# Scale-free graphs of increasing degree

Colin Cooper[*] Pawel Pralat[†]

Draft

**Abstract**

We study a scale-free random graph process in which the number of edges added at each step increases. This differs from the standard model in which a fixed number, $m$, of edges are added at each step.

Let $f(t)$ be the number of edges added at step $t$. In the standard scale-free model, $f(t) = m$ constant, whereas in this paper we consider $f(t) = [t^c],\ c > 0$. Such a graph process, in which the number of edges grows non-linearly with the number of vertices is said to have accelerating growth.

We analyze both an undirected and a directed process. The power law of the degree sequence of these processes exhibits widely differing behaviour.

For the undirected process, the terminal vertex of each edge is chosen by preferential attachment based on vertex degree. When $f(t) = m$ constant, this is the standard scale-free model, and the power law of the degree sequence is 3. When $f(t) = [t^c],\ c < 1$, the degree sequence of the process exhibits a power law with parameter $x = (3-c)/(1-c)$. As $c \to 0$, $x \to 3$, which gives a value of $x = 3$, as in standard scale-free model. Thus no more slowly growing monotone function $f(t)$ alters the power law of this model away from $x = 3$. When $c = 1$, so that $f(t) = t$, the expected degree of all vertices is $t$, the vertex degree is concentrated, and the degree sequence does not have a power law.

For the directed process, the terminal vertex is chosen proportional to in-degree plus an additive constant, to allow the selection of vertices of in-degree zero. For this process when $f(t) = m$ is constant, the power law of the degree sequence is $x = 2 + 1/m$. When $f(t) = [t^c]$, $c > 0$, the power law becomes $x = 1 + 1/(1 + c)$, which naturally extends the power law to $(1, 2]$.

[*]Department of Computer Science, King's College, University of London, London WC2R 2LS, UK. *e-mail:* `colin.cooper@kcl.ac.uk`

[†]Department of Mathematics, West Virginia University, Morgantown, WV 26506-6310, USA. *e-mail:* `pralat@math.wvu.edu`

# 1 Introduction

Discrete random graph processes exhibiting power law properties have been studied by many authors and in many contexts. The study of such processes dates back at least, to Yule [31] in 1924. Recent interest in preferential attachment models follows from the work of Barabási and Albert [1] who observed a power law degree sequence for a subgraph of the World Wide Web, and of Faloutsos, Faloutsos and Faloutsos [15] who observed power law behaviour for the internet graph. Many models of such process exist. For details see, for example, the surveys [3, 24] and the monograph [5].

A graph process is said to have accelerating growth if the number of edges grows non-linearly with the number of vertices. In this paper, we consider a model of accelerating growth in which the number of edges $f(t)$ added at step $t$ is an increasing function of $t$. Thus the model supposes that vertices are in some sense greedy, and that, as each newly added vertex has more existing vertices to join to, the number of edges added also increases.

The concept of accelerating growth was studied by Dorogovtsev and Mendes [12, 13, 14]. The papers distinguish between two types of power law degree distribution, stationary and non-stationary. In the stationary case, the proportion $P(k,t)$ of vertices of degree $k$ at step $t$, satisfies $P(k,t) \propto k^{-\gamma}$, whereas in the non-stationary case $P(k,t) \propto t^z k^{-\gamma}$. The authors make two types of analysis. (i) From an assumption of a power law degree sequence, they infer the (average) number of edges added at a given step, by using feasibility arguments. (ii) From a general condition that $t^a$ edges are added preferentially but arbitrarily at step $t$, they infer the power law, by using average case differential equation arguments in a continuous model. In this way, they derive relationships between the parameters for the undirected (non-stationary) and directed (stationary) processes considered in this paper, and a wide range of other models besides.

Dorogovtsev and Mendes assume the proportion of vertices of degree $k$ satisfies $P(k,t) \propto t^z k^{-\gamma}$, and use the notation $\overline{k}(t) = t^a$ for average degree. They obtain an expression for the average degree parameter $a$ as a function of $\gamma$. We note that this backward approach is also considered in [19]. The results of [12, 13, 14] correspond to those of this paper as follows: For the processes we describe, $f(t) = t^c$ edges are added at each step. Thus the average degree $\overline{k}(t) \propto t^c$. If $\gamma > 2$ then $\gamma = 1 + z/a$, which agrees with (5) for the undirected process. To see this put $z = (1+c)/(1-c) - 1$ and $a = c$. If $1 < \gamma \leq 2$ then $\gamma = 1 + (z+1)/(a+1)$, and when $z = 0$ (stationary process) this agrees with (6) for the directed process.

Dorogovtsev and Mendes also present substantial empirical results for many networks, whose relevance is as follows. For certain directed processes, e.g. the one studied in this paper, the power law for in-degree is above 2 when $f(t) = m$ constant, and below 2 when $f(t) \to \infty$. The power law for in-degree of the World Wide Web in 1999 was estimated at 2.1 ([6]). As this is close to 2, the authors of [13] ponder the evidence for accelerating growth in the Web. The paper [13] generated interest in models of accelerating growth [28, 30], and in evidence for accelerating growth in real life networks [23].

Another approach to networks where the edge density changes over time is taken by Leskovec, Kleinberg and Faloutsos [18, 19]. They look for evidence of increasing edge density (densification) and decreasing diameter in existing networks. Among networks found by the authors to exhibit increasing average out-degree over time were: ArXiv citation, patent citation and autonomous systems (internet routers). The authors are concerned to find causal models of densification, and propose explanatory mechanisms for this, such as community guided attachment, and the forest fire model. Simulations of the forest fire model show that densification itself does not necessarily cause the diameter to shrink. It is observed empirically that the densification of graphs which evolve over time often follows a relation[1] $e(t) = n(t)^a$. The authors of [18] mention that "while one could clearly define a graph model in which $e(t) = n(t)^a$ by simply having each node, when it arrives at time $t$, generate $n(t)^{a-1}$ out-links, [...] Such a model would not provide any insight into the origin of the exponent $a$ as the exponent is unrelated to the operational details by which the network is being constructed". Although this is undeniably true, we maintain that scale-free graph process are of independent interest as simple models of network growth, and that it is useful to formally determine the degree sequence as a function of the rate at which edges are added.

In this paper we are concerned with studying the limiting behaviour of scale-free processes when the degree of newly added vertices increases over time. We address the following questions:
(i) How fast does the number of edges $f(t)$ added to a scale-free process have to grow in order to escape from power law 3 for the degree sequence.
(ii) What is the effect of accelerating growth on the degree sequence of the process.

Before answering these questions, we discuss some existing models in which a constant number, $m$, of edges are added at each step.

In the simplest scale-free model, at each time step $t$, the graph grows by the addition of a new vertex $v_t$ which directs an edge $e_t = (v_t, u)$ towards the existing graph $G(t-1)$. The terminal vertex $u$ of this edge is chosen preferentially, that is, with probability proportional to its degree $d(u,t)$ at the start of step $t+1$. In this model

$$\mathbf{Pr}(v_{t+1} \text{ chooses } u \text{ as the terminal vertex of edge } e) = \frac{d(u,t)}{2t+1}, \tag{1}$$

where $d(v_{t+1}, t) = 1$ so that a loop can occur at $v_{t+1}$.

The simplest generalization of the process $G(t) = G_1(t)$ is the model $G_m(t)$ where a constant number $m \geq 2$ edges are added at each step. The easiest way to achieve this, is to allow a step to consist of $m$ steps in the one edge model $G(t) = G_1(t)$, so that $G_m(t) = G_1(mt)$ where vertex $u_t$ of $G_m(t)$ consists of vertices $v_{m(t-1)+1}, \ldots, v_{mt}$ of $G_1(mt)$. The scale-free model was studied using this approach by Bollobas, Riordan, Spencer, and Tusnady [2], see also Bollobas and Riordan [4].

It is well known that preferential attachment processes such as these have a degree sequence

[1] edges $e(t)$, vertices $n(t)$, parameter $a > 0$

which exhibits a power law. As $k \to \infty$ the expected number $\mathbf{E}N_k$ of vertices of degree $k$ satisfies

$$\mathbf{E}N_k \sim tCk^{-x}.$$

For the model $G_m(t)$ described above, the power law parameter is $x = 3$. It is known precisely from [2] that

$$\mathbf{E}N_k \sim t\,\frac{2m(m+1)}{k(k+1)(k+2)}, \tag{2}$$

and thus

$$\mathbf{E}N_k \sim t(2m(m+1))k^{-3}(1 + O(1/k)),$$

giving $C \sim 2m(m+1)$ in the asymptotic formula.

It is curious that the power law parameter $x$ in the above model exhibits no dependence whatsoever on the number of edges added at each step, however large but constant the value of $m$. With this in mind, a reasonable question is the following. Suppose we make $m$ is a non-decreasing function of $t$, that is, $m(t) = f(t)$. How fast does $f(t)$ have to grow before the dependence on $m(t)$ becomes apparent in the power law parameter $x$? For example is $f(t) \sim \log t$ large enough, or $f(t) \sim t^c, c < 1$, or $f(t) = t$? Moreover, is there a value of $f(t)$ growing with $t$ for which the graph no longer exhibits a power law degree sequence? The growth rate we consider is $f(t) = [t^c]$, (where $[y]$ denotes the rounding of $y$ to the nearest integer). Indeed, for the undirected model we show that no slower growth rate alters the power law away from $x = 3$.

