ON THE MEAN DISTANCE IN SCALE FREE GRAPHS

G. Hooghiemstra and P. Van Mieghem*

Abstract

We consider a graph, where the nodes have a pre-described degree distribution F, and where nodes are randomly connected in accordance to their degree. Based on a recent result [22], we improve the approximation of the mean distance between two randomly chosen nodes given by Newman *et al.* [32]. Our new expression for the mean distance involves the expectation of the logarithm of the limit of a super-critical branching process. We compare simulations of the mean distance with the results of Newman *et al.* and with our new approach.

Keywords and phrases: scale free graphs, mean distance

1 Introduction

In the first subsection of this introduction, we introduce the graph model, discuss its relevance, and present our main results. In Section 1.2, we review related work, whereas in Section 1.3 we explain the organization of this paper.

1.1 Definition and statement of results

In several disciplines it was observed that graph models for real-world phenomena often possess a specific degree distribution. In electrical engineering and computer science, Faloutsos *et al.* [20] have shown that the degree distribution of the Internet graph follows a power law. Since their paper, a large amount of additional measurements to infer properties of the graph of the Internet have been published (see e.g. [34], [36]). In physics, Newman *et al.* [32] have studied real-world networks described by various degree distributions. Two recent review papers on graphs and networks that include a discussion on scale free graphs are [5] and [31]. A general observed characteristic is that most of these 'real-world' degree distributions often differ considerably from the Poisson distribution that asymptotically (for a large number of nodes N) describes the degree distribution of the random graph $G_p(N)$ [8]. Barabasi [7] showed that preferential attachment of nodes gives rise to a class of

*Delft University of Technology, Electrical Engineering, Mathematics and Computer Science, P.O. Box 5031, 2600 GA Delft, The Netherlands. E-mail: G. Hooghiemstra@ewi.tudelft.nl, P.VanMieghem@ewi.tudelft.nl

graphs called 'scaled free networks'. Scale free graphs have a power law degree distribution which contrasts with the degree distribution of the classical random graph $G_p(N)$.

The interest in the understanding of real-world graphs has led to the introduction of a number of different graph models, which we review in Section 1.2 below.

We consider a model which was introduced in a series of papers ([32], [33] and [22]). For a given probability mass function f_j, $j = 1, \ldots,$ and an integer N we define an i.i.d. sequence $D_1, D_2, \ldots, D_N$ with (common) distribution

$$\mathbb{P}[D = j] = f_j, \quad j = 1, 2, \ldots. \tag{1}$$

The construction of the involved graph model starts with N nodes, where node j has D_j (random!) edges (initially called stubs). These stubs are connected by matching them randomly. Each matched pair of stubs form together one edge or link. If the total number of stubs $D_1 + D_2 + \ldots + D_N$ in the random matching construction is odd an additional stub is added to the N^{th} node, so that D_N is increased by one. This single stub will make will make no difference in what follows. We refer to [22, Introduction] for more details.

Let $F(x)$ be the cumulative distribution function of the discrete distribution f_j, i.e., $F(x) = \sum_{j=1}^{\lfloor x \rfloor} f_j$, where $\lfloor x \rfloor$ is the largest integer smaller than or equal to x. As in [22] we will assume that for some $\tau > 3$ and some positive c,

$$1 - F(x) \leq c x^{-\tau+1}, \quad x > 0. \tag{2}$$

This graph model that includes heavy tail degrees with finite variance is a variant of the configuration model, which, given a degree sequence, is the random graph with that given degree sequence. For a graph, the degree sequence of that graph is the vector, of which the k^{th} coordinate equals the frequency of nodes with degree k. In our model, the degree sequence is very close to the distribution of the nodal degree D of which $D_1, \ldots, D_N$ are i.i.d. copies. We further define

$$\mu = \mathbb{E}[D], \qquad \nu = \frac{\mathbb{E}[D(D-1)]}{\mathbb{E}[D]}. \tag{3}$$

The importance of the parameter ν will be clarified below.

The distance or path length between two nodes is defined as the minimal number of distinct edges that forms a path connecting the two nodes. If the two nodes are not connected by a path, we put this distance equal to $+\infty$. The mean of this distance is a characterizing property of a graph. Moreover, the computational complexity of a routing algorithm is nicely related to the distance distribution in a graph and in particular to the mean distance. For $\tau > 3$, the graph distance (or hopcount) H_N between node 1 and node 2 in the graph, conditionally on the event that these two nodes are connected, was shown heuristically in [32, (54)] and [5, Section V.C, (63) and (64)] to satisfy approximately

$$\mathbb{E}[H_N | H_N < \infty] \approx \frac{\log(N/\mu)}{\log \nu} + 1. \tag{4}$$

Concerning the asymptotic distribution of H_N, we proved together with R. van der Hofstad the following statement adapted from [22, Reference (1.9)].

Theorem 1.1 *Assume that $\tau > 3$ and that the expected value ν defined in (3) satisfies $\nu > 1$. For $n \geq 1$ define $a_n = \lfloor \log_\nu n \rfloor - \log_\nu n \in (-1, 0]$. Then there exist random variables $\left\{\hat{R}_a,\, a \in (-1,0]\right\}$ such that for each $\alpha > 0$, as $N \to \infty$,*

$$\begin{aligned} &\mathbb{P}\left[H_N - \lfloor \log_\nu N \rfloor > k\right] \\ &\qquad = \mathbb{P}\left[\hat{R}_{a_N} > k\right] + O((\log N)^{-\alpha}), \quad k \in \mathbb{Z}. \end{aligned} \tag{5}$$

Furthermore, if $\mathcal{W}^{(1)}$ and $\mathcal{W}^{(2)}$ are two independent copies of the limit random variable $\mathcal{W}$ introduced below, then for fixed $a \in (-1, 0]$, and $k \in \mathbb{Z}$,

$$\mathbb{P}\left[\hat{R}_a > k\right] = \mathbb{E}\left[e^{-\mu(\nu-1)^{-1}\nu^{a+k}\mathcal{W}^{(1)}\mathcal{W}^{(2)}}\right]. \tag{6}$$

The random variable $\mathcal{W}$ is the almost sure limit of the normalized (delayed) branching process

$$\mathcal{W}_n = \mathcal{Z}_n/\mu\nu^{n-1}, \tag{7}$$

where $\mathcal{Z}_0 = 1$, having offspring distribution in the first generation given by (1), whereas in the second and in all further generations the offspring distribution is given by

$$g_j = \frac{(j+1)f_{j+1}}{\mu}, \quad j = 0, 1, \ldots. \tag{8}$$

The mean value of the probability mass function $\{g_j\}$ is denoted by $\nu = \sum j g_j$ and satisfies (3). The involvement of $\{g_j\}$ becomes transparent if one realizes that the stubs are matched according to size. Since the number of stubs at each node follows distribution $\{f_j\}$, sampling in accordance with size means sampling proportional to jf_j. The approximating branching process $\{\mathcal{Z}_n\}$ that describes the expansion (the number of nodes on distance n from a given node) uses *one* stub for attachment and, hence, the offspring of this branching process has size j with probability proportional to $(j+1)f_{j+1}$, see also [32, (Section II.a, Reference (9)].

Theorem 1.1 suggests that the asymptotic mean distance satisfies

$$\mathbb{E}[H_N | H_N < \infty] = \lfloor \log_\nu N \rfloor + \mathbb{E}[\hat{R}_{a_N} | \hat{R}_{a_N} < \infty] + o(1), \tag{9}$$

and that the asymptotic variance satisfies

$$\mathrm{Var}\left[H_N | H_N < \infty\right] = \mathrm{Var}\left[\hat{R}_{a_N} | \hat{R}_{a_N} < \infty\right] + o(1), \tag{10}$$

as $N \to \infty$.

