n \in \mathbb{N},$$

and the indegrees as

$$\Phi(k) = \sum_{j=0}^{k-1} \frac{1}{f(j)} \quad \text{for } k \in \mathbb{N} \cup \{0\},$$

then a law of large numbers applies and shows that, for every $m \in \mathbb{N}$, we have

$$\lim_{n \rightarrow \infty} \frac{\Phi(\mathcal{Z}[m, n])}{\Psi(n)} = 1 \quad \text{almost surely.}$$

Moreover, denoting $\pi(t) = n$, where $n \in \mathbb{N}$ is maximal with $\Psi(n) \leq t$, we can write the evolution of the indegrees in the new scale

$$X_s(t) = \Phi(\mathcal{Z}[\pi(s), \pi(t)]) \quad \text{if } t \geq s \geq 0.$$

The fluctuations

$$M_s(t) = \begin{cases} 0 & \text{if } 0 \leq t \leq s \\ X_s(t) - (\Psi(\pi(t)) - \Psi(\pi(s))) & \text{if } t \geq s, \end{cases}$$

define a martingale. By estimating its quadratic variation process we see that this martingale converges almost surely if and only if the series in Theorem 3 converges.

First suppose the series converges. Then $M_0(t)$ converges to a finite limit, say $M[1]$, which means that the scaled indegree of the vertex labelled 1 at time $n \geq 1$ is

$$\Psi(n) + M[1] + o(1).$$

Now look at any competing vertex, say with label m , and note that the martingale $M_s(t)$ converges to a finite limit, say $M[m]$. Hence, at time $n \geq m$ its scaled indegree is

$$\Psi(n) - \Psi(m) + M[m] + o(1)$$

and in order to have a higher indegree than the first vertex for more than just a finite number of times, the martingale limit associated with vertex m needs to satisfy $M[m] \geq \Psi(m) + M[1]$. We see that the summand $\Psi[m]$ acts as a ‘handicap’, which is the larger the later a particle is born. A technical a-priori estimate ensures that there is a (random) threshold $M \in \mathbb{N}$ such that vertices born after time M can no longer compete. We then show that the distributions of the martingale limits have no atoms, which ensures that there is a unique vertex $m \in \{1, \dots, M\}$ maximising $M[m] - \Psi(m)$, and this vertex is therefore the perpetual hub.

Now suppose that the series does not converge. By the martingale central limit theorem, for any label m and $s = \Psi(m)$, the martingale $(M_s(t) : t \geq 0)$ satisfies

$$\left(\frac{1}{\sqrt{T}} M_s(\phi_{Tt}^{-1}) : t \geq 0\right) \Longrightarrow (W_t : t \geq 0),$$

where convergence is in law on the Skorokhod space, the process on the right is standard Brownian motion and

$$\phi_t = \int_0^{\Phi^{-1}(t)} \frac{1}{f(\lfloor u \rfloor)^2} du,$$

is scaling independent of the choice of m . This implies that also

$$\left(\frac{1}{\sqrt{T}} (X_0(\phi_{Tt}^{-1}) - X_s(\phi_{Tt}^{-1})) : t \geq 0\right)$$

converges to Brownian motion, and thus changes sign infinitely often. Hence for every pair of vertices, the lead in the competition for the highest indegree changes hands infinitely often, and no vertex can be a perpetual hub.

5. EXISTENCE OF A GIANT COMPONENT

We now address the question of connectedness in our network.

Definition 4. Denote by \mathcal{C}_n the largest connected component in the network at time n (as defined in Definition 1). We say that a giant component exists in the network if there exists $p > 0$ such that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \#\mathcal{C}_n = p \text{ in probability.}$$

Note that not the existence of the limit is the crucial part in this definition, but the fact that the largest connected component comprises a *positive fraction* of the nodes.