Turning to directed processes, perhaps the simplest model is one where initially, at $t = 0$ the graph is an isolated vertex, and at each step $t = 1, 2, 3, \ldots$ each new vertex adds $m$ edges, where

$$\mathbf{Pr}(v_{t+1} \text{ chooses } u \text{ as the terminal vertex of edge } e) = \frac{d^-(u,t)+1}{(m+1)t+1}. \tag{3}$$

The factor of 1 in the numerator allows vertices of in-degree zero to be chosen. A similar but more general model is given by

$$\mathbf{Pr}(v_{t+1} \text{ chooses } u \text{ as the terminal vertex of edge } e) = \alpha\frac{d^-(u,t)}{|E(t)|} + (1-\alpha)\frac{1}{|V(t)|}, \tag{4}$$

where $0 < \alpha \leq 1$ is a parameter. Choosing $\alpha = m/(m+1)$ we recover (3) (asymptotically). This model has been studied by, for example, [9, 16, 17]. The in-degree sequence for the process (4) has power law $x = 1 + 1/\alpha$. The model given in (3) is thus equivalent to choosing $\alpha = m/(m+1)$ in (4), and the model (3) has power law $x = 2 + 1/m$. In contrast to the undirected scale-free model, the process now exhibits a dependence on $m$, the number of edges added. As $m \to \infty$ and $x \to 2$, and this dependency appears to vanish.

The simplest generalization of the undirected scale-free model equivalent to (4), also chooses a terminal vertex proportional to vertex degree with probability $\alpha$. This model has power law $x = 1 + 2/\alpha$ for the degree sequence (see, for example, [9] or [11]). Thus a dependence on $m$

in the undirected case can be introduced by using the generalized model with $\alpha = m/(m+1)$. Our main interest in this paper, however, is to move the scale-free model ($\alpha = 1$) away from power law $x = 3$ by using a growth rate $t^c$.

To summarize, when $m$ edges are added at each step, the power laws for the models given above are:

(i) Scale free model, $x = 3$,

(ii) Generalized undirected model, $x = 1 + 2/\alpha$,

(iii) Generalized directed model, $x = 1 + 1/\alpha$.

We next state the main properties of the scale-free and directed models, on the assumption that $f(t) = [t^c]$, and highlight the differences in power law properties of their degree sequences.

## 1.1 Undirected Scale Free Model

The model $G_f(t)$ used in this paper is a variant on the scale-free model. We refer to $G_f(t)$ as an $\boldsymbol{f}$-process, where $\boldsymbol{f}(t) = (f(v_1), f(v_2), \ldots, f(v_t), \ldots)$ and $f(v_j)$ is the number of edges directed from $v_j$ to $G(v_j - 1)$. Thus in our notation $G_m(t)$ is an $m$-process $\boldsymbol{m} = (m, \ldots, m, \ldots)$.

The process $(G_f(t))_{t \geq 0}$ is constructed as follows: Let $G_f(t) = (V(t), E(t))$, where $G_f(0)$ is a single vertex with a loop, $V(t) = V(t-1) \cup \{v_t\}$, and $E(t) = E(t-1) \cup \{e_1, \ldots, e_{f(t)}\}$. The vertex $v_t$, $t \geq 1$, (which is usually referred to by its step label $t$), directs the $f(t)$ edges $e_i$, $i = 1, \ldots, f(t)$ to $G_f(t-1) \cup \{e_1, \ldots, e_{i-1}\}$ and

$$\mathbf{Pr}(v_t \text{ chooses } u \text{ as terminal vertex of edge } e_i) = \frac{d(u, t, i-1)}{2|E(t, i-1)|},$$

where $|E(t, i-1)|$ is the number of edges in the graph after $e_{i-1}$ has been added to $v_t$, so that $E(t, i-1) = E(t-1) \cup \{e_1, \ldots, e_{i-1}\}$, and $d(v_t, t, 0) = 0$. The first edge $e_1$ of $v_t$ chooses its terminal vertex in $G_f(t-1)$ ensuring that the graph is connected. For $i > 1$ the edge $e_i$ can choose $v_t$ preferentially as its terminal vertex $u$. This follows as $d(v_t, t, i-1) \geq i - 1$. Thus the model allows loops and parallel edges. The model differs from the scale-free one of (1), in that an edge cannot 'choose itself' to form a self loop.

For positive $y$ let $[y]$ denote the rounding of $y$ to the nearest integer. We consider the model where $f(t) = [t^c]$ edges are added at each step. We prove that for $0 < c < 1$, constant, there is a power law degree sequence, whereas if $c = 1$, all vertices have expected degree close to $t$ and the vertex degree is sharply concentrated about this value. Essentially, for $0 < c < 1$, the number $N_{k,t}$ of vertices of degree $k$ at step $t$ has expectation $\mathbf{E}N_{k,t} \sim n_k$ where

$$n_k = \frac{2\sqrt{1 - \xi(k)}}{1 - c} \left( \frac{t^{1+c}}{k^{3-c}} \right)^{\frac{1}{1-c}},$$

and $\xi(k) \to 0$ for large enough $k$. This gives a power law parameter of $x = (3-c)/(1-c)$. If $c \to 0$ the power law parameter $x \to 3$. Thus we need to add at least $m(t) = [t^c]$ ($c > 0$ constant) edges at each step, to move the power law parameter $x$ away from the value $x = 3$ which is obtained when $m$ is constant. The value $(1+c)/(1-c)$ exponent of $t$ in $n_k$ is a consequence of the total degree of the graph $(\Theta(t^{1+c}))$ when $t^c$ edges are added at each step. The same value for the exponent of $t$ is obtained in [13] using purely heuristic normalizing arguments.

Let

$$\omega = \log t,$$

a notation we preserve throughout this paper. The following theorem summarizes properties of the degree sequence.

**Theorem 1.** *Let* $0 < c < 1$, *and let* $A = \max(5, 2/(1-c))$. *For integer* $k$, *let*

$$n_k = \frac{2\sqrt{1-\xi(k)}}{1-c}\left(\frac{t^{1+c}}{k^{3-c}}\right)^{\frac{1}{1-c}}, \tag{5}$$

*where* $\xi(k) = (s_0/t)^{(1+c)/2}$ *and* $s_0$ *is the real solution to* $k = s_0^c(t/s_0)^{(1+c)/2}$.

*Let* $f(t) = [t^c]$, *then the following results hold for the undirected* $\boldsymbol{f}$*-process* $G_f(t)$:

*(i) Let* $N_{k,t}$ *be the number of vertices of degree* $k$ *at step* $t$. *Let* $k_1 = t^c(1+\varphi/\omega)$, *where* $\varphi \to \infty$ *arbitrarily slowly and* $\varphi = o(\omega)$. *Let* $k_2 = t^{(1+c)/(3-c)}/\omega^A$. *For* $k$ *integer,* $k_1 \le k \le k_2$ *then*

$$\mathbf{E}N_k = (1+o(1))n_k.$$

*(ii) All but* $O(\omega^A)$ *vertices have degree at least* $t^c - O(\omega)$ **whp**.

For vertex $v_s$ added at step $s$, the degree $d(s,t)$ of $v_s$ at step $t$ has an expected value $\mathbf{E}d(s,t) \sim s^c(t/s)^{(1+c)/2}$, (see Lemma 1(i)). Thus the expected minimum degree $\sim t^c$ and the expected maximum degree $\sim t^{(1+c)/2}$. To see this, note that as $c < 1$, $c < (1+c)/2$, the function $h(s) = s^c(t/s)^{(1+c)/2}$ is monotone decreasing in $s$.

The explanation of $s_0$ is as follows: Choosing an integer $k$ we can find a real $s_0$ for which $k$ acts as the equivalent expected degree. Vertices of degree $d(s,t) = k$ should have been added at values of $s$ close to $s_0$. When $s_0 = o(t)$, the value $\xi(k)$ is $o(1)$. When $s_0 = \Omega(t)$ then $\xi$ corrects the power law for the 'low degree'vertices, in the same way as (2) above. Essentially, the theorem characterizes the degree sequence of vertices added between steps $t^{(1+c)/(3-c)+o(1)}$ and $t - o(t)$. The value $k_1 = t^c(1+\varphi/\omega)$ in the theorem corresponds (approximately) to the degree of vertices added around step $t(1-\Theta(\varphi/\omega))$, i.e. $t - o(t)$ or earlier. The value $k_2$ is somewhat arbitrary, and has been chosen so that $n_{k_2} \to \infty$ faster than $\omega^A$, the order of the error term in the proof of Theorem 1(i).

We remark that in the range $[k_1, k_2]$ given in Theorem 1 above, $N_k$ is concentrated about the mean $n_k$. We do not include a full proof of this, but briefly explain how this can be obtained. For any $c$ it can be proved, using the methods of [11], that the vertex degrees are asymptotically independent, i.e.,

$$\mathbf{Pr}(d(v,t) = a,\ d(w,t) = b) \ = \ (1 + o(1))\, \mathbf{Pr}(d(v,t) = a)\, \mathbf{Pr}(d(w,t) = b).$$

The Chebychev inequality then gives concentration of $N_k$ within the stated range.

Turning to the case $c = 1$ we have:

**Theorem 2.** *Let $c = 1$, so that $f(t) = t$. The following hold for $G_t(t)$:*

*(i)* $\mathbf{E}d(v,t) = t(1 + O(1/v))$,

*(ii) All but $O(\omega^5)$ vertices $v$ have degree $d(v,t)$ in $[t - O(\omega^2),\ t + O(\omega^2)]$* **whp**.

## 1.2 Directed Model.

The following theorem summarizes properties of the degree sequence.

**Theorem 3.** *Let $c > 0$. Let $N_{k,t}$ be the number of vertices of in-degree $k$ at step $t$. Let $f(t) = [t^c]$. The following results hold for the directed $\boldsymbol{f}$-process $G_f(t)$:*

*(i) Expected degree sequence:*

$$\mathbf{E}(N_{k,t}) = c_k t(1 + o(1)), \qquad k \geq 0,$$

*where,*

$$c_k = \frac{k!\, \Gamma(1 + 1/(1+c))}{(1+c)\, \Gamma(k + 2 + 1/(1+c))}.$$

*(ii) Power law for degree sequence:*

$$\mathbf{E}(N_{k,t}) \sim B(1 + O(1/k))tk^{-\left(1+\frac{1}{1+c}\right)}, \tag{6}$$

*where $B$ is a positive constant.*

*(iii) Concentration of degree sequence:*
*For $k$ integer, $0 \leq k \leq \left(n/\log^8 n\right)^{\frac{1+c}{6+4c}}$* **whp**

$$N_{k,t} = \mathbf{E}(N_{k,t})(1 + o(1)).$$

Directed models of this type were used by Dorogovtsev and Mendes [13] to model networks where growth in the number of links accelerates as the network evolves. Their analysis has a different approach, that of mean field theory, but their results are the same as those reported here.