Note that the random variables $\{R_a,\, a \in (-1, 0]\}$, defined in [22, (1.8)], as

$$\mathbb{P}[R_a > k] = \mathbb{E}\left[e^{-\mu(\nu-1)^{-1}\nu^{a+k}\mathcal{W}^{(1)}\mathcal{W}^{(2)}} | \mathcal{W}^{(1)}\mathcal{W}^{(2)} > 0\right], \tag{11}$$

are connected to the (possibly defective) random variables $\{\hat{R}_a,\, a \in (-1,0]\}$ defined in (6), through

$$R_a = \hat{R}_a | \hat{R}_a < \infty, \tag{12}$$

because $\{\hat{R}_a < \infty\} = \{\mathcal{W}^{(1)}\mathcal{W}^{(2)} > 0\}$.

We will not prove (9) or (10) rigorously, but we give the following motivation. Observe that (9) (as the result (10)), can only hold for two nodes that are connected, as is indicated by the conditioning on the event $\{H_N < \infty\}$. However, the error term in (5), does not hold for the event $\{H_N < \infty\}$. Instead, in statements involving the event $\{H_N < \infty\}$ the error term, in [22], was replaced by a $o(1)$ term, as $N \to \infty$. Even if we demand that $\mathbb{P}(D \geq 2) = 1$, in which case $\mathcal{W}^{(1)}\mathcal{W}^{(2)} > 0$, with probability 1 and hence the distribution of H_N has for $N \to \infty$ no mass at $+\infty$, the error term in (5) does not suffice to obtain (9). However, when $\mathbb{P}(D \geq 2) = 1$ and defining $H_N^* = \min\{H_N, \alpha_N\}$, where $\alpha_N = C \log_\nu N$, for some multiplication factor $C > 1$, we obtain for $N \to \infty$,

$$\begin{aligned}
\mathbb{E}[H_N^*] &= \sum_{n=1}^{\alpha_N} n\mathbb{P}[H_N = n] \\
&= \sum_{n=-\lfloor \log_\nu N \rfloor}^{\alpha_N - \lfloor \log_\nu N \rfloor} (\lfloor \log_\nu N \rfloor + k)\mathbb{P}[H_N - \lfloor \log_\nu N \rfloor = k] \\
&= \lfloor \log_\nu N \rfloor \sum_{n=-\lfloor \log_\nu N \rfloor}^{\alpha_N - \lfloor \log_\nu N \rfloor} \mathbb{P}[R_{a_N} = k] + \sum_{n=-\lfloor \log_\nu N \rfloor}^{\alpha_N - \lfloor \log_\nu N \rfloor} k \left(\mathbb{P}[R_{a_N} = k] + O((\log N)^{-\alpha}) \right) \\
&= \lfloor \log_\nu N \rfloor + \mathbb{E}[R_{a_N}] + O((\log N)^{1-\alpha}),
\end{aligned}$$

as can be obtained from (5) and (11). The choice of H_N^* is motivated by the centering constant $\lfloor \log_\nu N \rfloor$ of H_N, and it is expected that H_N will not exceed a large multiple of $\log_\nu N$. We can not prove (9) rigorously, with the tools of [22], instead we are satisfied by the simulations in Section 5, which convincingly show that the approximation following from (9) is superior compared to the crude approximation (4).

Combining (9) and the intermediate result (14), formulated and proved in Theorem 2.1 below, yields the main result of this paper,

$$\mathbb{E}[H_N | H_N < \infty] \approx \frac{\log N}{\log \nu} + \frac{1}{2} - \left(\frac{\gamma + \log \mu - \log(\nu - 1)}{\log \nu} \right) - 2 \frac{\mathbb{E}[\log \mathcal{W} | \mathcal{W} > 0]}{\log \nu}, \tag{13}$$

where γ denotes Euler's constant ([1, 6.1.3]).

1.2 Related work

There is a wealth of related work which we will now summarize. The model investigated here was also studied in [33], with $1 - F(x) = x^{-\tau+1}L(x)$, where $\tau \in (2,3)$ and L denotes a slowly varying function. It was shown in [33] that the average distance is bounded from above by $2\frac{\log\log N}{|\log(\tau-2)|}(1 + o(1))$. The

asymptotics of the distances and the connected component sizes for $\tau < 3$ is also treated in three publications [23, 24, 25].

There is substantial work on random graphs that are, although different from ours, still similar in spirit. In [2], random graphs were considered with a degree sequence that is *precisely* equal to a power law, meaning that the number of nodes with degree k is precisely proportional to $k^{-\tau}$. Aiello *et al.* [2] show that the largest connected component is of the order of the size of the graph when $\tau < \tau_0 = 3.47875\ldots$, where τ_0 is the solution of $\zeta(\tau-2) - 2\zeta(\tau-1) = 0$, and where ζ is the Riemann Zeta function. When $\tau > \tau_0$, the largest connected component is of smaller order than the size of the graph and more precise bounds are given for the largest connected component. When $\tau \in (1,2)$, the graph is with high probability connected. The proofs of these facts use couplings with branching processes and strengthen previous results due to Molloy and Reed [29, 30]. For this same model, Dorogovtsev *et al.* [17, 18] investigate the leading asymptotics and the fluctuations around the mean of the distance between arbitrary nodes in the graph from a theoretical physics point of view, using mainly generating functions.

A second related model can be found in [14] and [15], where edges between nodes i and j are present with probability equal to $w_i w_j / \sum_l w_l$ for some 'expected degree vector' $w = (w_1, \ldots, w_N)$. Chung and Lu [14] show that when w_i is proportional to $i^{-\frac{1}{\tau-1}}$ the mean distance between pairs of nodes is $\log_\nu N(1+o(1))$ when $\tau > 3$, and $2\frac{\log\log N}{|\log(\tau-2)|}(1+o(1))$ when $\tau \in (2,3)$. The difference between this model and ours is that the nodes are not exchangeable in [14], but the observed phenomena are similar. This result can be understood heuristically as follows. Firstly, the actual degree vector in [14] should be close to the expected degree vector. Secondly, for the expected degree vector, we can compute that the number of nodes for which the degree is at least k equals

$$|\{i : w_i \geq k\}| = |\{i : ci^{-\frac{1}{\tau-1}} \geq k\}| \propto k^{-\tau+1}.$$

Thus, one expects that the number of nodes with degree at least k decreases as $k^{-\tau+1}$, similarly as in our model. In [15], Chung and Lu study the sizes of the connected components in the above model. The advantage of their model is that the edges are *independently* present, which makes the resulting graph closer to a classical random graph.

All the models described above are *static*, i.e., the size of the graph is *fixed*, and we have not modelled the *growth* of the graph. There is a large body of work investigating *dynamical* models for complex networks, often in the context of the World-Wide Web. In various forms, preferential attachment has been shown to lead to power law degree sequences. Therefore, such models intend to *explain* the occurrence of power law degree sequences in random graphs. See [3, 4, 5, 9, 10, 11, 12, 13, 16, 27] and the references therein. In the preferential attachment model, nodes and/or edges are added sequentially. The edges are attached to nodes with a probability proportionally to the degree of the receiving node, thus favoring nodes with large degrees. For these growth models, it is shown

that the number of nodes with degree k decays according to a power law. In some papers this is done heuristically, using simulations, in some others models a rigorous mathematical proof has been given. See also [11] for a survey.

It can be expected that our model is a snapshot of the above models, i.e., a realization of the graph growth processes at the time instant that the graph has a certain prescribed size. Thus, rather than to describe the growth of the model, we investigate the properties of the model at a given time instant. This is suggested in [5, Section VII.D], and it would be very interesting indeed to investigate this further mathematically, i.e., to investigate the relation between the configuration and the preferential attachment models.

The reason why we study the random graphs at a given time instant is that we are interested in the topology of the Internet. In [34], and inspired by the observed power law degree sequence in [20], the configuration model with i.i.d. degrees is proposed as a model for the AS-graph in Internet, and it is argued on a qualitative basis that this simple model serves as a better model for the Internet topology than currently used topology generators.