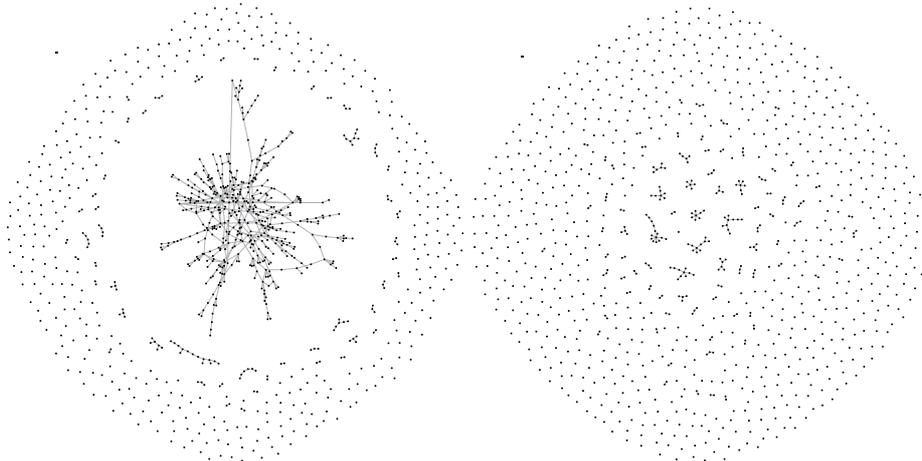


FIGURE 2. Simulation of the network with $f(k) = \frac{1}{2}\sqrt{k} + x$, for $x = \frac{2}{5}, \frac{1}{10}$ and 1000 vertices, generated using the *Network Workbench Tool*, see [Tea06]. On the left, the giant component is displayed in the centre surrounded by the smaller components. On the right, no giant component exists.

In network models where a new vertex establishes an edge to at least one old vertex, the entire network is connected. In analogy to the Erdős-Rényi model, we also expect in relatively homogeneous network models that if the mean number of edges per vertex exceeds one, there exists a giant component. We will see in Example 4 below that this is indeed the case in our model. If however the expected number of edges per vertex is less than one, in general the question of existence of a giant component becomes very subtle, in particular in preferential attachment models where the principles underlying the construction are beneficial to the global connectivity. Before stating our most general result on the existence of a giant component, we look at the linear case $f(k) = \gamma k + \beta$, in which we obtain an explicit answer.

Theorem 4 (Existence of a giant component: linear case, see Proposition 1.2 in [DM10]). *If $f(k) = \gamma k + \beta$ for some $0 \leq \gamma < 1, 0 < \beta \leq 1$ then there exists a giant component if and only if*

$$\gamma \geq \frac{1}{2} \text{ or } \beta > \beta_{\text{crit}} := \frac{(\frac{1}{2} - \gamma)^2}{1 - \gamma}.$$

Remark 4. This result corresponds to the following intuition: If $\gamma \geq \frac{1}{2}$ the preferential attachment is sufficiently strong to establish a giant component in the network for purely topological reasons and regardless of the edge density. However if $\gamma < \frac{1}{2}$ then the preferential attachment is weak and a giant component exists only if the edge density is large enough, i.e. if β exceeds the critical value β_{crit} . \diamond

The formulation of the result for the nonlinear case requires the study of some operators, whose precise rôle in the network will become apparent in the sketch of the proof. For the formal definition we use a pure birth Markov process $(Z_t: t \geq 0)$ started in zero with generator

$$Lg(k) = f(k) \Delta g(k),$$

which means that the process leaves state k with rate $f(k)$. Given a suitable parameter $0 < \alpha < 1$ we define a linear operator A_α on the Banach space $\mathbf{C}(\mathcal{S})$ of continuous, bounded functions on $\mathcal{S} := \{\ell\} \cup [0, \infty]$, by

$$A_\alpha g(\tau) := \int_0^\infty g(t) e^{\alpha t} dM(t) + \int_0^\infty g(\ell) e^{-\alpha t} dM^\tau(t),$$

where increasing functions M , resp. M^τ , $\tau \in \mathcal{S}$ are given by

$$M(t) = \int_0^t e^{-s} \mathbb{E}[f(Z_s)] ds, \quad M^\ell(t) = \mathbb{E}[Z_t],$$

$$M^\tau(t) = \mathbb{E}[Z_t | \Delta Z_\tau = 1] - \mathbf{1}_{[\tau, \infty)}(t) \quad \text{for } \tau \in [0, \infty),$$

$$M^\infty(t) = \lim_{\tau \rightarrow \infty} M^\tau(t).$$

If A_α is well-defined it is a compact operator. We denote by \mathcal{I} the set of parameters where this is the case, and note that it is an open (but possibly empty) subinterval of $(0, 1)$.

Theorem 5 (Existence of a giant component, see Theorem 1.1 in [DM10]).

A giant component exists if and only if for all $\alpha \in \mathcal{I}$ the spectral radius $\rho(A_\alpha)$ exceeds one.