## 1.3 Comparison of results.

As can be seen from the above theorems, the power law behaviour of the models is as different as it could possibly be. In terms of the notation of [12] the degree distribution of the directed model is stationary ($P(k,t) \propto k^{-\gamma}$) whereas the undirected model is non-stationary ($P(k,t) \propto t^z k^{-\gamma}$).

For the directed model, when $f(t) = m$ constant, we have (as $k \to \infty$) that

$$\mathbf{E}(N_{k,t}) \sim tk^{-\left(2+\frac{1}{m}\right)},$$

giving a power law in $(2, 3]$ for the in-degree, which tends to 2 as $m \to \infty$. For the corresponding $f(t) = [t^c]$-process, $c > 0$, we have

$$\mathbf{E}(N_{k,t}) \sim tk^{-\left(1+\frac{1}{1+c}\right)},$$

as $k \to \infty$. This gives a power law $x \in (1, 2)$ for the in-degree which tends to 1 as $c \to \infty$, and thus extending the range of the constant $m$-process. The fact that the directed model is stationary arises from the presence, in the long run, of a constant proportion of vertices of in-degree zero. The structure of the digraph $G(T)$ at some much earlier time $T$ does not **whp** strongly influence the structure of $G(t)$.

For the undirected scale-free model, when $f(t) = m$ constant, the power law is invariant at $x = 3$. For the corresponding $f(t) = [t^c]$-process, $0 < c < 1$, for a suitable range of $k$ we have

$$\mathbf{E}(N_{k,t}) \sim \frac{2}{1-c}\left(\frac{t^{1+c}}{k^{3-c}}\right)^{\frac{1}{1-c}},$$

giving a power law of the form $x = (3 - c)/(1 - c)$. As $c \to 0$ the power law coefficient $x$ tends to 3, from above, and we have $x > 3$ in contrast to the directed model where $x$ tends to 2 from below.

For the undirected model, we see from Lemma 1(i) that for any $c > 0$ the expected degree $\mathbf{E}d(v,t) \sim v^c(t/v)^{(1+c)/2}$. Thus for $c < 1$, the oldest vertices have the highest expected degree $t^{(1+c)/2}$. When $c = 1$, the degree distribution is concentrated about $\sim t$. Thus power laws only occur for $0 < c < 1$ in the undirected model, in contrast to the directed case where power laws persists for all $c > 0$. For $c > 1$, the most recently added vertices have highest expected degree, $\sim t^c$.

For the directed model, it can easily be proved that for any $c > 0$ the expected in-degree $\mathbf{E}d^-(v,t) \sim (t/v)^{1+c+o(1)}$. Thus the oldest vertices have highest expected in-degree $t^{1+c+o(1)}$, which is much higher than the value in the undirected case ($t^{(1+c)/2}$). The out-degree of vertex $v$ is $[v^c]$, so the expected total degree $\mathbf{E}d(v,t) \sim v^c + (t/v)^{1+c+o(1)}$. The value of $\mathbf{E}d(v,t)$ is minimum when $v = t^{(1+c)/(1+2c)+o(1)}$, and the process is similar to the Protean Process [22, 27, 26, 21].

Finally, for the purpose of comparison, we state, without proof, the results for the generalized undirected model. These follow directly by applying the proof method of Section 2 to the results for the single edge undirected process with parameter $\alpha$ (see [11] for details).

Let $\eta = \alpha(1+c)/2$, and suppose that $\eta - c > 0$. The minimum degree is effectively $t^c$, and for $k > t^c$, the expected number of vertices of degree $k$ (in a suitable range) is given by,

$$n_k \sim \frac{1}{\eta - c}\, t^{1+\frac{c}{\eta-c}}\, k^{-\left(1+\frac{1}{\eta-c}\right)}.$$

Putting $\alpha = 1$, as in the scale-free model, we obtain (5). When $c \to \eta$, i.e. when $\alpha \to 2c/(1+c)$ from above, all vertices have (approximately) the same degree, which is of order $t^c$, and there is no power law degree sequence. Thus choosing $\alpha < 1$ (i.e. allowing edges to choose terminal vertices u.a.r.) reduces the value of $c$ below which a power law holds, from $c = 1$ for the scale-free model, to $c = \alpha(2-\alpha) < 1$.

As a closing remark, it seems reasonable to wonder if the degree sequence power law parameter, $x$, is the right quantity to measure experimentally in accelerating growth networks. The fact that $x < 2$ for directed and $x > 3$ for undirected processes seems uninformative, and gives limited insight.

# 2 Undirected model: Proof of results

Let $G_f(t)$ be a $\boldsymbol{f}$-process, and let $d_f(s,t)$ denote the degree of vertex $v_s$ at the end of step $t$ in the process.

An important observation is the following:
**Equivalence of Processes.** Suppose vertex $\nu$ has degree $d_f(\nu,\sigma) = \phi$ at the end of step $\sigma$ in the $\boldsymbol{f}$-process $G_f(\sigma)$ when $F(\sigma)$ edges have been added. Then the evolution of degree of $\nu$ at step $\tau \geq \sigma$ when $F(\tau)$ edges have been added is equivalent to the evolution of degree of a vertex $v$ in the 1-process $G_1(t)$, at step $t = F(\tau)$, when at step $s = F(\sigma)$ vertex $v$ has degree $d_1(v,s) = \phi$. Thus

$$\mathbf{Pr}_f(d_f(\nu,\tau) = \phi + \ell \mid d_f(\nu,\sigma) = \phi) = \mathbf{Pr}_1(d_1(v,F(\tau)) = \phi + \ell \mid d_1(v,F(\sigma)) = \phi). \quad (7)$$

Using this, the proofs of Theorems 1 and 2 are obtained as follows:

In Lemma 8, we derive precise results for the degree distribution of $G_1(t)$. In particular, we obtain

$$\mathbf{Pr}(d_1(v,t) = a + l \mid d_1(v,s) = a)$$

for suitable $a, l$ and $v \leq s < t$. This is done using the methods of [11], and is included here as an appendix, for completeness, as a special case of the proofs in that paper.

Using (7), we convert back from the 1-process to the equivalent $\boldsymbol{f}$-process. to obtain Lemma 1 of Section 2.1. Theorems 1 and 2 follow in Section 2.2, by summing the distribution of vertices of given degree to obtain

$$\mathbf{E}N_k = \sum_{a,\sigma,\tau} \mathbf{Pr}(d_f(\sigma,\tau) = k, d_f(\sigma,\sigma) = a),$$

and deriving an asymptotic approximation.

The distribution of $d_f(\sigma,\tau)$ is obtained in Section 2.1, and the asymptotics for $\mathbf{E}N_k$ in Section 2.2. Between them they complete the proof of Theorem 1(i).

## 2.1 Distribution of vertex degree

Where there is no ambiguity, we write $d(v,t)$ for $d_f(v,t)$. In the Appendix we make an asymptotic estimate of the degree distribution of the tree process which underlies $G_f(t)$. We use the following shorthand notation. Let

$$\pi(\phi,l,v,t) = \binom{\phi+l-1}{l}\left(\frac{v(1+O(v^{-c}))}{t}\right)^{\left(\phi\frac{1+c}{2}\right)}\left(1-\left(\frac{v(1+\delta(v))}{t}\right)^{\frac{1+c}{2}}\right)^{l}, \tag{8}$$

where $\omega = \log t$ as usual, and $\delta(v) = O(\min(\phi+l, \phi\omega^4)/v^{1+c})$. Thus the precise value of $\pi(\phi,l,v,t)$ lies in an interval which can be determined by inserting the minimum/maximum values of $O(v^{-c})$, $\delta(v)$ arising in the proof given in the Appendix.

**Lemma 1.** *Let $f(t) = [t^c]$. Let $K > 0$ and $A = \max(5, 2/(1-c))$. The following results hold for the $\boldsymbol{f}$-process $G_f(t)$ at step $t$:*

*(i) Expected degree. For $v \geq 1$ and any $c > 0$*

$$\mathbf{E}(d(v,t)) = v^c\left(\frac{t}{v}\right)^{\frac{1+c}{2}}(1+O(v^{-c})).$$

*(ii) Upper bound on vertex degree. For $v \geq 1$, and any $c > 0$*

$$\mathbf{Pr}\left(d(v,t) \geq v^c\left(\frac{t}{v}\right)^{\frac{1+c}{2}} K\log^4 t\right) \leq t^{-K}.$$

*(iii) Occurrence of loops. For $l \geq 1$, $\mathbf{Pr}(d(v,v) = [v^c]+l) = O\left(\frac{1}{lv^{1-c}}\right)^l$.*

*(iv) Distribution of vertex degree for $0 < c < 1$.*
*If $\omega^A \leq v \leq t(1-\delta^*)$, where $\delta^* = o(\omega^5/t^{(1-c)/2})$, then*

$$\mathbf{Pr}(d(v,t) = \phi + l \mid d(v,v) = \phi) = \pi(\phi,l,v,t)\,(1+O(1/\omega)) + O(v^{-K+1}). \tag{9}$$

*(v) Upper bound on distribution of vertex degree.*
*(a) For $0 < c < 1$, and $\omega^A \le v \le t$,*

$$\mathbf{Pr}(d(v,t) = \phi + l \mid d(v,v) = \phi) \le \pi(\phi, l, v, t)\ (1 + O(1/\omega)) + O(v^{-K+1}).$$

*(b) For $c = 1$, and $\omega^5 \le v \le t$,*

$$\mathbf{Pr}(d(v,t) = \phi + l \mid d(v,v) = \phi) \le \pi(\phi, l, v, t)\ (1 + O(1)) + O(v^{-K+1}).$$

**Proof** Using $G_f(t) = (V(t), E(t))$, $|V(0)| = 1$, $|E(0)| = 1$ let $F(\tau) = |E(0)| + \sum_{\nu=1}^{\tau} f(\nu)$. Thus $|V(\tau)| = \tau + 1$ and $|E(\tau)| = F(\tau)$. If $f(t) = [t^c]$ and $F(t) = 1 + \sum_{s=1}^{t} f(s)$ then from Lemma 10 $F(t) = \frac{t^{1+c}}{1+c}(1 + O(t^{-c}))$.