In [31, Table II], many more examples are given of real networks that have power law degree sequences. Interestingly, there are also many examples where power laws are *not* observed, and often the degree law falls off faster than a power law. These observed degrees can be described by a degree distribution as in (1) with $1 - F(x)$ smaller than any power, and the results in this paper thus apply. Such examples are described in more detail in [5, Section II]. Examples where the tails of the degree distribution are lighter than power laws are power and neural networks [5, Section II.K], where the tails are observed to be exponential, and protein folding [5, Section II.L], where the tails are observed to be Gaussian. In other examples, a degree distribution is found that for small values is a power law, but has an exponential cut off. An example of such a degree distribution is

$$f_k = Ck^{-\delta}e^{-k/\kappa},$$

for some $\kappa > 0$ and $\delta \in \mathbb{R}$. The size of κ indicates up to what degree the power law still holds, and where the exponential cut off starts to set in. For this example, our results apply since the exponential tail ensures that (2) holds for *any* $\tau > 3$ by picking $c > 0$ large enough.

1.3 Organization of the paper

The paper is organized as follows. In Section 2 we compute $\mathbb{E}[R_a]$ within very tight bounds. Section 3 discusses the computation of $\mathbb{E}[\log \mathcal{W} | \mathcal{W} > 0]$. Apart from a numerical way to obtain an approximation to $\mathbb{E}[\log \mathcal{W} | \mathcal{W} > 0]$, we discuss upper and lower bounds for this quantity. In Section 4 we discuss the two types of random graphs for which we compare the mean hopcount. Section 5 compares the value of (13) with simulations of the mean hopcount in the mentioned random graphs. The important

conclusion drawn in Section 6 from the material presented in Section 4 and Section 5 is that the expression (13) for the mean hopcount between two arbitrary nodes in the configuration model is surprisingly accurate, even for a relatively small number of nodes N around $N = 500$. Mathematical and numerical details on the computation of $\mathbb{E}[\log \mathcal{W}|\mathcal{W} > 0]$ are deferred to Appendices A, B and C.

2 The first and second moment of R_a

In this section we give tight bounds for the first moment $\mathbb{E}[R_a]$ and present a parallel result for $\mathrm{Var}[R_a]$. Since we actually include an expression for the generating function $\varphi_{R_a}(z) = \mathbb{E}[z^{R_a}]$, extensions to higher moments can also be obtained. For the statement of the theorem we introduce the non-negative function

$$T(\nu) = \frac{2}{\sqrt{\log \nu}} \sum_{k=1}^{\infty} \frac{1}{\sqrt{2k \sinh \frac{2k\pi^2}{\log \nu}}}, \qquad \nu > 1.$$

The graph of this function is given in Figure 1. The function $T(\nu)$ is increasing and for $1 < \nu \leq 5$, its maximum value $T(5)$ is smaller than 0.0035.

Theorem 2.1 *For $\nu > 1$,*

$$\left| \mathbb{E}[R_a] - \left[\frac{1}{2} - a - \left(\frac{\gamma + \log \mu - \log(\nu - 1)}{\log \nu} \right) - 2 \frac{\mathbb{E}[\log \mathcal{W}|\mathcal{W} > 0]}{\log \nu} \right] \right| \leq T(\nu), \tag{14}$$

where $\mathcal{W}$ is the limit of the delayed branching process introduced in (7), the parameters μ and ν are defined in (3) and where γ denotes Euler's constant ([1, 6.1.3]).

Proof: Denote by φ_{R_a} the generating function of the random variable R_a, introduced in (12). Then

$$\varphi_{R_a}(z) = \sum_{k=-\infty}^{\infty} z^k \mathbb{P}[R_A = k] = (z-1) \sum_{k=-\infty}^{\infty} z^k \mathbb{P}[R_A > k].$$

Let $a \in (-1, 0]$ be fixed. Denote by X a random variable having the same distribution as the random variable

$$\mu(\nu - 1)^{-1} \nu^a \mathcal{W}^{(1)} \mathcal{W}^{(2)} | \mathcal{W}^{(1)} \mathcal{W}^{(2)} > 0,$$

then from (11), using Fubini's theorem [28],

$$\varphi_{R_a}(z) = (z-1)\, \mathbb{E}\left[\sum_{k=-\infty}^{\infty} z^k \exp\{-\nu^k X\} \right].$$

Introducing $G(x, z) = G_1(x, z) + G_2(x, z)$, where

$$G_1(x, z) = \sum_{k=0}^{\infty} z^k e^{-x\nu^k},$$

$$G_2(x, z) = \sum_{k=1}^{\infty} z^{-k} e^{-x\nu^{-k}} = \sum_{j=0}^{\infty} \frac{(-x)^j}{j!} \frac{1}{z\nu^j - 1},$$

yields

$$\varphi_{R_a}(z) = (z-1)\,\mathbb{E}[G(X,z)] = (z-1)\,\{\mathbb{E}[G_1(X,z) + G_2(X,z)]\}. \tag{15}$$

Let us first consider $G_1(x,z)$. For $\sigma = Re(s) > 0$ and $k \geq 0$, we have [1, 6.1.1],

$$\frac{\Gamma(s)}{\nu^{ks}} = \int_0^\infty t^{s-1}\, e^{-\nu^k t}\, dt,$$

and

$$\Gamma(s) \sum_{k=0}^{\infty} \frac{z^k}{\nu^{ks}} = \int_0^\infty t^{s-1} \sum_{k=0}^{\infty} z^k e^{-\nu^k t}\, dt,$$

or, assuming $|z/\nu^s| < 1$,

$$\frac{\Gamma(s)\nu^s}{\nu^s - z} = \int_0^\infty t^{s-1} G_1(t,z)\, dt.$$

By the inverse Mellin transform, an alternative expression of $G_1(x,z)$ is, with $c > 0$,

$$G_1(x,z) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{\Gamma(s)}{\nu^s - z} \left(\frac{\nu}{x}\right)^s ds.$$

The integrand

$$g(s) = \frac{\Gamma(s)}{\nu^s - z} \left(\frac{\nu}{x}\right)^s,$$

is analytic in the entire complex plane ($s \in \mathbb{C}$), except for the points $s = -k$ ($k \geq 0$), where we encounter simple poles of $\Gamma(s)$ and the points $s = \frac{\log z + 2k\pi i}{\log \nu}$ ($k \in \mathbb{Z}$), which are the simple poles of $\frac{1}{\nu^s - z}$. Consider the contour C_R, consisting of that part of the circle $|s-c| = R$ with real part smaller than or equal to c, and the vertical line segment $[c - ib, c + ib]$, connecting the points of intersection of the line $Re(s) = c$ and the circle. Then, with $\Gamma_R = \{s : |s-c| = R, Re(s) \leq c\}$

$$\int_{C_R} g(s)\, ds = \int_{c-ib}^{c+ib} g(s)\, ds + \int_{\Gamma_R} g(s) ds.$$

Since

$$\begin{aligned}\left|\int_{\Gamma_R} g(s) ds\right| &= \left| iR \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} g(R\, e^{i\theta}) e^{i\theta} d\theta \right| \\ &\leq R \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} \left| g(R\, e^{i\theta}) \right| d\theta,\end{aligned}$$

and, since from [1, (6.1.17), (6.1.26)], we have

$$\lim_{R\to\infty} R \left| g(R\, e^{i\theta}) \right| = 0,$$

uniformly for $\theta \in \left[\frac{\pi}{2}, \frac{3\pi}{2}\right]$, this results in

$$\lim_{R\to\infty} \int_{C_R} g(s)\, ds = \int_{c-i\infty}^{c+i\infty} g(s)\, ds.$$

Hence, by Cauchy's residue theorem,

$$\int_{c-i\infty}^{c+i\infty} g(s)\, ds = 2\pi i \sum_k \mathrm{Res}_{s=s_k}\{g(s)\}$$

where the sum is over all simple poles in the left half plane $Re(s) \leq 0$, so that

$$G_1(x,z) = \sum_{k=0}^{\infty} \frac{1}{k!} \frac{(-x)^k}{1-z\nu^k} + \frac{1}{z\log\nu} \sum_{k=-\infty}^{\infty} \Gamma\left(\frac{\log z + 2k\pi i}{\log\nu}\right) \left(\frac{\nu}{x}\right)^{\frac{\log z + 2k\pi i}{\log\nu}}$$
$$= -G_2(x,z) + \frac{z^{-\log_\nu x}}{\log\nu} \sum_{k=-\infty}^{\infty} \Gamma\left(\frac{\log z + 2k\pi i}{\log\nu}\right) e^{-2k\pi i \log_\nu x}.$$

This implies the following representation for $G(x,z)$,

$$G(x,z) = \frac{z^{-\log_\nu x}}{\log\nu} \sum_{k=-\infty}^{\infty} \Gamma\left(\frac{\log z + 2k\pi i}{\log\nu}\right) e^{-2k\pi i \log_\nu x}.$$

Hence, from (15),

$$\varphi_{R_a}(z) = \frac{(z-1)}{\log\nu} \mathbb{E}\left[\sum_{k=-\infty}^{\infty} \Gamma\left(\frac{\log z + 2k\pi i}{\log\nu}\right) e^{-(\log z + 2k\pi i)\log_\nu X}\right],$$

which shows the relation between the random variables R_a and $\log_\nu X$.