Remark 5. Observe that the criterion above holds when $\mathcal{I} = \emptyset$. ◇

The drawback of Theorem 5 is that the general criterion for existence of a giant component is infinite-dimensional. However at what appears to be a small prize in accuracy we can replace it by explicit, finite-dimensional necessary or sufficient criteria. The criterion for existence in the following theorem is very accurate, but the criterion for nonexistence is quite crude. It should suffice to give a first impression.

Theorem 6 (Sufficient/necessary criteria, derived from Proposition 1.9. in [DM10]).

If

$$\sum_{k=0}^{\infty} \prod_{j=0}^k \frac{f(j)}{\frac{1}{2} + f(j)} > \frac{1}{2},$$

then there exists a giant component. However, if

$$\sum_{k=0}^{\infty} \prod_{j=0}^k \frac{f(j+1)}{\frac{1}{2} + f(j+1)} \leq \frac{1}{2},$$

then there exists no giant component.

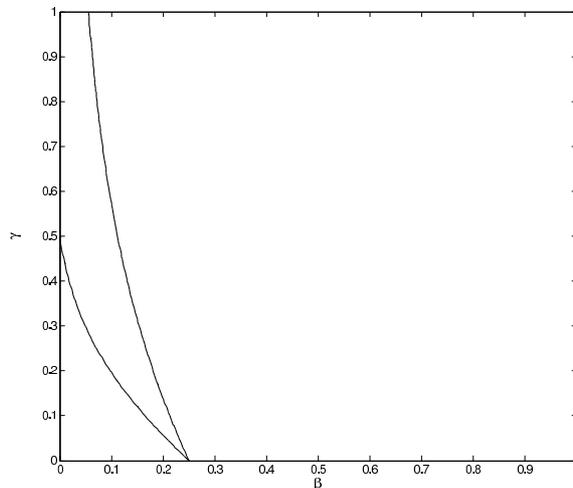


FIGURE 3. For the attachment function $f(k) = \gamma\sqrt{k} + \beta$ we are interested in the boundary between the phase of nonexistence and the phase of existence of the giant component in the (β, γ) -plane. The figure shows a lower and upper bound for this boundary, using the criteria of Theorem 6.

Example 4. If the asymptotic mean outdegree of a vertex in our network is at least $\frac{1}{2}$, and hence the asymptotic mean degree per vertex is at least one, a giant component exists. Indeed, by Theorem 2, the asymptotic mean outdegree distribution of a vertex equals

$$\langle f, \mu \rangle = \sum_{k=0}^{\infty} \prod_{l=0}^k \frac{f(l)}{1 + f(l)},$$

and hence the assumption implies

$$\sum_{k=0}^{\infty} \prod_{l=0}^k \frac{f(l)}{\frac{1}{2} + f(l)} > \sum_{k=0}^{\infty} \prod_{l=0}^k \frac{f(l)}{1 + f(l)} \geq \frac{1}{2},$$

so that a giant component exists by Theorem 6.

Sketch of proof. The key step is to explore the neighbourhood of a vertex in the network. Starting from a vertex v chosen uniformly from all vertices, we consider all the vertices which are connected to v by a single edge as the offspring of v . For each offspring vertex, we consider the vertices other than v which are connected to it by a single edge as its offspring, and for a while we can go on without creating a circle, describing the local neighbourhood of v by a tree. The key step in the proof is to define a *coupling* of this tree with the genealogical tree of a branching random walk, called the *idealised neighbourhood tree*, in such a way that the probability that the two trees match up to the occurrence of one of three stopping events explained below converges to one. The three stopping events are

- (A) the number of vertices discovered exceeds a threshold value,
- (B) a very old vertex is discovered, and
- (C) there are no more vertices left to discover.

The local nature of the coupling allows us only to use thresholds of order $\log n$. Stopping in event (A) ensures that vertex v is in a component of at least threshold value, as does stopping in event (B), as a rough a-priori estimate shows that sufficiently old vertices are always in components exceeding the threshold size. Stopping in event (C) however ensures that v is not in such a component. Denoting by $C_n(v)$ the size of the connected component of \mathcal{G}_n containing v , we thus infer that, as $n \uparrow \infty$, we have

$$\mathbb{E} \left[\frac{1}{n} \sum_{v=1}^n \mathbf{1} \{ C_n(v) \text{ exceeds threshold} \} \right] \sim \mathbb{P}(\text{coupling ends in (A) or (B)}).$$

A (tricky) variance estimate ensures that we can omit the expectation on the left and keep the equivalence with probability going to one. We use this to obtain information about the giant component using a technique called *sprinkling*: if a sufficiently large proportion of vertices are contained in medium sized components, then for a slightly increased attachment function these vertices are all connected. Roughly speaking we obtain from this that, with probability going to one, the proportion of vertices in the largest component of the network is asymptotically equivalent to the probability that our coupling ends in events (A) or (B).