To convert the results of the Appendix for the 1-process, to the $\boldsymbol{f}$-process, we use the notation $\sigma, \tau$ for vertices of $G_f(\tau)$ and $v, s, t$ for vertices of $G_1(t)$.

*Proof of (i), (ii).* Use Lemma 7(i), (ii) with $s = F(\sigma)$ etc. We have e.g.

$$\mathbf{E} d_f(\sigma, \tau) = (1 + O(\omega \sigma^{-c})) f(\sigma) \left( \frac{F(\tau)}{F(\sigma)} \right)^{\frac{1}{2}}. \tag{10}$$

*Proof of (iii).* The degree of vertex $\sigma$ immediately after its addition is $\phi(\sigma) = f(\sigma) + g(\sigma)$ where $g(\sigma)$ is the number of loops arising at $\sigma$ during the addition of the edges $e_i, i = 1, \ldots, [s^c]$. The probability of $e_i$ forming a loop is at most $2(i-1)/2F(\sigma - 1) \le 3/s$. Thus the number of loops is stochastically dominated by Binomial $B([s^c], 3/s)$, and (iii) follows.

*Proof of (iv), (v).* For the proof of (iv) below, we assume **whp** that for $c < 1$, the number of loops $g(\sigma) \le L = K/(1-c)$; and for $c = 1$, $g(s) \le \log t$. In both cases the probability of this not occurring is $O(t^{-K})$.

If $s = F(\sigma) \sim \sigma^{1+c}$ then $d(\sigma, \sigma) = \phi(\sigma) \sim \sigma^c$ and $a = \phi^2 = O(s)$. Thus we can apply Lemma 8 to the 1-process, with $\theta = \min(\phi + l, \phi K (\log F(\tau))^4)$. Using (7), this gives

$$\mathbf{Pr}(d_f(\sigma, \tau) = \phi(\sigma) + l \mid d_f(\sigma, \sigma) = \phi) = O(\sigma^{-(c+1)(K-1)}) \tag{11}$$

$$+ (1 + \epsilon) \binom{\phi + l - 1}{l} \left( \frac{F(\sigma)}{F(\tau)} \right)^{\frac{\phi}{2}} \left( 1 - \left( \frac{F(\sigma)}{F(\tau)} \right)^{\frac{1}{2}} \left(1 + O\left(\frac{\theta}{F(\sigma)}\right)\right) \right)^{l}, \tag{12}$$

and

$$\epsilon = O\left( \frac{\phi^2}{F(\sigma)} \right) - l^2 \left| O\left( \frac{\left( \log \frac{F(\tau)}{F(\sigma)} + \frac{\theta}{F(\sigma)} \right)}{(\sqrt{F(\tau)} - \sqrt{F(\sigma)})^2} \right) \right|. \tag{13}$$

Given that $\sigma \ge \log^5 t$ then $\left( \frac{\theta}{F(\sigma)} \right) = O(1/\omega)$ which deals with the error term in line (12).

The first term in (13) is $O(\sigma^{-(1-c)/2}) = O(1/\omega)$ for $\sigma \ge \omega^{2/(1-c)}$. We note that the second error term on line (13) is negative, and thus provided $\sigma \ge \omega^A$ ignoring it gives an upper bound on the required probability for part (v).

For part (iv), we argue as follows: Let $\sigma = \tau(1-\delta)$. For $0 \le \delta \le 1$, and $0 < a \le 1$, $(1-\delta)^a \le 1 - a\delta$, so $(\sqrt{F(\tau)} - \sqrt{F(\sigma)})^2 \ge (1+o(1))\tau^{1+c}\delta^2(1+c)/4$.

Let $L$ be the **whp** maximum degree given by part (ii). Thus $l \le L = \sigma^c(t/\sigma)^{(1+c)/2}K\omega^4$. We require the error term in (13) to be $\epsilon = O(1/\omega)$, i.e. that $L^2/(\sqrt{F(\tau)} - \sqrt{F(\sigma)})^2 = 1/\omega$. We find this is satisfied by $\delta = B\omega^{9/2}/\tau^{(1-c)/2}$ for some constant $B > 0$. As $\delta^* > \delta$ part (iv) holds. □

We next determine the (**whp**) minimum degree of vertices of the $\boldsymbol{f}$-process added after step $\omega^A$.

**Lemma 2.** *Let $K$ be a large constant. Let $f(t) = [t^c]$. The following results hold for the $\boldsymbol{f}$-process $G_f(t)$ at step $t$:*

*(i) Case $0 < c < 1$: There are at most $\omega^A$ vertices $v$ such that $d(v,t) \le t^c - K\omega$,* **whp**.

*(ii) Case $c = 1$: At most $\omega^5$ vertices $v$ do not satisfy $(t - K\omega^2 \le d(v,t) \le t + K\omega^2)$,* **whp**.

**Proof** To find $\mathbf{Pr}(d_f(v,t) \le (1-\epsilon)t^c)$ we proceed as follows. Referring to Lemma 1(v) and (8) we see that the upper bound for the probability distribution of vertex degree decreases from its central term near the mean, given in Lemma 1(i). Thus we can upper bound $\mathbf{Pr}(d_f(v,t) = (1-\epsilon)t^c)$, and then multiply this by $t^c$ to include lower values. If the formula (8) is expanded using Stirling's approximation we have the following cases.

*Case (i) ($c < 1$).* We condition on the number of loops at $v$ being at most a constant $L$. Using $\pi(\phi, \ell, v, t)$ from (8) with $v^c \le \phi \le v^c + L$, $\phi + \ell = (1-\epsilon)t^c$ where $\epsilon = K\omega/t^c$. Assume $v \ge \omega^A$, and $v \le (1-\delta)t$ where $\delta \ge 2\epsilon/(1-c)$. Let $b = (1+c)/2$ and $v' = v(1+o(1))$. Then

$$\begin{aligned}
\pi(\phi, \ell, v, t) &= o(t^{cL}) \frac{((1-\epsilon)t^c)^{(1-\epsilon)t^c}}{v^{cv^c}((1-\epsilon)t^c - v^c)^{(1-\epsilon)t^c - v^c}} \left(\frac{v'}{t}\right)^{bv^c} \left(1 - \left(\frac{v'}{t}\right)^b\right)^{(1-\epsilon)t^c - v^c} \\
&\le o(t^{cL}) \left( {}_{(1+o(1))} \frac{(1-\epsilon)^{1-\epsilon}}{(1-\delta)^{c(1-\delta)}(\delta-\epsilon)^{\delta-\epsilon}} (1-\delta)^{b(1-\delta)} (1 - (1-\delta_{(1+o(1))})^b)^{\delta-\epsilon} \right)^{t^c} \\
&\le o(t^{cL}) \left( {}_{(1+o(1))}(1-\epsilon)^{1-\epsilon}(1-\delta)^{\frac{1-c}{2}(1-\delta)} \left(\frac{b\delta}{\delta-\epsilon}\right)^{\delta-\epsilon} \right)^{t^c} \\
&\le o(t^{cL}) \exp -t^c{}_{(1+o(1))} \left( \epsilon(1-\epsilon) + \frac{1-c}{2}\delta(1-\delta) \right) \\
&\le o(t^{cL}) \exp(-\epsilon t^c) = O(t^{-K}).
\end{aligned}$$

Part (i) now follows.

*Case (ii) ($c = 1$).* The proof is similar. Let $|\epsilon| = 2K\omega^2/t$, $\delta \ge \omega\epsilon$ and let $L = \log t$. Then from the above expression

$$\pi(\phi, \ell, v, t) = O(t^L) \left( (1-\epsilon)^{1-\epsilon} \left(\frac{\delta}{\delta-\epsilon}\right)^{\delta-\epsilon} \right)^t. \tag{14}$$

Note that $\log(1+x) \leq x - x^2/2 + \max(0, x^3/3)$ for $|x| < 1$. Let $\Psi$ be given by

$$\begin{aligned}
\Psi =&(1-\epsilon)\log(1-\epsilon) + (\delta-\epsilon)\log\left(1+\frac{\epsilon}{\delta-\epsilon}\right) \\
=&(1-\epsilon)(-\epsilon - O(\epsilon^2)) + (\delta-\epsilon)\left(\frac{\epsilon}{\delta-\epsilon} - \frac{1}{2}\left(\frac{\epsilon}{\delta-\epsilon}\right)^2 + O\left(\frac{\epsilon}{\delta-\epsilon}\right)^3\right) \\
=&-\frac{1}{2}\left(\frac{\epsilon^2}{\delta-\epsilon}\right) + O(\epsilon^2) + O\left(\frac{\epsilon^3}{\delta^2}\right) \\
\geq&-\frac{1}{2}\frac{\epsilon}{\omega}\left(1 - O\left(\frac{1}{\omega}\right)\right).
\end{aligned}$$

Thus

$$\mathbf{Pr}(d_f(v,t) \leq (1-\epsilon)t^c) = O(t^L)\exp - O(t\epsilon/\omega) = O(t^{-K}).$$

For $\epsilon \geq 1$, $\Psi$ becomes

$$\Psi = (1+\epsilon)\log(1+\epsilon) + (\delta+\epsilon)\log\left(1 - \frac{\epsilon}{\delta+\epsilon}\right)$$

giving the same result, and thus $\mathbf{Pr}(d_f(v,t) \leq (1+\epsilon)t) \leq O(t^{-K})$ as before. □

Finally, we use the following notation and concentration results in the proofs of the next section. Let $b = (1+c)/2$ and let

$$\theta(v) = \left(\left(\frac{t}{v}\right)(1+\delta(v))\right)^b, \tag{15}$$

where $\delta(v) = O(\min(\phi + l, \phi\omega^4)/v^{1+c})$.