It is convenient to treat the term with index $k=0$ separately,

$$\varphi_{R_a}(z) = \mathbb{E}\left[\frac{z^{-\log_\nu X}(z-1)}{\log\nu}\Gamma\left(\frac{\log z}{\log\nu}\right)\right] + \frac{(z-1)}{\log\nu}\sum_{k\neq 0}\Gamma\left(\frac{\log z + 2k\pi i}{\log\nu}\right)$$
$$\times\, \mathbb{E}\left[e^{-(\log z + 2k\pi i)\log_\nu X}\right]. \tag{16}$$

We start by investigating the term associated with the first term in (16),

$$\frac{z^{-\log_\nu X}(z-1)}{\log\nu}\Gamma\left(\frac{\log z}{\log\nu}\right) = \frac{(1-e^{-\log z})}{\log z}\Gamma\left(\frac{\log z}{\log\nu}+1\right) e^{(1-\log_\nu X)\log z}$$
$$= \frac{(e^{(1-\log_\nu X)\log z} - e^{-\log_\nu X \log z})}{\log z}\Gamma\left(\frac{\log z}{\log\nu}+1\right).$$

Expanding in a series yields

$$\frac{z^{-\log_\nu X}(z-1)}{\log\nu}\Gamma\left(\frac{\log z}{\log\nu}\right) = \sum_{k=0}^{\infty} \frac{(1-\log_\nu X)^{k+1} - (-\log_\nu X)^{k+1}}{(k+1)!}\log^k z$$
$$\times \sum_{m=0}^{\infty} \frac{\Gamma^{(m)}(1)}{m!\log^m\nu}\log^m z.$$

From this we find explicitly

$$\frac{d}{dz}\frac{z^{-\log_\nu X}(z-1)}{\log\nu}\Gamma\left(\frac{\log z}{\log\nu}\right)\Big|_{z=1} = \frac{1}{2} - \frac{\gamma}{\log\nu} - \frac{\log X}{\log\nu},$$

where as before $\gamma = .5772...$ denotes Euler's constant. Including the second term yields

$$\mathbb{E}\left[R_a\right] = \frac{1}{2} - \frac{\gamma}{\log\nu} - \mathbb{E}\left[\frac{\log X}{\log\nu}\right] + 2\mathbb{E}\left[\sum_{k=1}^{\infty} \frac{\cos\left(\frac{2k\pi}{\log\nu}\log\left(\frac{\nu}{X}\right) + \arg\Gamma\left(\frac{2k\pi i}{\log\nu}\right)\right)}{\sqrt{\log\nu}\sqrt{2k\sinh\frac{2k\pi^2}{\log\nu}}}\right].$$

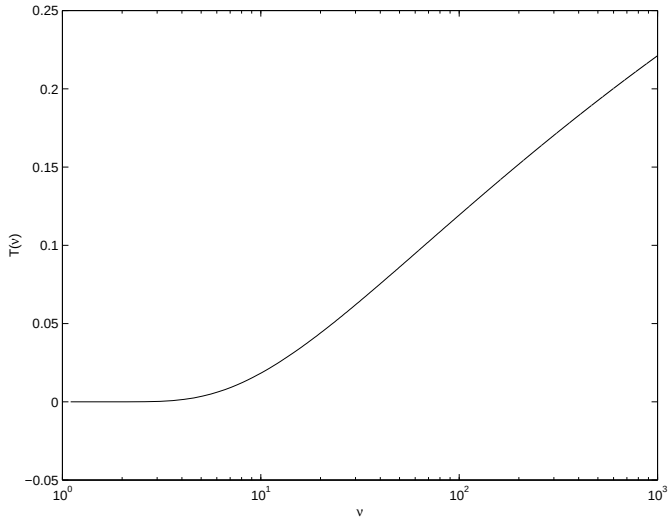

Figure 1: The function $T(\nu)$.

For $x > 0$, the rapidly converging series

$$\frac{2}{\sqrt{\log \nu}} \sum_{k=1}^{\infty} \frac{|\cos\left[\frac{2k\pi}{\log \nu} \log\left(\frac{\nu}{x}\right) + \arg\Gamma\left(\frac{2k\pi i}{\log \nu}\right)\right]|}{\sqrt{2k \sinh \frac{2k\pi^2}{\log \nu}}} \leq \frac{2}{\sqrt{\log \nu}} \sum_{k=1}^{\infty} \frac{1}{\sqrt{2k \sinh \frac{2k\pi^2}{\log \nu}}} = T(\nu),$$

increases with ν. From the above inequality we obtain

$$\frac{1}{2} - \frac{\gamma}{\log \nu} - \mathbb{E}\left[\frac{\log X}{\log \nu}\right] - T(\nu) \leq \mathbb{E}\left[R_a\right] \leq \frac{1}{2} - \frac{\gamma}{\log \nu} - \mathbb{E}\left[\frac{\log X}{\log \nu}\right] + T(\nu).$$

Since $X = \mu(\nu-1)^{-1}\nu^a \mathcal{W}^{(1)}\mathcal{W}^{(2)} | \mathcal{W}^{(1)}\mathcal{W}^{(2)} > 0$, and $\mathcal{W}^{(i)}$ for $i = 1, 2$ are independent, we end up with the statement of the theorem.

□

Differentiating twice with respect to z, and omitting the tedious calculation which runs parallel to the case above, we obtain

$$\mathrm{Var}\left(R_a\right) \approx \frac{1}{12} + \frac{\pi^2}{6 \log^2 \nu} + 2\mathrm{Var}\left[\log_\nu \mathcal{W} | \mathcal{W} > 0\right].$$

3 Computation of $\mathbb{E}[\log \mathcal{W} | \mathcal{W} > 0]$

3.1 Numerical approach

Formula (13) requires the expected value $\mathbb{E}[\log \mathcal{W} | \mathcal{W} > 0]$. Unfortunately, little is known about the distribution of the limit random variable $\mathcal{W}$. The functional equation that determines the generating function of $\mathcal{W}$ seems difficult to solve in general. For the classical super-critical Galton-Watson branching process (our delayed branching process is a simple extension of this model) the distribution of $\mathcal{W}$ is only explicitly known for geometrically distributed offspring (see [21, p. 92]). In this case $\mathcal{W}$

possesses an atom at 0 and conditionally on the event $\{\mathcal{W} > 0\}$, the distribution of $\mathcal{W}$ is exponential. Appendix C solves the functional equation numerically for Poisson distributed offspring.

Here we use the Fast Fourier Transform (FFT) to approximate the distribution of $\mathcal{W}_n$ and from that distribution the value of $\mathbb{E}[\log \mathcal{W}_n | \mathcal{W}_n > 0]$. This approximation is based on the fact that the generating function of the size of generation n in a branching process is the functional iterate of the generating function of the offspring distribution. We show in Appendix A that indeed $\mathbb{E}[\log \mathcal{W}_n | \mathcal{W}_n > 0]$ converges to $\mathbb{E}[\log \mathcal{W} | \mathcal{W} > 0]$. However, conclusions about the involved error in this approximation are beyond the scope of this paper. We therefore extrapolate our results numerically.