To investigate the limit of the latter quantity we turn our attention to the idealised neighbourhood tree and first state its definition in terms of a multitype branching random walk. The particles in this process have *locations* on the real line and *types* in the space $\mathcal{S} = \{\ell\} \cup [0, \infty]$. The type of a particle is ℓ if its parent is located to its left, and otherwise it is the distance between the locations of the particle and its parent. The initial particle is of type ℓ with starting position $-X$, where X is standard exponentially distributed. Recall the definition of the pure birth Markov process $(Z_t: t \geq 0)$ and, for any $\tau \geq 0$, let $(Z_t^{(\tau)}: t \geq 0)$ be the same process conditioned to have a birth at time τ .

Each particle of type ℓ in position x generates offspring

- to its left with relative positions distributed according to a Poisson point process on $(-\infty, 0]$ with intensity measure $e^t \mathbb{E}[f(Z_{-t})] dt$;
- to its right with relative positions distributed like the jumps of the process $(Z_t: t \geq 0)$.

Each particle of type $\tau \geq 0$ in position x generates offspring

- to its left in the same manner as with a parent of type ℓ ;
- to its right with relative positions distributed like the jumps of the process $(Z_t^{(\tau)} - \mathbf{1}_{[\tau, \infty)}(t): t \geq 0)$.

Note that in this branching random walk every particle has finite offspring to the left, but *infinite* offspring to the right. The idealised neighbourhood tree arises when particles in the branching random walk (and all their descendants) are killed when they are located to the right of the origin.

It is not hard to figure out how this branching random walk arises as a limit of the local neighbourhoods in our network. Representing the location of a vertex $m \leq n$ at time n as $\Psi(m) - \Psi(n)$, where Ψ is the function used to rescale time in Section 4, we obtain that the initial vertex v , chosen uniformly from $\{1, \dots, n\}$, is mapped onto $\Psi(v) - \Psi(n)$, which is asymptotically distributed like $-X$. The pure birth processes describing the relative positions of its offspring to the *left* has independent number of jumps in disjoint sets and hence converges in law to a Poisson point process on $(-\infty, 0]$ with intensity measure $e^t \mathbb{E}[f(Z_{-t})] dt$. The pure birth processes describing the relative positions of its offspring to the *right* converges in law to the process $(Z_t: t \geq 0)$. The only relevant interference from previous exploration which influences the offspring distributions of later vertices is the existence of an edge connecting them to their parent vertex, if it is located to the right. It leads to an extra jump in the process $(Z_t: t \geq 0)$ at the relative position of the father, which influences the distribution but has to be discarded from the asymptotic offspring law.

The probability that our coupling ends in events (A) or (B) converges, as the network size goes to infinity, to the probability that the idealised neighbourhood tree is infinite. A giant component exists, if this ‘survival’ probability is positive, and in this case this is also the asymptotic proportion of vertices in the giant component. Techniques from branching random walks then enable us to give a necessary and sufficient criterion for a positive survival probability: These are based on introducing a score of $e^{-\alpha t}$ for a particle in the population located at position $t \in \mathbb{R}$, i.e. the further a particle is to the left the higher its score.

Loosely speaking, $A_\alpha \mathbf{1}_\sigma(\tau)$ is the expected score of the offspring of type σ by a particle of type τ located at the origin. If, for some α , the operator A_α has a spectral radius less or equal to one, then the overall score of the population goes to zero, which means that the positions of all particles go to infinity and the killed process dies out. Conversely, if for all α the spectral radius of A_α exceeds one, this can be used to construct a particle whose location goes to $-\infty$, meaning that the killed process survives if the initial particle has started sufficiently far to the right. This characterises survival of the idealised neighbourhood tree in terms of the spectral radii of the operators ($A_\alpha: \alpha \in \mathcal{I}$), finishing the sketch of the proof of Theorem 5.