**Lemma 3.** *Let $\omega^A \leq v \leq (1-\epsilon)t$ and $\pi(\phi, l, v, t)$ given by (8). Provided $h = o(\epsilon)$,*

$$\pi(\phi, l, v, t) = O(1)\exp\left(-\frac{\phi\theta}{2(\theta-1)}h^2(1+O(h))\right),$$

**Proof** Expanding (8) using Stirlings approximation, we have

$$\pi = O(1)\frac{(\phi+l)^{\phi+l}}{l^l\ \phi^\phi}\frac{1}{\theta^\phi}\left(1-\frac{1}{\theta}\right)^l.$$

Let $\phi + l = a\phi\theta$, let

$$f(a) = \frac{(a\theta)^{a\theta}}{(a\theta-1)^{a\theta-1}}\frac{1}{\theta}\left(1-\frac{1}{\theta}\right)^{a\theta-1},$$

and let

$$y(a) = \theta\log a\theta + \theta\log(\theta-1) - \theta\log\theta - \theta\log(a\theta-1).$$

The derivatives of $f(a)$ are

$$\begin{aligned} f'(a) &= f(a)\cdot y(a) \\ f''(a) &= f'(a)\cdot y(a) + f(a)\cdot\left(\frac{-\theta}{a(a\theta-1)}\right) \\ f'''(a) &= f''(a)\cdot y(a) + 2f'(a)\cdot\left(\frac{-\theta}{a(a\theta-1)}\right) + f(a)\cdot\left(\frac{\theta(2a\theta-1)}{a^2(a\theta-1)^2}\right), \end{aligned}$$

and $y(1)=0,\ f(1)=1,\ f'(1)=0,\ f''(1)=-\theta/(\theta-1),\ f'''(1)=\theta(2\theta-1)/(\theta-1)^2$. When $a=1+h$, and $\zeta\in[0,1]$ we have

$$f'''(1+\zeta h) = O\left(\frac{\theta^2}{(\theta(1-h)-1)^2}\right).$$

Thus provided $v\le t(1-\epsilon)$, then $\theta-1\ge b\epsilon$ and as $h=o(\epsilon)$ $f'''(1+\zeta h)=f'''(1)(1+o(1))$. It follows that

$$f(1+h) = 1-\frac{h^2\theta}{2(\theta-1)} - \frac{h^3}{6}\frac{\theta(2\theta-1)}{(\theta-1)^2}(1+o(1)).$$

Using $h=o(\epsilon)$ again (i.e. $h/(\theta-1)=o(1)$), the lemma follows. □

The concentration of vertex degree will follow from:

**Lemma 4.** *Let*

$$h(v)=\sqrt{\frac{(\theta(v)-1)2K\log t}{\theta(v)\phi}}, \qquad \mu=\phi[(t/v)(1+\delta(v)]^b.$$

*Then* **whp***, for* $v\le t(1-\varphi/\omega)$

$$\mathbf{Pr}(d(v,t)\notin[(1-h)\mu,(1+h)\mu]) = O(t^{-K+1}).$$

**Proof** As $h(v)=o(\epsilon)=\varphi/\omega$ in Lemma 3, we have

$$\pi(\phi,\mu(1+h),v,t) = O(1)\exp\left(-\frac{\phi\theta}{2(\theta-1)}h^2(1+O(h))\right)\sim O(t^{-K}).$$

As $\pi$ is monotone decreasing about the mode $\sim\mu$, the result follows. □

## 2.2 Proof of Theorem 1

Theorem 1(ii) is proved in Lemma 2. For the proof of Theorem 1(i) we proceed as follows.

Let $s_0=s_0(k)$ be the unique real solution of

$$k = s_0^c\left(\frac{t}{s_0}\right)^{\frac{1+c}{2}}, \tag{16}$$

then $n_k = \frac{2\sqrt{1-\xi}}{1-c}\frac{s_0}{k}$. By Lemma 1(i), for integer $s_0$, the expected degree of vertex $s_0$ is approximately $k$ at time $t$, so intuitively we might expect most vertices of degree $k$ to have labels $s$ close to $s_0$.

Let

$$h(s_0) = \sqrt{\frac{(\theta(s_0)-1)}{\theta(s_0)} \; \frac{2K\log t}{s_0^c}}, \tag{17}$$

for some large constant $K$, and $\theta(v)$ given by (15). It follows from Lemma 4 with that with

$$s_1 = s_0(1-h)^{2/(1-c)}, \quad s_2 = s_0(1+h)^{2/(1-c)},$$

the interval $[s_1,\ldots,s_2]$ contains all vertices $v$ with $d(v,t)=k$ simultaneously for all $v,k$ with probability $1-O(t^{-K+2})$. Conditional on this, define an interval $I(k)$ for vertices of degree $k$ by $I(k) = [(1-\varepsilon)s_0,\ldots,(1+\varepsilon)s_0]$ where $\varepsilon = \max((1\pm h)^{2/(1-c)}-1)$. Note that, for some $a>0$ constant, $\varepsilon = ah$.

For $k_1 \le k \le k_2$, we write

$$\mathbf{E}N_k = \sum_{s>\omega^A} \mathbf{Pr}(d_f(s,t)=k) + O(\omega^A).$$

We prove the first term on the right hand side is asymptotic to $n_k$.

Let $L$ be a large constant. From Lemma 1(iii), the probability that $\phi(v) \ge v^c + L$ is $O(v^{-L(1-c)})$. Thus we can have $\phi(v) = v^c(1+O(v^{-c}))$ with probability $1-O(v^{-L(1-c)})$.

Let $b=(1+c)/2$, let $s' = s(1+O(\delta(s))$, let $s'' = s(1+O(s^{-c}))$ then from (8)

$$\mathbf{Pr}(d_f(s,t)=k \mid d_f(s,s)=\phi(s)) = (1+o(1))\binom{k-1}{k-\phi(s)}\left(\frac{s''}{t}\right)^{b\phi(s)}\left(1-\left(\frac{s'}{t}\right)^b\right)^{k-\phi(s)} + O(s^{-K+1}).$$

From Sterling's formula we have, for $k>\phi$, that

$$\binom{k-1}{k-\phi} = \left(1+O(\tfrac{1}{k}+\tfrac{1}{\phi}+\tfrac{1}{(k-\phi)})\right)\frac{\sqrt{\phi}}{\sqrt{2\pi k(k-\phi)}}\left(\frac{k}{\phi}\right)^{\phi}\left(1-\frac{\phi}{k}\right)^{\phi-k}.$$

Thus

$$\mathbf{Pr}(d_f(s,t)=k \mid \phi) = \left(1+O(\tfrac{1}{k}+\tfrac{1}{\phi}+\tfrac{1}{(k-\phi)})\right)\frac{\sqrt{\phi}}{k\sqrt{2\pi}}\left(e^{O(1/k)}\frac{k}{\phi}\left(\frac{s''}{t}\right)^b\right)^{\phi}\left(\frac{1-\left(\frac{s'}{t}\right)^b}{1-\frac{\phi}{k}}\right)^{k-\phi}.$$

From (16) define $\xi \le 1$ by

$$\xi = \frac{{s_0}^c}{k} = \left(\frac{s_0}{t}\right)^b.$$

From the choice of $k_1$, $\xi \le 1 - \varphi/\omega$. We see from (17) that $h(s_0) = o(1-\xi)$, a condition we require below. Let $s = \lambda s_0$ where $(1-\varepsilon) \le \lambda \le (1+\varepsilon)$, and where $\varepsilon = O(h(s_0))$ comes from $I(k)$ above. Substitute $(s/t)^b = \lambda^b \xi$, $\phi(s) = s^c(1+o(1))$ and $s^c/k = \lambda^c \xi$ to obtain

$$\mathbf{Pr}(d_f(s,t) = k \mid \phi(s)) = {\scriptstyle (1+o(1))} \frac{s_0^{c/2}}{k} \frac{1}{\sqrt{2\pi}} \left( \lambda^{(b-c)\lambda^c} \left( \frac{1-\lambda^b \xi}{1-\lambda^c \xi} \right)^{\frac{1-\lambda^c\xi}{\xi}} {\scriptstyle (1+o(1))} \right)^{s_0^c(1+o(1))} . \quad (18)$$

Let

$$g(\lambda) = (b-c)\lambda^c \log \lambda + \frac{1-\lambda^c \xi}{\xi} \log \left( \frac{1-\lambda^b \xi}{1-\lambda^c \xi} \right),$$

and let

$$G(\lambda) = \frac{1}{\sqrt{2\pi}} \exp(s_0^c\, g(\lambda)(1+o(1))).$$

Thus

$$\sum_{s \ge \omega^A} \mathbf{Pr}(d_f(s,t) = k) = (1+o(1)) \frac{s_0^{1+c/2}}{k} \int_{1-\varepsilon}^{1+\varepsilon} G(\lambda)\, d\lambda. \quad (19)$$

We prove that

$$\int G(\lambda)\, d\lambda = (1+o(1)) \frac{\sqrt{1-\xi}}{(b-c) s_0^{c/2}} \quad (20)$$

and thus $\mathbf{E}N_k = (1+o(1))n_k$ which completes the proof of Theorem 1.