Since in a super-critical Galton-Watson process the size of generation n grows exponentially with n, the memory in current computers confines the accuracy of the computation of $\mathbb{E}[\log \mathcal{W} | \mathcal{W} > 0]$, especially if the mean of the offspring distribution is substantially greater than 1. We expect that this drawback will be less pronounced in future implementations of the FFT. The exact expression of $\mathbb{E}[\log \mathcal{W} | \mathcal{W} > 0]$ for geometric offspring is used, in Appendix B, as a benchmark to evaluate the accuracy of the FFT.

3.2 Bounds

In a simple Galton-Watson branching process the extinction probability π_0 is the smallest non-negative solution of the equation:

$$s = p_0 + p_1 s + p_2 s^2 + \ldots, \tag{17}$$

where $\{p_k\}$ denotes the offspring distribution (see [21]). Given $\{p_k\}$ this solution can easily be obtained numerically. In case of a delayed branching process with offspring $\{f_k\}$ in the first generation and offspring distribution $\{g_k\}$ given in (8) for all further generations, we find conditionally on $\mathcal{Z}_1 = j$ that:

$$\mathbb{P}(\mathcal{W} = 0 | \mathcal{Z}_1 = j) = \lim_{n \to \infty} \mathbb{P}(\mathcal{Z}_n = 0 | \mathcal{Z}_1 = j) = \pi_0^j,$$

where π_0 is the smallest non-negative solution of (17), with the $\{p_k\}$ replaced by $\{g_k\}$. Hence the extinction probability of the delayed branching process $\mathcal{Z}_k$ equals:

$$\pi = \mathbb{P}(\mathcal{W} = 0) = \sum_{j=1}^{\infty} (\pi_0)^j f_j, \tag{18}$$

where as before π_0 is the smallest non-negative solution of (17), with the $\{p_k\}$ replaced by $\{g_k\}$, and $\{f_k\}$ is the discrete distribution defined in (1).

Since $\mathbb{E}[\mathcal{W} | \mathcal{W} > 0] = \mathbb{E}[\mathcal{W}] / (1 - \pi)$, Jensen's inequality for concave functions provides us with an upperbound,

$$\mathbb{E}[\log \mathcal{W} | \mathcal{W} > 0] \leq \log(\mathbb{E}[\mathcal{W} | \mathcal{W} > 0]) = \log(\mathbb{E}[\mathcal{W}]) - \log(1 - \pi).$$

Using $\mathbb{E}[\mathcal{W}] = 1$, the Jensen's inequality becomes

$$\mathbb{E}[\log \mathcal{W} | \mathcal{W} > 0] \leq -\log(1-\pi). \tag{19}$$

From (19), we obtain a lower bound for the mean distance (or hopcount)

$$\mathbb{E}[H_N | H_N < \infty] \geq \log_\nu N + \frac{1}{2} - \left(\frac{\gamma + \log \mu - \log(\nu - 1)}{\log \nu} \right) + \frac{2 \log(1-\pi)}{\log \nu}. \tag{20}$$

An obvious lower bound for $\mathbb{E}[\log \mathcal{W} | \mathcal{W} > 0]$, derived in Appendix B, is

$$\mathbb{E}[\log \mathcal{W} | \mathcal{W} > 0] \geq -\int_0^1 \frac{\mathbb{P}(0 < \mathcal{W} < y)}{1-\pi} \frac{dy}{y}.$$

Since the distribution of $\mathcal{W}$ is unknown, this lower bound is not very useful. It would be of interest to have an attainable lower bound for $\mathbb{E}[\log \mathcal{W} | \mathcal{W} > 0]$; in turn this would imply an upper bound for the mean distance.

In case $\pi = 0$, conditioning on $\mathcal{W} > 0$, boils down to conditioning on the sure event and the Jensen inequality gives:

$$\mathbb{E}[\log \mathcal{W} | \mathcal{W} > 0] \leq \log(\mathbb{E}[\mathcal{W}]) = 0.$$

Since in this case the sequence $\{\log \mathcal{W}_n | \mathcal{W}_n > 0\} = \{\log \mathcal{W}_n\}$ is a supermartingale, because $\mathcal{W}_n$ is a martingale and the function $x \mapsto \log x$ is concave, we conclude that for $\pi = 0$, the sequence $\mathbb{E}[\log \mathcal{W}_n | \mathcal{W}_n > 0]$ is monotone decreasing. This result is of interest to compute $\mathbb{E}[\log \mathcal{W} | \mathcal{W} > 0]$ via the Fourier method, for the distribution (21) considered below.

4 Examples of two types of degree graphs

We performed the numerical calculation of $\mathbb{E}[\log \mathcal{W} | \mathcal{W} > 0]$ for the delayed branching process with a power law offspring (specified by (21) below) and for the binomial offspring distribution.

First consider the distribution function

$$F(k) = \mathbb{P}(D \leq k) = 1 - \left(\frac{1}{k} \right)^{\tau - 1}, \quad k = 1, 2, 3, \ldots, \tag{21}$$

with $\tau > 3$. Observe that $F(1) = 0$ and consequently that $\mathbb{P}(D \geq 2) = 1$. This implies that the delayed branching process has extinction probability $\pi = 0$. Introducing Riemann's zeta function $\zeta(s) = \sum_{j=1}^{\infty} j^{-s}$, $s > 1$, we find

$$\mu = \mathbb{E}[D] = 1 + \zeta(\tau - 1) \quad \text{and} \quad \nu = \frac{2\zeta(\tau - 2)}{1 + \zeta(\tau - 1)}. \tag{22}$$

In Table 1 we take $\tau = 4.5$ and find

$$\mu = 2.1267\ldots \qquad \nu = 1.2615\ldots.$$

Table 1: Results for $\tau = 4.5$

n	$\mu\nu^{n-1}$	FFT(approx)	$\mathbb{E}[\log \mathcal{W}_n \vert \mathcal{W}_n > 0]$
13	34.5570	34.5433	-0.5334
14	43.5952	43.5697	-0.5522
15	54.9972	54.9498	-0.5686
16	69.3813	69.2935	-0.5828

We only display the results in the last 4 generations, where the difference between numerical and theoretical mean is less than 0.1, which is our rule of thumb as explained in Appendix B. Using a numerical extrapolation we will adopt the value $\mathbb{E}[\log \mathcal{W} | \mathcal{W} > 0] \approx -0.6$, as the correct value for $\tau = 4.5$ and will compute in the next section our estimate of $\mathbb{E}[H_N | H_N < \infty]$. Similarly, for $\tau = 5.5$, we have that

$$\mu = 2.0547\ldots \qquad \nu = 1.0967\ldots,$$

and the numerical results for the last 4 generations are given in Table 2. Extrapolating we adopt the value $\mathbb{E}[\log \mathcal{W} | \mathcal{W} > 0] \approx -0.45$.

Table 2: Results for $\tau = 5.5$

n	$\mu\nu^{n-1}$	FFT(approx)	$\mathbb{E}[\log \mathcal{W}_n \vert \mathcal{W}_n > 0]$
43	99.3130	99.3086	-0.4263
44	108.9199	108.9138	-0.4283
45	119.4562	119.4475	-0.4301
46	131.0117	130.9993	-0.4319

Next, we consider the binomial probabilities $f_j = \binom{N-1}{j} p^j (1-p)^{N-1-j}$, with mean $\mu = (N-1)p$, for which

$$g_j = \frac{(j+1) f_{j+1}}{Np} = \binom{N-2}{j} p^j (1-p)^{N-2-j},$$

so that $\{g_j\}$ has a binomial distribution with parameters $N-2$ and p. Hence $\nu = (N-2)p$. It is well known that for $N \to \infty$ and $p_N \to 0$, such that $Np_N \to \nu$, the binomial distribution with parameters N and p_N is excellently approximated by the Poisson distribution with mean ν. Therefore we used the Poisson distribution with mean ν as offspring distribution in *all* generations of the branching process. Exact series for the probability density and generating function for the limit random variable $\mathcal{W}$ of a branching process with Poisson offspring distribution are presented in [35]. The last 4 relevant generations for the Poisson distribution with mean $\nu = 1.5$, are given in Table 3.