In the linear case, it turns out that conditioning ($Z_t: t \geq 0$) on having a birth at time $\tau \geq 0$ and then subtracting $\mathbf{1}_{[\tau, \infty)}(t)$ leads to the same process for all values of τ . Hence one can collapse the typespace \mathcal{S} to just two elements, corresponding to left and right descendants, and calculate the spectral radius of the operator A_α for any α explicitly, which leads to the result stated in Theorem 4. In the nonlinear case relatively rough upper and lower bounds for the spectral radius of A_α give the criteria in Theorem 6.

6. ROBUSTNESS AND PERCOLATION

As the discussions of the previous section indicate, there are two different reasons why we may have a giant component: In some situations there exist a relatively small number of highly connected vertices linking up the vertices in the giant component of the network regardless of the overall edge density. In other situations there is a critical edge density marking the transition between nonexistence and existence of a giant component. Bollobás et al. [BR03] have shown that for their particular linear preferential attachment model the first situation prevails. To do this they use percolation, i.e. they randomly remove edges from the network, to reduce the edge density and then test whether there is still a giant component. If this is the case, they call the giant component robust. In this section we will perform this test of robustness for our model and obtain a full and explicit characterization of the robust case.

Definition 5. *For a fixed parameter $0 < p < 1$, remove every edge in the network independently with probability $1 - p$ and call the resulting network the percolated network. We say that the giant component in a network is robust, if, for every $0 < p < 1$, the percolated network has a giant component.*

Theorem 7 (Robustness, see Theorem 1.5 in [DM10]). *The giant component in the network is robust if and only if*

$$\inf_{n \in \mathbb{N}} \Delta f(n) \geq \frac{1}{2}.$$

In the non-robust case we can identify the critical percolation parameter in terms of the spectral radii of the operators ($A_\alpha: \alpha \in \mathcal{I}$).

Theorem 8 (Percolation, see Remark 1.6 in [DM10]). *The percolated network has a giant component if and only if the percolation parameter satisfies*

$$p > \frac{1}{\min_{\alpha \in \mathcal{I}} \rho(A_\alpha)}.$$

Remark 6. In the linear case $f(k) = \gamma k + \beta$ the network is robust if and only if $\gamma \geq \frac{1}{2}$. Moreover the percolated network has a giant component if and only if

$$p > \left(\frac{1}{2\gamma} - 1\right) \left(\sqrt{1 + \frac{\gamma}{\beta}} - 1\right),$$

where the right hand side has to be interpreted as $\frac{1}{4\beta}$ if $\gamma = 0$. \diamond

Sketch of proof. As in the previous section we can couple the neighbourhood of a uniformly picked vertex in the percolated network and the percolated idealised neighbourhood tree, i.e. the component of the root in the remainder of the idealised neighbourhood tree when every edge is removed independently with probability $1 - p$. An argument similar to the one sketched in Section 5 shows that the percolated network has a giant component if and only if this tree is infinite, which is the case if and only if

$$p > \frac{1}{\min_{\alpha \in \mathcal{I}} \rho(A_\alpha)},$$

with the right hand side interpreted as zero if $\mathcal{I} = \emptyset$.

By Theorem 6, the value of the series

$$2 \sum_{k=0}^{\infty} \prod_{j=0}^k \frac{f(j)}{\frac{1}{2} + f(j)}$$

is a lower bound for the minimal spectral radius. This series diverges if

$$\inf_{n \in \mathbb{N}} \Delta f(n) = \lim_{n \rightarrow \infty} \frac{f(n)}{n} \geq \frac{1}{2},$$

where the equality follows from concavity of f . Conversely, if $\inf_n \Delta f(n) < \frac{1}{2}$, the series

$$2 \sum_{k=0}^{\infty} \prod_{j=0}^k \frac{f(j+1)}{\frac{1}{2} + f(j+1)}$$

converges. Again by Theorem 6 this series is an upper bound for the minimal spectral radius. Hence for small values of p the idealised neighbourhood tree fails to be infinite, and the network is not robust.

7. SMALL WORLDS

An analysis of the *small world phenomenon* for the given class of preferential attachment models is currently undertaken by Dereich, Mönch and Mörters in [DMM10] and [DMM11]. Recall that the distance of two vertices $v, w \in \mathcal{V}$ in a network with edge set $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ is

$$d(v, w) = \min \left\{ n \geq 0 : \exists v = v_0, v_1, \dots, v_n = w \text{ with } (v_{i-1}, v_i) \in \mathcal{E} \text{ or } (v_i, v_{i-1}) \in \mathcal{E} \forall i \right\}.$$

The investigation in [DMM10, DMM11] is naturally concentrated on the case where a giant component exists, and focuses on the typical distance between vertices in the giant component. We expect the occurrence of three different regimes, depending on the strength of the preference in the attachment rule.