Note that[2]

$$g(1) = 0, \quad g'(1) = 0, \quad g''(1) = -\frac{(b-c)^2}{1-\xi}, \quad g'''(1) = -\frac{(b-c)^2(\xi b + b - 3 + 3\xi - c\xi + 2c)}{(1-\xi)^2},$$

and provided $h = o(1-\xi)$, $g'''(1+\zeta h) = O((1-\xi)^{-2})$ for $\zeta \in [0,1]$. As $1-\xi = \Omega(\varphi/\omega)$, then as remarked before $h = o(1-\xi)$, and so

$$g(1+h) = -\frac{h^2}{2} \frac{(b-c)^2}{1-\xi}(1+o(1)).$$

Thus

$$G(1+h) = \frac{1}{\sqrt{2\pi}} \exp\left( -\frac{s_0^c h^2 (b-c)^2}{2(1-\xi)}(1+o(1)) \right).$$

As long as $s_0 \le t(1 - b\varphi/\omega)$ where $\varphi \to \infty$ slowly, and $\omega = \log t$, we have $\theta(s_0) - 1 \ge b'\varphi/\omega$. Let $\lambda = 1+h$, and $x = s_0^{c/2}(b-c)h/\sqrt{1-\xi}$. When $h = \varepsilon$, and the upper and lower limits of integration of the standardized normal integral are at least $x = (b-c)\sqrt{\varphi} \to \infty$. This gives the required result (and explains our choice of $k_1$). □

[2] The values of $g'(\lambda), \ldots, g'''(\lambda)$ have been checked in Maple

# 3 Directed model: Proof of results

The model produces a sequence $\{G(t)\}_{t=1}^{\infty} = \{(V(t), E(t))\}_{t=1}^{\infty}$ of directed graphs, where $t$ denotes time. The model has one parameter: $c > 0$. Let $G_1 = (V(1), E(1)) = (\{v_1\}, \emptyset)$ be a fixed initial graph with a single isolated vertex. For $t > 1$, form $G(t)$ from $G(t-1)$ by adding a new vertex $v_t$ together with $f(t) = [t^c]$ edges from $v_t$ directed towards existing vertices chosen randomly with weighted probabilities as follows:
The edges, $e_i, i = 1, \ldots, [t^c]$, are added in $[t^c]$ independent sub-steps. In each sub-step, one edge is added, and the probability that $u$ is chosen as its endpoint (the link probability), given by

$$\mathbf{Pr}(v_t \text{ chooses } u \text{ as terminal vertex of edge } e_i) = \frac{d^-(u, t-1) + 1}{|E(t-1)| + |V(t-1)|}. \tag{21}$$

Our model allows multiple edges; there seems no reason to exclude them. Indeed, for $c > 1$ multiple edges cannot be avoided. The expected number of parallel edges added to a vertex of in-degree $d$ at step $t$ is $O(d^2/t^2)$. Thus modifying the model to exclude them (for small $c$) will not to significantly affect the results of Theorem 3.

We approximate (21) as follows. For any $c > 0$, define the function $g_c : \mathbb{N} \to \mathbb{R}$:

$$g_c(t) = |V(t)| + |E(t)| = t + \sum_{j=2}^{t} [j^c] = t + \frac{t^{1+c}}{1+c} + O(1) = \frac{t^{1+c}}{1+c}(1 + O(t^{-c}))\,. \tag{22}$$

See Lemma 10 for a proof of this.

Let $N_{i,t}$ denote the number of vertices of in-degree $i$ in $G(t)$. The equations relating the random variables $N_{i,t}$ are described as follows. As $G_1$ consists of one isolated vertex, $N_{0,1} = 1$, and $N_{i,1} = 0$ for $i > 0$. At time $t + 1 > 0$, we add a new vertex $v_{t+1}$ which initially has in-degree zero and $[(t+1)^c]$ edges. The expected number of these edges which choose a vertex of in-degree zero (at time $t$) as their endpoint is $[(t+1)^c]N_{0,t}/g_c(t)$. Some of those vertices can be chosen more then once during the $[(t+1)^c]$ sub-steps. The expected number of vertices that are chosen exactly $k$ times is $O(([(t+1)^c])^k N_{0,t}/g_c(t)^k) = O(N_{0,t}/t^k)$. Therefore, we get

$$\begin{aligned}
\mathbf{E}(N_{0,t+1} - N_{0,t} \mid G_t) &= 1 - N_{0,t}\frac{1}{g_c(t)}[(t+1)^c] + O(N_{0,t}/t^2) \\
&= 1 - N_{0,t}\frac{1+c}{t}(1 + O(t^{-\min\{c,1\}})).
\end{aligned} \tag{23}$$

Similarly, one can analyze the expected change in $N_{i,t}$ for $i > 0$. The fact that some of these vertices can be chosen more than once affects the formula as before. Note that vertices of degree $i-k$ can receive $k$ parallel edges so all vertices of degree at most $i$ have to be considered. The function $h(i,t) = h(t, N_{0,t}, N_{1,t}, \ldots, N_{i,t})$ introduced below takes care of vertices that are chosen at least twice.

$$\begin{aligned}\mathbf{E}(N_{i,t+1}-N_{i,t}\mid G_t) &= N_{i-1,t}\frac{(i-1)+1}{g_c(t)}[(t+1)^c]-N_{i,t}\frac{i+1}{g_c(t)}[(t+1)^c]+h(i,t)\\ &= \left(N_{i-1,t}\frac{i(1+c)}{t}-N_{i,t}\frac{(i+1)(1+c)}{t}\right)(1+O(t^{-\min\{c,1\}}))+h(i,t).\end{aligned}$$

It can be seen that

$$\begin{aligned}h(i,t) &= O\left(([(t+1)^c])^2 N_{i,t}\left(\frac{i+1}{g_c(t)}\right)^2\right)+O\left(([(t+1)^c])^2 N_{i-1,t}\left(\frac{i}{g_c(t)}\right)^2\right)\\ &\quad +O\left(\sum_{k=2}^{i}([(t+1)^c])^k N_{i-k,t}\left(\frac{i-k+1}{g_c(t)}\right)^k\right)\\ &= O\left((N_{i,t}+N_{i-1,t})\left(\frac{i}{t}\right)^2\right)+O\left(\sum_{k=2}^{i}N_{i-k,t}\left(\frac{i(1+c)}{t}\right)^k\right).\end{aligned}$$

We show below (inductively) that **whp** $N_{i-k,t}=\Theta(N_{i-k-1,t})$ $(1\le k\le i-1)$ which implies that $h(i,t)=O(N_{i,t}i^2/t^2)$ and thus

$$\mathbf{E}(N_{i,t+1}-N_{i,t}\mid G_t) = \left(N_{i-1,t}\frac{i(1+c)}{t}-N_{i,t}\frac{(i+1)(1+c)}{t}\right)(1+O(it^{-\min\{c,1\}})). \quad (24)$$

To do this we first establish the expected value of $N_{0,t}$ using Lemma 5, and then prove the variable $N_{0,t}$ is concentrated using Theorem 4. Assuming this, we then use Lemma 5 and Corollary 5, to establish the same result for $N_{1,t}$ and so on.

Assuming $h(i,t)=O(N_{i,t}i^2/t^2)$, recurrence relations for the expected values of $N_{i,t}$ are obtained by taking the expectation of the above equations. To solve these relations, we use the following lemma on real sequences, which is Lemma 3.1 from [7].

**Lemma 5.** *If* $(\alpha_t)$, $(\beta_t)$ *and* $(\gamma_t)$ *are real sequences satisfying the relation*

$$\alpha_{t+1}=\left(1-\frac{\beta_t}{t}\right)\alpha_t+\gamma_t,$$

*where* $\lim_{t\to\infty}\beta_t=\beta>0$ *and* $\lim_{t\to\infty}\gamma_t=\gamma$, *then* $\lim_{t\to\infty}\frac{\alpha_t}{t}$ *exists and equals* $\frac{\gamma}{1+\beta}$.

Taking expectations again, from (23) it follows that

$$\mathbf{E}(N_{0,t+1})=\left(1-\frac{1+c}{t}(1+O(t^{-\min\{c,1\}}))\right)\mathbf{E}(N_{0,t})+1.$$

Applying Lemma 5 with $\alpha_t=\mathbf{E}(N_{0,t})$, $\beta_t=(1+c)(1+o(1))$, and $\gamma_t=1$ gives that $\mathbf{E}(N_{0,t})=c_0t(1+o(1))$ with

$$c_0=\frac{1}{2+c}.$$

For $i > 0$, the lemma can be inductively applied with $\alpha_t = \mathbf{E}(N_{i,t})$, $\beta_t = (i+1)(1+c)$, and $\gamma_t = \mathbf{E}(N_{i-1,t})\frac{i(1+c)}{t}$ to show that $\mathbf{E}(N_{i,t}) = c_i t(1+o(1))$, where

$$c_i = c_{i-1}\frac{i(1+c)}{1+(i+1)(1+c)},$$

and thus for $i \geq 0$,

$$c_i = \frac{i!\,\Gamma(1+1/(1+c))}{(1+c)\,\Gamma(i+2+1/(1+c))}.$$

It follows that $c_i = (1+O(1/i))Ai^{-(1+\frac{1}{1+c})}$, where $A = \Gamma(1+1/(1+c))/(1+c)$. This shows that for large $i$, the expected proportion $\mathbf{E}(N_{i,n})/n$ has a power law with exponent $x = 1+\frac{1}{1+c}$.

Let

$$i_f = i_f(n) = \left(\frac{n}{\log^8 n}\right)^{\frac{1+c}{6+4c}}.$$

We show that the random variables $N_{i,n}$, $i < i_f$, are **whp** well concentrated around their mean. In order to sketch the technique that can be used, we consider $N_{0,t}$, the number of nodes of in-degree zero. We use the super-martingale method of Pittel et al. [25], as described in [29]. We say that an event holds *with extreme probability* (**wep**) if it holds with probability at least $1-\exp(-\Omega(\log^2 n))$ as $n \to \infty$. In this section, we will often use the stronger notion of **wep** in favor of the more commonly used **whp**, since it simplifies some of our proofs. If we consider a polynomial number of events that each holds **wep**, then **wep** all these events hold.