Table 3: Results for the Poisson with $\nu = 1.5$

n	ν^n	FFT(approx)	$\mathbb{E}[\log \mathcal{W}_n \mid \mathcal{W}_n > 0]$
11	86.4976	86.4976	0.1048
12	129.7463	129.7463	0.1016
13	194.6195	194.6194	0.0992
14	291.9293	291.8817	0.0973

We approximate:

$$\mathbb{E}[\log \mathcal{W} | \mathcal{W} > 0] \approx 0.09.$$

For the binomial distribution, the degree distribution $\{f_j\}$ coincides with the degree distribution of the random graph $G_p(N)$; both distributions are binomial with parameters $N-1$ and p. In contrast to the random graph $G_p(N)$, the degree graph can have self-loops by construction. However, the hopcount in both graphs (the degree graph with binomial degrees with parameters $N-1$ and p, and the random graph $G_p(N)$) are asymptotically equal in distribution as follows from the coupling construction in [22]. This implies that our results for the mean graph distance (hopcount) also hold for the random graph $G_p(N)$.

5 Comparison of different formulas for $\mathbb{E}[H_N | H_N < \infty]$

5.1 Power law distributions

For several values of N we compare a large sample confidence interval of the simulated mean hopcount based on 1000 repetitions with the results (4) and (13).

For the power-law distribution (21) with $\tau = 4.5$, $\mu = 2.1267$, $\nu = 1.2615$ and $\mathbb{E}[\log \mathcal{W} | \mathcal{W} > 0] \approx -0.6$, we present the results in Table 4.

Table 4: Comparison for $\tau = 4.5$

N	simulation	$\mathbb{E}[H_N H_N < \infty](4)$	$\mathbb{E}[H_N \mid H_N < \infty](13)$
500	(20.10, 21.34)	24.5	20.9
1000	(23.05, 24.22)	27.5	23.9
5000	(30.78, 32.01)	34.4	30.8
10,000	(33.51, 34.70)	37.4	33.8

For $\tau = 5.5$, $\mu = 2.0547$, $\nu = 1.0967$ and $\mathbb{E}[\log \mathcal{W} | \mathcal{W} > 0] \approx -0.45$, we present our findings in

Table 5.

Table 5: Comparison for $\tau = 5.5$

N	simulation	$\mathbb{E}[H_N H_N < \infty](4)$	$\mathbb{E}[H_N \vert H_N < \infty](13)$
500	(36.55, 39.19)	60.52	38.21
1000	(43.42, 45.96)	68.03	45.72
5000	(61.70, 64.32)	85.47	63.16
10,000	(68.98, 71.64)	92.98	70.66

In all cases the value (13) for $\mathbb{E}[H_N|H_N < \infty]$ is contained in the confidence interval obtained from the simulations, whereas the expression (4) overestimates the actual mean. These values demonstrate the accuracy of (13) for $\mathbb{E}[H_N H_N < \infty]$ in degree graphs with a finite variance degree distribution.

5.2 The Binomial (Poisson) distribution

For the degree graph with a binomial distribution with parameters $N-1$ and p, with $Np = 1.5$ for which $\mathbb{E}[\log \mathcal{W}|\mathcal{W} > 0] \approx 0.09$, we refer to Table 6.

Table 6: Comparison for binomial with $Np = 1.5$

N	simulation	$\mathbb{E}[H_N H_N < \infty](4)$	$\mathbb{E}[H_N \vert H_N < \infty](13)$
500	(10.35, 11.34)	15.3	11.2
1000	(12.25, 13.49)	17.0	12.9
5000	(16.10, 17.34)	21.0	16.9
10,000	(18.25, 19.51)	22.7	18.6

Figure 2 shows the meane hopcount as a function of ν. The simulated values for the random graph with $N = 500$ and average degree equal to $\nu = (N-1)p$ have error bars specified by one standard deviation. We observe that, although the number of nodes $N = 500$ is reasonably small, Formula (13) gives a nearly perfect match with the simulations. The lower bound (20) is relatively close to the simulated mean hopcount for larger values of the average degree and is superior to Newman's formula. As already demonstrated above in the tables, the Figure 2 also illustrates that for small values of ν, Newman's formula (4) is inadequate.

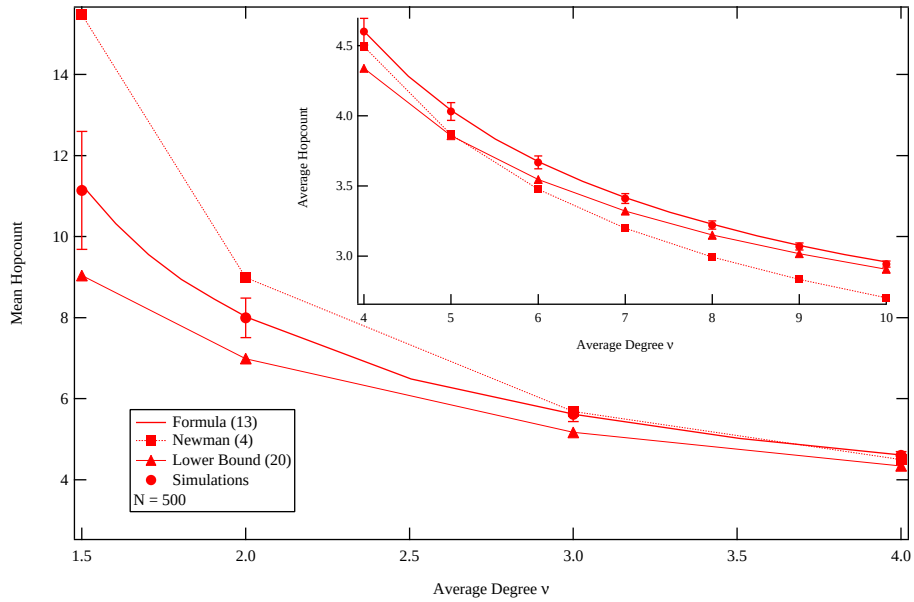

Figure 2: The mean hopcount for the degree graph with binomial degree distribution.

6 Conclusions

We conclude that (13) for $\mathbb{E}[H_N|H_N < \infty]$ gives a genuine improvement over Newman's formula (4). Inspection of both approximations shows that (4) overestimates the correct value of the mean $\mathbb{E}[H_N|H_N < \infty]$. The error in (4) is increasingly large when the Malthusian parameter ν is close to 1 because the factor $\log(N/\mu)$ is divided by $\log \nu$, which approaches zero as $\nu \to 1$. Apparently addition of the term

$$-\frac{1}{2} - \left(\frac{\gamma - \log(\nu - 1)}{\log \nu}\right) - 2\frac{\mathbb{E}[\log \mathcal{W}|\mathcal{W} > 0]}{\log \nu},$$

corrects the overshoot.

We have shown how the difficult quantity $\mathbb{E}[\log \mathcal{W}|\mathcal{W} > 0]$ can be computed and have presented in (20) a lower bound for $\mathbb{E}[H_N|H_N < \infty]$. Even for moderate values of the number of nodes N, Formula (13) for $\mathbb{E}[H_N|H_N < \infty]$ seems surprisingly accurate. For the binomial degree distribution the distance in the degree graph is asymptotically equal to the distance in the *random graph* $G_p(N)$. For this case we developed a method to compute numerically the expected value $\mathbb{E}[\mathcal{W}|\mathcal{W} > 0]$ for all values of $\nu > 1$ and demonstrated that Formula (13) is capable of computing the mean hopcount in the random graph even for values N is as low as 500, with remarkable precision.

Acknowledgment. We thank Remco van der Hofstad for valuable discussions concerning branching processes, and Ramin Hekmat for the simulation of the mean hopcount in the random graph (Figure 2). We also thank the referee for suggestions that improved the readability of this paper.

References

[1] M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions,* Dover 1968.