In the case of strongest preferential attachment we have found that typical distances in the random network are doubly logarithmic. Moreover it turns out that the constant in front of the log log-term depends only on the asymptotic growth of the attachment rule.

Theorem 9 (Ultrasmall networks, Theorem 1 in [DMM10]). *If the attachment rule satisfies*

$$\gamma := \inf_{n \in \mathbb{N}} \Delta f(n) > \frac{1}{2},$$

and V, W are two vertices independently picked from the largest connected component in \mathcal{G}_n , then, for every $\varepsilon > 0$,

$$(4 - \varepsilon) \frac{\log \log n}{\log \frac{\gamma}{1-\gamma}} \leq d(V, W) \leq (4 + \varepsilon) \frac{\log \log n}{\log \frac{\gamma}{1-\gamma}},$$

with probability converging to one.

Remark 7. (a) Precisely the same asymptotic behaviour of the typical distance can be shown in a variety of different preferential attachment models, see [DMM10]. This robustness of typical distances is in stark contrast to the situation for the *diameter* of the giant component, which is heavily depend on modeling details. For example, in our model even in the regime of strongest preferential attachment the diameter is logarithmic in the number of nodes, but it is doubly logarithmic in the corresponding regime for the preferential attachment models studied by Dommers et al [DHH10].

(b) Doubly logarithmic typical distances are expected to hold essentially whenever the asymptotic degree sequence $\mu_k \asymp k^{-\tau}$ for $2 < \tau < 3$. In our model this is equivalent to $\gamma > \frac{1}{2}$ (recall formula (3.1) to see this in the case of linear f). This behaviour of typical distances is also confirmed for some other models, in particular the configuration model, see van der Hofstad et al. [HHZ07]. The crucial difference is that in these models the constant 4 in the precise result needs to be replaced by 2, and this corresponds to a different internal structure of the short paths in the network. A full discussion of this interesting phenomenon will be given in [DMM10]. \diamond

Results for other regimes are at present still on a conjectural level and will be discussed in the forthcoming work [DMM11]. We conjecture the existence of two further regimes:

- a regime of *weak* preference, which occurs when

$$\inf_{n \in \mathbb{N}} \Delta f(n) < \frac{1}{2}$$

and which is characterised by the fact that typical distances in the giant component are *logarithmic* in the number of nodes. This case is similar to the configuration model in the case of power-law exponents $\tau > 3$ and all tails that are lighter than

power-laws, see [CL02, HHM05]. A more detailed conjecture is that for vertices V, W picked independently from the largest connected component in \mathcal{G}_n and every $\varepsilon > 0$,

$$\frac{1 - \varepsilon}{\inf_{0 < \alpha < 1} \log \varrho(A_\alpha)} \log n \leq d(V, W) \leq \frac{1 + \varepsilon}{\inf_{0 < \alpha < 1} \log \varrho(A_\alpha)} \log n,$$

with probability converging to one.

- a *critical* regime, which occurs when

$$\inf_{n \in \mathbb{N}} \Delta f(n) = \frac{1}{2},$$

in which very rich and complex behaviour of the typical distances in the giant component is possible, depending on the fine structure of f . We expect however that, for every $\varepsilon > 0$,

$$d(V, W) \leq (1 + \varepsilon) \frac{\log n}{\log \log n},$$

with probability converging to one and that a corresponding lower bound holds in the linear case.

Some progress in these questions has been made in the linear case for different variants of the model by Bollobás and Riordan [BR04] and by Dommers et al. [DHH10]. The set-up of our model appears to be more favourable and therefore one would hope to obtain more precise results covering also truly nonlinear cases.

8. SOME REMARKS ON CLUSTERING

The *clustering properties* of preferential attachment networks with a concave attachment rule also warrant a thorough discussion. Note that our model has nontrivial edge dependence: conditional on the existence of the edge $(v, w) \in \mathcal{E}$ we have an increased likelihood of the existence of edges of the form (v, w') for $w' \in \mathcal{V}$. But this does not lead to large clustering coefficients, which are too low compared to those empirically found in real networks. It can be suspected that while the preferential attachment paradigm leads to networks with macroscopic features that match those of the examples listed in the introduction rather closely, quantities relying strongly on the local neighbourhoods of vertices like the clustering coefficients may not be well represented in our model.