**Lemma 6.** *Let $G_0, G_1, \ldots, G_n$ be a random graph process and $X_t$ a random variable determined by $G_0, G_1, \ldots, G_t$, $0 \leq t \leq n$. Suppose that for some real $\beta$ and constants $\gamma_t$,*

$$\mathbf{E}(X_t - X_{t-1}|G_0, G_1, \ldots, G_{t-1}) < \beta$$

*and*

$$|X_t - X_{t-1} - \beta| \leq \gamma_t$$

*for $1 \leq t \leq n$. Then for all $\alpha > 0$,*

$$\mathbf{Pr}\big(\textit{For some } t \textit{ with } 0 \leq t \leq n : X_t - X_0 \geq t\beta + \alpha\big) \leq \exp\left(-\frac{\alpha^2}{2\sum_{i=1}^n \gamma_i^2}\right).$$

**Theorem 4. Wep** *for every $1 \leq t \leq n$, we have that*

$$N_{0,t} = \frac{t}{2+c} + O(n^{1/2}\log^3 n) = c_0 t + O(n^{1/2}\log^3 n).$$

*Proof.* We first transform $N_{0,t}$ into something close to a martingale. It provides insight if we define a real function $f(x)$ to model the behaviour of the scaled random variable $\frac{1}{n}N_{0,xn}$. If we presume that changes in the function correspond to the expected change of a random variable (see (23)), we obtain the following differential equation

$$f'(x) = 1 - f(x)\frac{1+c}{x}$$

with the initial condition $f(0) = 0$. The general solution of this equation can be put in the form

$$f(x)x^{1+c} - \frac{x^{2+c}}{2+c} = C.$$

Consider the following real-valued function

$$H(x, y) = x^{1+c}y - \frac{x^{2+c}}{2+c} \tag{25}$$

(note that we expect $H(t, N_{0,t})$ to be close to zero). Let $\mathbf{w}_t = (t, N_{0,t})$, and consider the sequence of random variables $(H(\mathbf{w}_t) : 1 \le i \le n)$. The second-order partial derivatives of $H$ evaluated at $\mathbf{w}_t$ are all $O(t^c)$. Therefore, we have

$$H(\mathbf{w}_{t+1}) - H(\mathbf{w}_t) = (\mathbf{w}_{t+1} - \mathbf{w}_t) \cdot \text{grad } H(\mathbf{w}_t) + O(t^c), \tag{26}$$

where "$\cdot$" denotes the scalar product and grad $H(\mathbf{w}_t) = (H_x(\mathbf{w}_t), H_y(\mathbf{w}_t))$.

Observe that from our choice of $H$, we have that

$$\begin{aligned}
\mathbf{E}(\mathbf{w}_{t+1} - \mathbf{w}_t \mid G_t) \cdot & \text{ grad } H(\mathbf{w}_t) \\
&= \left(1, 1 - N_{0,t}\frac{1+c}{t}(1 + O(t^{-\min\{c,1\}}))\right) \cdot \left((1+c)t^c N_{0,t} - t^{1+c}, t^{1+c}\right) \\
&= O(t^c).
\end{aligned}$$

Hence, taking the expectation of (26) conditional on $G_t$, we obtain that

$$\mathbf{E}(H(\mathbf{w}_{t+1}) - H(\mathbf{w}_t) \mid G_t) = O(t^c).$$

In order to estimate the maximum change in $H(\mathbf{w}_t)$ we need to bound the change in $N_{0,t}$. It is clear that for each $t \in [n]$, $|N_{0,t} - N_{0,t-1}|$ can be as big as $[t]^c$ (extreme case). However, using Chernoff's inequality (see, for instance Theorem 2.1 in [20]), we show that $|N_{0,t} - N_{0,t-1}| < \log^2 n$ **wep**. $N_{0,t} - N_{0,t-1}$ is a random variable that can be expressed as a sum of independent random variables with expected value $O(1)$.

Indeed, let $X(t, j)$ be a random indicator variable for an event that vertex $v_t$ joins vertex of in-degree zero (at the end of step $t-1$) at sub-step $j$ of step $t$ ($t = 2, 3, \ldots, n$, $j = 1, 2, \ldots, [t^c]$). It is clear that

$$\mathbf{Pr}(X(t, j) = 1) = 1 - \mathbf{Pr}(X(t, j) = 0) = N_{0,t-1}/g_c(t-1) = O(t^{-c})$$

and

$$\mathbf{E}(N_{0,t} - N_{0,t-1} | N_{0,t-1}) \leq 1 + \sum_{j=1}^{[t^c]} \mathbf{Pr}(X(t,j) = 1) = O(1).$$

Since Lemma 6 requires a deterministic bound, we introduce the stopping time

$$T = \min\{t \geq 1 : (|N_{0,t} - N_{0,t-1}| \geq \log^2 n) \quad \vee \quad (t = n)\} .$$

A stopping time is any random variable $T$ with values in $\{0, 1, \dots\} \cup \{\infty\}$ such that it is determined whether $T = \hat{t}$ for any time $\hat{t}$ from knowledge of the process up to and including time $\hat{t}$.

From (26), with $i \wedge T$ denoting $\min\{i, T\}$ and noting that

$$\text{grad } H(\mathbf{w}_t) = \left((1+c)t^c N_{0,t} - t^{1+c}, t^{1+c}\right) ,$$

we have that

$$|H(\mathbf{w}_{(t+1)\wedge T}) - H(\mathbf{w}_{t\wedge T})| \leq O(t^{1+c}) + O(t^{1+c} \log^2 n) = O(t^{1+c} \log^2 n).$$

If $T < n$, then the value of $H(\mathbf{w}_{t\wedge T})$ remains the same from that point on, and the conditions of Lemma 6 hold.

Recalling that $\mathbf{w}_t = (t, N_{0,t})$ and choosing $X_t = H(t \wedge T, N_{0,t\wedge T})$, we apply Lemma 6, with $\alpha = n^{3/2+c} \log^3 n$, $\beta = O(t^c)$ and $\gamma_t = O(t^{1+c} \log^2 n)$, to the sequence $(H(\mathbf{w}_{t\wedge T}) : 1 \leq t \leq n)$, and also symmetrically to $(-H(\mathbf{w}_{t\wedge T}) : 1 \leq t \leq n)$. We obtain that **wep**

$$|H(\mathbf{w}_{t\wedge T}) - H(\mathbf{w}_0)| = O(n^{3/2+c} \log^3 n)$$

for $1 \leq t \leq n$. As $H(\mathbf{w}_0) = 0$, this implies from the definition (25) of the function $H$, that **wep**

$$N_{0,t\wedge T} = \frac{t \wedge T}{2+c} + O(n^{1/2} \log^3 n)$$

for $1 \leq t \leq n$.

To complete the proof we need to show that **wep**, $T = n$ but, as we already mentioned, this follows immediately from the Chernoff bound. □

**Corollary 5.** *Let* $i_f(n) = (n/(\log^8 n))^{\frac{1+c}{6+4c}}$. **Wep** *for every* $1 \leq t \leq n$ *and* $0 \leq i \leq i_f(n)$ *we have that*

$$N_{i,t} = c_i t + O(i n^{1/2} \log^3 n) .$$

*Proof.* We may repeat (recursively) the argument as in the proof of Theorem 4 for $N_{i,t}$ with $i \geq 1$. Since the expected change for $N_{i,t}$ is slightly different now (see (24)), we obtain our result by considering the following function:

$$H(x,y) = x^{(i+1)(1+c)} y - c_{i-1} \frac{i(1+c)}{1+(i+1)(1+c)} x^{1+(i+1)(1+c)}.$$

Using this function, we may show by similar arguments as in the case $i = 0$ that **wep**

$$N_{i,n} = c_i n + O(in^{1/2} \log^3 n).$$

Since

$$\begin{aligned} i_f\, n^{1/2} \log^3 n &= n^{\frac{4+3c}{6+4c}} \log^{\frac{10+4c}{6+4c}} n \\ &= o(n^{1-\frac{2+c}{6+4c}} \log^{\frac{10+4c}{6+4c}+1} n) \\ &= o\left(i_f^{-(1+\frac{1}{1+c})} n\right) \quad = \quad o(c_{i_f} n), \end{aligned}$$

we obtain concentration for all degrees $i$ up to $i_f$. □

# References


[1] A. Barabási and R. Albert, Emergence of scaling in random networks, *Science* **286** (1999), 509–512.

[2] B. Bollobás, O. Riordan, J. Spencer, and G. Tusnády, The degree sequence of a scale-free random graph process, *Random Structures and Algorithms* **18** (2001), 279–290.

[3] B. Bollobás and O. Riordan, Mathematical results on scale-free graphs, *Handbook of Graphs and Networks*, S. Bornholdt, H. Schuster (eds), Wiley-VCH, Berlin (2002).

[4] B. Bollobás and O. Riordan, The diameter of a scale-free random graph, *Combinatorica* **24** (2004), 5–34.

[5] A. Bonato, *A course on the web graph*, American Mathematical Society Graduate Studies in Mathematics 2008, Volume 89, (2008).

[6] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener, Graph structure in the web, *Proceedings of the 9th International World Wide Web Conference* (2000), 309-320,

[7] F.R.K. Chung and L. Lu, *Complex Graphs and Networks*, American Mathematical Society, Providence, Rhode Island, (2006).

[8] C. Cooper, Classifying special interest groups in web graphs, *Proceedings of Random 2002* 263–275.

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

[10] C. Cooper, A. Frieze, and J. Vera, Random vertex deletion in a scale free random graph, *Internet Mathematics* **1** (2004), 463–483.

[11] C. Cooper, The age specific degree distribution of web-graphs, *Combinatorics Probability and Computing* **15(5)** (2006), 637–661,.

[12] S. Dorogovtsev and J. Mendes, Accelerated growth of networks.

[13] S. Dorogovtsev and J. Mendes, Effect of the accelerating growth of communications networks on their structure, *Physical Review E* **63** (2001), 025101(R).

[14] S. Dorogovtsev and J. Mendes, Evolution of networks, *Adv. Phys.* **51** (2002), 1079.

[15] M. Faloutsos, P. Faloutsos, and C. Faloutsos, On Power-law Relationships of the Internet Topology, *SIGCOMM* (1999) 251–262.