[2] W. Aiello, F. Chung and L. Lu. A random graph model for power law graphs. *Experiment. Math.* **10**, no. 1, 53–66, 2001.

[3] W. Aiello, F. Chung and L. Lu. Random evolution of massive graphs. In Handbook of Massive Data Sets, J. Abello, P.M. Pardalos and M.G.C. Resende, eds., Kluwer Academic, Dordrecht, 97–122, 2002.

[4] R. Albert and A.-L. Barabási. Emergence of Scaling in Random Networks. *Science* **286**: 509-512, 1999.

[5] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Rev. Mod. Phys.,* **74**, 47-97, 2002.

[6] S. Asmussen, Some martingale methods in the limit theory of supercritical branching processes, in *Branching Processes*, edited by A. Joffe and P. Ney, Advances in Applied Probability **5**, Marcel Dekker, New York 1978.

[7] A.-L. Barabasi. *Linked, the New Science of Networks.* Perseus Publishing, Cambridge, Massachusetts 2002.

[8] B. Bollobás. *Random Graphs* 2nd edition. Academic Press, 2001.

[9] B. Bollobás, C. Borgs, J.T. Chayes and O. Riordan. Directed scale-free graphs. Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms (Baltimore, MD, 2003), 132–139, ACM, New York, 2003.

[10] B. Bollobás and O. Riordan. The diameter of scale-free random graphs. Preprint Dept. of Math. Sciences, University of Memphis, 2003.

[11] B. Bollobás and O. Riordan. Mathematical results on scale-free random graphs. *Handbook of graphs and networks*, 1–34, Wiley-VCH, Weinheim, 2003.

[12] B. Bollobás and O. Riordan. Coupling scale-free and classical random graphs. Preprint, 2003.

[13] 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.

[14] F. Chung and L. Lu. The average distances in random graphs with given expected degrees, PNAS, **99**(25), 15879–15882, 2002.

[15] F. Chung and L. Lu. Connected components in random graphs with given expected degree sequences, *Annals of Combinatorics*, **6**, 125–145, 2002.

[16] C. Cooper and A.A. Frieze. A general model of web graphs. *Random Structures Algorithms*, **22** (3), 311–335, 2003.

[17] S.N. Dorogovtsev, A.V. Goltsev and J.F.F. Mendes. Pseudofractal scale-free web, *Phys. Rev. E* **65**, 066122, 2002.

[18] S.N. Dorogovtsev, J.F.F. Mendes and A.N. Samuhkin. Metric structure of random networks, *Nucl. Phys. B.* 653, 307, 2003.

[19] M. S. Dubuc, La densité de la loi-limite d'un processus en cascade expansif, Z. Wahrscheinlichkeitstheorie Verw. Gebiete **19**, 281-290, 1971.

[20] C. Faloutsos, P. Faloutsos and M. Faloutsos, On power-law relationships of the Internet topology, Proceedings of the ACM SIGCOMM, 1999.

[21] G.R. Grimmett and D.R. Stirzacker, *Probability and Random Processes,* Clarendon Press, Oxford 1982.

[22] R. van der Hofstad, G. Hooghiemstra and P. Van Mieghem, Random graphs with finite variance degrees, To appear in Random Graphs and Structures, 2004.

[23] R. van der Hofstad, G. Hooghiemstra and D. Znamenski. Distances in random graphs with finite mean and infinite variance degrees. Preprint 2004.

[24] R. van der Hofstad, G. Hooghiemstra and D. Znamenski. Distances in random graphs with infinite mean degrees. Submitted to Extremes.

[25] R. van der Hofstad, G. Hooghiemstra and D. Znamenski. Connected components in random graphs with i.i.d. degrees. In preparation.

[26] P. Jagers, *Branching Processes with Biological Applications,* John Wiley and Sons, London 1975.

[27] R. Kumar, P. Raghavan, S. Rajagopalan D. Sivakumar, A. Tomkins and E. Upfal. Stochastic Models for the Web Graph. *42st Annual IEEE Symposium on Foundations of Computer Science*, 57–65, 2000.

[28] M. Loève, *Probability Theory,* second edition, Van Nostrand, New York 1960.

[29] M. Molloy and B. Reed. A critical point for random graphs with a given degree sequence, *Random Structures and Algorithms*, **6**, 161-179, 1995.

[30] M. Molloy and B. Reed. The size of the giant component of a random graph with a given degree sequence, *Combin. Probab. Comput.*, **7**, 295-305, 1998.

[31] M.E.J. Newman. The structure and function of complex networks. *SIAM Rev.* **45**, no. 2, 167–256, 2003.

[32] M.E.J. Newman, S.H. Strogatz, and D.J. Watts. Random graphs with arbitrary degree distribution and their application, *Phys. Rev. E*, **64**, 026118, 1-17, 2001.

[33] H. Reittu and I. Norros. On the power law random graph model of massive data networks, *Performance Evalution*, **55** (1-2), 3-23, 2004.

[34] H Tangmunarunkit, R. Govindan, S. Jamin, S. Shenker, and W. Willinger. Network topology generators: Degree-based vs. structural. *ACM Sigcomm'02, Pittsburgh, Pennsylvania, pp. 19-23, USA*, 2002.

[35] P. Van Mieghem, The limit random variable W of a branching process. Submitted to Journal of Applied Probability, 2004.

[36] R. Pastor-Satorras, A. Vázquez and A. Vespignani, Dynamical and correlation properties of the Internet, Phys. Rev. Lett. **87**, 258701, 2001.

APPENDIX A: On $\mathbb{E}\left[\log \mathcal{W} | \mathcal{W} > 0\right]$.

Consider a supercritical Galton-Watson process with $\mathcal{Z}_0 = 1$,

$$\mathcal{Z}_{n+1} = \xi_1 + \xi_2 + \ldots + \xi_{\mathcal{Z}_n}, \qquad n = 0, 1, \ldots,$$

where $\xi_1, \xi_2, \ldots$ is an i.i.d. sequence with mean $1 < \mathbb{E}[\xi] = \nu < \infty$ and generating function

$$f(s) = \sum_{k=0}^{\infty} s^k \mathbb{P}\left[\xi = k\right].$$

The martingale sequence is given by $\mathcal{W}_n = \mathcal{Z}_n/\nu^n$, $n = 1, 2, \ldots$ and $\mathcal{W}_n \to \mathcal{W}$, almost surely. The expectation $\mathbb{E}[\mathcal{W}] = 1$ if and only if [26],

$$\sum_{k=1}^{\infty} (k \log k) \mathbb{P}[\xi = k] < \infty.$$

Denote by $I_A(\cdot)$ the indicator function of the set A. Concerning the existence of $\mathbb{E}\left[\log \mathcal{W} | \mathcal{W} > 0\right]$, we have

$$\begin{aligned}
\mathbb{E}[I_{(0,\infty)}(\mathcal{W}) \log \mathcal{W}] &= \mathbb{E}\left[I_{(0,\infty)}(\mathcal{W}) \int_1^{\mathcal{W}} \frac{dy}{y}\right] \\
&= \mathbb{E}\left[-\int_0^1 I_{(0,y)}(\mathcal{W}) \frac{dy}{y}\right] + \mathbb{E}\left[\int_1^{\infty} I_{(y,\infty)}(\mathcal{W}) \frac{dy}{y}\right] \\
&= -\int_0^1 \mathbb{P}[0 < \mathcal{W} < y] \frac{dy}{y} + \int_1^{\infty} \mathbb{P}[\mathcal{W} > y] \frac{dy}{y},
\end{aligned}$$

by Fubini's theorem [28]. Both integrals on the right-hand side are finite. The second because $\mathbb{E}[\mathcal{W}] = 1$, the first one because [19, Theorem 1] implies that $\mathbb{P}(0 < \mathcal{W} < y) \leq Cy^{\alpha}$, $0 < y < 1$, for some positive α and some positive C, if we assume that $f'(\pi_0) \neq 0$, with π_0 the extinction probability. Similarly we have

$$\mathbb{E}[I_{(0,\infty)}(\mathcal{W}_n) \log \mathcal{W}_n] = -\int_0^1 \mathbb{P}[0 < \mathcal{W}_n < y] \frac{dy}{y} + \int_1^{\infty} \mathbb{P}[\mathcal{W}_n > y] \frac{dy}{y}.$$

From estimates of the difference

$$\mathbb{P}[\mathcal{W}_n > y] - \mathbb{P}[\mathcal{W} > y],\ y > 0,$$

([6], and the references therein) and the form of the density of $\mathcal{W}|\mathcal{W} > 0$ [19, Theorem 1], one can conclude that $\mathbb{E}[\log \mathcal{W}_n|\mathcal{W}_n > 0]$ converges to $\mathbb{E}[\log \mathcal{W}|\mathcal{W} > 0]$.