The natural solution to this problem is to enhance the preferential attachment model, for example by giving nodes characteristic features that influence the attachment of other nodes or by favouring attachment to nodes which are directly linked to vertices of high degree. This may well lead to richer models which are still within the scope of rigorous mathematical treatment. Needless to say that scientists have pursued this line already to some extent, using nonrigorous analysis as well as empirical and numerical studies. But a lot of very interesting rigorous mathematics remains still to be done on this question.

Acknowledgements: We would like to thank Christian Mönch for providing Figure 2 and Figure 3 and for his permission to include unpublished joint results. The second author would like to thank EPSRC for support through an *Advanced Research Fellowship*.

REFERENCES

- [BA99] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.
- [BC78] E. A. Bender and E. R. Canfield. The asymptotic number of labelled graphs with a given degree sequence. *J. Combin. Theory*, 24:296–307, 1978.
- [BR03] B. Bollobás and O. Riordan. Robustness and vulnerability of scale-free random graphs. *Internet Math.*, 1:1–35, 2003.
- [BR04] B. Bollobás and O. Riordan. The diameter of a scale-free random graph. *Combinatorica*, 24:5–34, 2004.
- [BRST01] B. Bollobás, O. Riordan, J. Spencer, and G. Tusnády. The degree sequence of a scale-free random graph process. *Random Structures Algorithms*, 18:279–290, 2001.
- [CL02] F. Chung and L. Lu. The average distances in random graphs with given expected degrees. *Proc. Natl. Acad. Sci. USA*, 99:15879–15882, 2002.
- [DHH10] S. Dommers, R. van der Hofstad, and G. Hooghiemstra. Diameters in preferential attachment models. *J. Stat. Phys.*, 139:72–107, 2010.
- [DM09] S. Dereich and P. Mörters. Random networks with sublinear preferential attachment: degree evolutions. *Electron. J. Probab.*, 14:1222–1267, 2009.
- [DM10] S. Dereich and P. Mörters. Random networks with sublinear preferential attachment: the giant component. *Preprint arXiv:1007.0899*, 2010.
- [DMM10] S. Dereich, C. Mönch, and P. Mörters. Distances in ultrasmall random networks. *In preparation*, 2010.
- [DMM11] S. Dereich, C. Mönch, and P. Mörters. Random networks with sublinear preferential attachment: small worlds. *In preparation*, 2011.
- [HHM05] R. van der Hofstad, G. Hooghiemstra, and P. van Mieghem. Distances in random graphs with finite variance degrees. *Random Structures Algorithms*, 27:76–123, 2005.
- [HHZ07] R. van der Hofstad, G. Hooghiemstra, and D. Znamenski. Distances in random graphs with finite mean and infinite variance degrees. *Electron. J. Probab.*, 12:703–766, 2007.
- [Hof09] R. van der Hofstad. *Random graphs and complex networks*. Eindhoven, 2009. Lecture Notes.
- [KM06] Z. Katona and T. F. Móri. A new class of scale free random graphs. *Statist. Probab. Lett.*, 76:1587–1593, 2006.
- [KRL00] P. L. Krapivsky, S. Redner, and F. Leyvraz. Connectivity of growing random networks. *Phys. Rev. Lett.*, 85:4629, 2000.
- [Mór02] T. F. Móri. On random trees. *Studia Sci. Math. Hungar.*, 39:143–155, 2002.
- [OS05] R. Oliveira and J. Spencer. Connectivity transitions in networks with super-linear preferential attachment. *Internet Math.*, 2:121–163, 2005.
- [RTV07] A. Rudas, B. Tóth, and B. Valkó. Random trees and general branching processes. *Random Structures and Algorithms*, 31:186–202, 2007.
- [Sin95] K. B. Singer. *Random intersection graphs*. PhD thesis, Johns Hopkins University, 1995.
- [Sta04] D. Stark. The vertex degree distribution of random intersection graphs. *Random Structures Algorithms*, 24:249–258, 2004.
- [Tea06] NWB Team. Network workbench tool. *Indiana University, Northeastern University, and University of Michigan.*, page <http://nwb.slis.indiana.edu>, 2006.
- [WS98] D. J. Watts and S. H. Strogatz. Collective dynamics of ‘small-world’ networks. *Nature*, 393:440–442, 1998.