Appendix B: Accuracy test.

We tested the speed of convergence of $\mathbb{E}[\log \mathcal{W}_n|\mathcal{W}_n > 0]$ to $\mathbb{E}[\log \mathcal{W}|\mathcal{W} > 0]$, using the `FFT`, for geometrically distributed offspring:

$$\mathbb{P}[\xi = k] = p(1-p)^k, \quad k \geq 0, \quad \nu = \mathbb{E}[\xi] = \frac{1-p}{p}.$$

In this case $\mathcal{W}|\mathcal{W} > 0$ has an exponential density on $(0, \infty)$ and

$$\mathbb{E}[\log \mathcal{W}|\mathcal{W} > 0] = -\gamma - \log(1 - \nu^{-1}),$$

where $\gamma = 0.5772\ldots$ is Euler's constant (see [21, p.92]). Moreover, the extinction probability $\pi_0 = \nu^{-1}$. We checked our numerical calculations for the case $\nu = 1.25$. As remarked in the introduction the maximum number of generations that can accurately be computed is limited since the support $\mathcal{Z}_n$ grows exponentially with n. The results with the `FFT` remain reliable only if the theoretical value of $\nu^n = \mathbb{E}[\mathcal{Z}_n]$ does not deviate too much from its numerical value using the `FFT`. As a rule of thumb we use that ν^n (second column of Table 7) does not differ more than 0.1 from its numerical value (third column of Table 7)

An accurate estimate for $\mathbb{E}[\log \mathcal{W}|\mathcal{W} > 0]$ is the value 1.03 given in generation $n = 20$. We see that starting with generation $n = 21$ the numerical values of $\mathbb{E}[\log \mathcal{W}|\mathcal{W} > 0]$ are decreasing, whereas up to and including generation $n = 20$, the values are increasing. The decreasing values are most likely caused by the deviation between the numerically computed distribution and its theoretical counterpart (see the differences in the second and third column).

If we use extrapolation on the values $\mathbb{E}[\log \mathcal{W}_n|\mathcal{W}_n > 0]$ for the generations $n = 12, 13, \ldots, 20$, we obtain $\mathbb{E}[\log \mathcal{W}|\mathcal{W} > 0] = 1.032$. The theoretical value is $\mathbb{E}[\log \mathcal{W}|\mathcal{W} > 0] = -\gamma + \log(5) = 1.0322\ldots$.

Appendix C: The limit random variable $\mathcal{W}$.

Table 7: Results for the geometric distribution with $\nu = 1.25$

n	ν^n	ν^n(num.)	$\mathbb{E}[\log \mathcal{W}_n \vert \mathcal{W}_n > 0]$
12	14.5519	14.5519	1.0087
13	18.1899	18.1899	1.0147
14	22.7374	22.7374	1.0192
15	28.4217	28.4217	1.0226
16	35.5271	35.5271	1.0251
17	44.4089	44.4089	1.0270
18	55.5112	55.5109	1.0285
19	69.3839	69.3836	1.0295
20	86.7362	86.6768	1.0301
21	108.4202	108.0140	1.0296
22	135.5253	133.6421	1.0256
23	169.4066	162.9979	1.0135
24	211.7582	194.4921	0.9869
25	264.6978	226.1251	0.9399
26	330.8722	255.9882	0.8690

The limit random variable $\mathcal{W}$ has an atom at 0 and possesses a density on $(0, \infty)$. It's moment generating function $\chi_{\mathcal{W}}(t) = \mathbb{E}\left[e^{-t\mathcal{W}}\right]$ obeys the functional equation

$$\varphi_{\mathcal{W}}(\nu t) = f\left(\varphi_{\mathcal{W}}(t)\right), \tag{23}$$

where $f(s) = \mathbb{E}\left[s^{\xi}\right]$ is the generating function of the offspring distribution. If $f(s)$ possesses a Taylor expansion around $s = 1$, all Taylor coefficient of $\varphi_{\mathcal{W}}(t)$ can be computed [35]. Using $\log x = \int_0^\infty t^{-1}(e^{-t} - e^{-tx})dt$, we obtain that

$$\mathbb{E}\left[\log \mathcal{W} | \mathcal{W} > 0\right] = \frac{1}{1-\pi_0} \int_0^\infty \frac{(1-\pi_0)\, e^{-t} - \varphi_{\mathcal{W}}(t) + \pi_0}{t} dt, \tag{24}$$

where π_0 is the extinction probability, e.g. the smallest non-negative solution of $s = f(s)$.

In Figure 3 we have computed $\mathbb{E}[\log \mathcal{W} | \mathcal{W} > 0]$ in case of a Poisson offspring distribution ($f(s) = e^{\nu(s-1)}$), for various values ν. For $1.1 \leq \nu \leq 1.9$ we have used the Fast Fourier Transform, whereas for $\nu \geq 2$ relation (24) has been used. The power series of $\varphi_{\mathcal{W}}(t)$ has been transformed by Euler's transformation into a numerically well-converging series that allows the computation of the integral in (24).

The values obtained here are plugged into (13) for the numerical values of $\mathbb{E}[H_N | H_N < \infty]$ in Figure 2.

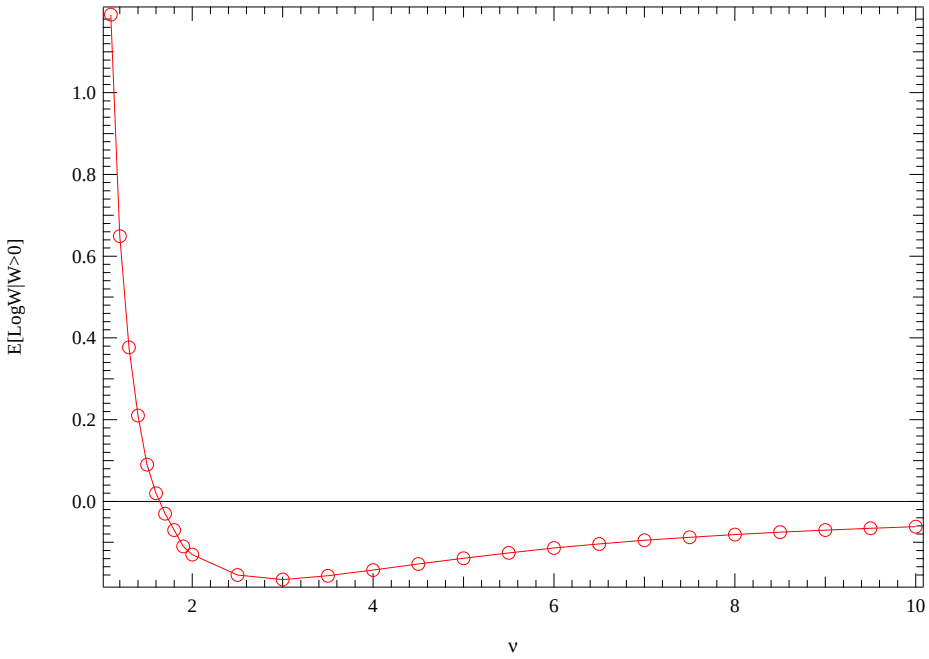

Figure 3: The value of $\mathbb{E}\left[\log \mathcal{W} | \mathcal{W} > 0\right]$ as a function of ν for Poisson(ν) offspring.