[16] R. Kumar, P. Raghavan, S. Rajagopalan, D. Sivakumar, A. Tomkins, and E. Upfal, The web as a graph, *Proceedings of the 19th ACM Symposium on Principles of Database Systems* (2000), 1–10.

[17] R. Kumar, P. Raghavan, S. Rajagopalan, D. Sivakumar, A. Tomkins, and E. Upfal, Stochastic models for the web graph, *Proceedings of the 41th IEEE Symp. on Foundations of Computer Science* (2000), 57–65.

[18] J. Leskovec, J. Kleinberg, and C. Faloutsos, Graphs over time: densification laws, shrinking diameters and possible explanations, *KDD '05: Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining* (2005), 177–187.

[19] J. Leskovec, J. Kleinberg, and C. Faloutsos, Graph evolution: Densification and shrinking diameters, *ACM TKDD07, ACM Transactions on Knowledge Discovery from Data* (2007), 1(1).

[20] S. Janson, T. Luczak, and A. Rucinski, *Random Graphs*, Wiley, New York, (2000).

[21] J. Janssen and P. Pralat, Protean graphs with a variety of ranking schemes, *Theoretical Computer Science*, 22pp, in press.

[22] T. Luczak and P. Pralat, Protean graphs, *Internet Mathematics* **3** (2006), 21–40.

[23] J. Mattick and M. Gagen, Accelerating networks, *Science* **307** (2005), 856.

[24] M. Mitzenmacher, A brief history of generative models for power law and lognormal distributions, *Proc. of the 39th Annual Allerton Conf. on Communication, Control, and Computing* (2001), 182–191.

[25] B. Pittel, J. Spencer, and N. Wormald, Sudden emergence of a giant $k$-core in a random graph, *Journal of Combinatorial Theory, Series B* **67** (1996), 111–151.

[26] P. Pralat, A note on the diameter of protean graphs, *Discrete Mathematics* **308** (2008), 3399–3406.

[27] P. Pralat and N. Wormald, Growing protean graphs, *Internet Mathematics* **4** (2009), 1–16.

[28] P. Sen, Accelerated growth in outgoing links in evolving networks: deterministic versus stochastic picture, *Phys. Rev. E* **69** (2004), 46107.

[29] N. Wormald, The differential equation method for random graph processes and greedy algorithms, In: *Lectures on Approximation and Randomized Algorithms*, eds. M. Karoński and H. J. Prömel, PWN, Warsaw, (1999), 73–155.

[30] X. Yu, Z. Li, D. Zhang, F. Liang, X. Wang, and X. Wu, The topology of an accelerated growth network, *J. Phys A: Math Gen.* **39** (2006) 14343–14351.

[31] G. Yule, A mathematical theory of evolution based on the conclusions of Dr. J.C. Willis, *Philosophical Transactions of the Royal Society of London (Series B)* **213** (1924), 21–87.

# 4 Appendix

## 4.1 Degree distribution of the 1-process

We consider a 1-process $(G_1(t))_{t>s}$ in which, at each step $t = s+1, \dots$ a new vertex $v_t = t$ is added, together with an edge $e_t$ from $v_t$ to $G_1(t-1)$. The terminal vertex of edge $e_t$ is chosen preferentially using the following rule,

$$\mathbf{Pr}(v \text{ chosen at step } t+1) = \frac{d(v,t)}{2t}.$$

The initial subgraph $G(s)$ is an arbitrary graph with $s$ edges. For our applications $G(s)$ will be $G_f(\sigma)$ the existing subgraph of the $\boldsymbol{f}$-process immediately after the addition of vertex $\sigma$, and where $s = F(\sigma)$ is the size of the edge set of $G_f(\sigma)$.

The following approximation is used frequently in the subsequent proofs. If $\phi(x)$ is positive and monotone decreasing then

$$\sum_{j=1}^{n-1} \phi(j) = \int_1^n \phi(x)\, dx + a\phi(1) \qquad 0 \le a < 1.$$

**Lemma 7.** *For the 1-process $(G_1(t))_{t>s}$,*

*(i)*

$$\mathbf{E}\left(d(v,t) \mid d(v,s)\right) = d(v,s)\left(\frac{t}{s}\right)^{\frac{1}{2}}\left(1 + O\left(\tfrac{1}{s}\right)\right).$$

*(ii) Let $K$ be a positive constant. For $1 \leq v \leq s \leq t$*

$$\mathbf{Pr}\left(d(v,t) \geq d(v,s)\left(\frac{t}{s}\right)^{\frac{1}{2}} K \log^3 t\right) = O(t^{-K}).$$

This or similar results are given in, for example, [2, 8, 9, 10, 11].

The proof of the following lemma derives from [11].

**Lemma 8.** *Let $\log t < s < t$, $a \leq s$ and $a^2 = O(s)$. For the 1-process $(G_1(t))_{t>s}$,*

$$\begin{aligned}\mathbf{Pr}(d(v,t) =& a+l \mid d(v,s)=a) \qquad (27)\\ &= (1+\epsilon)\binom{a+l-1}{l}\left(\frac{s}{t}\right)^{\frac{a}{2}}\left(1-\left(\frac{s}{t}\right)^{\frac{1}{2}}\left(1+O\left(\tfrac{\theta(t)}{s}\right)\right)\right)^{l} + O(s^{-K+1})\end{aligned}$$

*where*

$$\epsilon = O\left(\frac{a^2}{s}\right) - l^2\left|O\left(\frac{\log(t/s)+\theta(t)/s}{\left(\sqrt{t}-\sqrt{s}\right)^2}\right)\right|, \qquad (28)$$

*$\theta(t) = \min(a+l, aK(\log t)^4)$, and $K$ is a positive constant.*

**Proof** Let $\boldsymbol{\tau} = (\tau_j : \ j = 1,\ldots,l)$ where $\tau_j$ is the step occurring before the degree of vertex $v$ changes from $a+j-1$ to $a+j$. Thus

$$\begin{aligned}\mathbf{Pr}(d(v,t) =& a+l \mid d(v,s)=a,\ \boldsymbol{\tau})\\ =&\left(1-\frac{a}{2s}\right)\cdots\left(1-\frac{a}{2(\tau_1-1)}\right)\frac{a}{2\tau_1}\left(1-\frac{a+1}{2(\tau_1+1)}\right)\cdots\\ &\cdots\frac{a+j-1}{2\tau_j}\left(1-\frac{a+j}{2(\tau_j+1)}\right)\cdots\left(1-\frac{a+j}{2(\tau_{j+1}-1)}\right)\frac{a+j}{2\tau_{j+1}}\cdots\left(1-\frac{a+l}{2(t-1)}\right)\\ =&\frac{a(a+1)\cdots(a+l-1)}{2^l\ \tau_1\cdots\tau_l}\Phi(\boldsymbol{\tau}),\end{aligned}$$

where $\Phi(\boldsymbol{\tau})$ is given by

$$\Phi(\boldsymbol{\tau}) = \left(1-\tfrac{a}{2s}\right)\cdots\left(1-\tfrac{a}{2(\tau_1-1)}\right)\left(1-\tfrac{a+1}{2(\tau_1+1)}\right)\cdots\left(1-\tfrac{a+j}{2(\tau_j+1)}\right)\cdots\left(1-\tfrac{a+j}{2(\tau_{j+1}-1)}\right)\cdots\left(1-\tfrac{a+l}{2(t-1)}\right).$$

From Lemma 9, proved below,

$$\Phi(\boldsymbol{\tau}) = \left(1+O\left(\tfrac{a^2}{s}\right)\right)\ \left(\frac{s}{t}\right)^{\frac{a}{2}}\prod_{j=1}^{l}\left(\frac{\tau_j}{t}\right)^{\frac{1}{2}}\left(1+O\left(\tfrac{a+j}{\tau_j}\right)\right). \qquad (29)$$

Note that $(a+j)/\tau_j \leq 1$ as required by Lemma 9, because $\tau_j = s+\sigma$ where $s \geq v-1+a$ and $\sigma \geq j$, as it takes at least $j$ steps to add $j$ edges to $v$.

Thus

$$\mathbf{Pr}(d(v,t)=a+l \mid d(v,s)=a) = \binom{a+l-1}{l}\; l!\; \sum_{\boldsymbol{\tau}} \frac{\Phi(\boldsymbol{\tau})}{2^l\; \tau_1 \ldots \tau_l}, \tag{30}$$

where

$$\sum_{\boldsymbol{\tau}} \frac{\Phi(\boldsymbol{\tau})}{2^l\; \tau_1 \cdots \tau_l} = \left(1+O\left(\tfrac{a^2}{s}\right)\right)\; \left(\frac{s}{t}\right)^{\frac{a}{2}} \sum_{\tau_1<\cdots<\tau_l} \prod_{j=1}^{l} \frac{1}{2\sqrt{\tau_j t}} \left(1+O\left(\frac{a+j}{\tau_j}\right)\right).$$

Provided $b_j \geq 0$ we have

$$(b_s+\cdots+b_T)^k - ({b_s}^2+\cdots+{b_T}^2)\binom{k}{2}(b_s+\cdots+b_T)^{k-2} \leq k! \sum_{i_1<\cdots<i_k} b_{i_1}\cdots b_{i_k} \leq (b_s+\cdots+b_T)^k.$$

Let

$$b_\tau = \frac{1}{2\sqrt{t\tau}}\left(1+O\left(\frac{\Delta(\tau)}{\tau}\right)\right),$$

where $\Delta(\tau)=\min(a+l, a(\tau/s)^{1/2}K(\log\tau)^3)$. The second term follows from Lemma 7(ii), as the probability of the event $\mathcal{E}(\tau)$ that the degree $j$ exceeds $a(\tau/s)^{1/2}K(\log\tau)^3$ at step $\tau$ is $O(\tau^{-K})$. In the estimate of $\Phi(\boldsymbol{\tau})$ in (29) we can replace $1+O\left(\frac{a+j}{\tau_j}\right)$ by $1+O\left(\frac{\Delta(\tau_j)}{\tau_j}\right)$ provided we add a term of $O(s^{-K+1})$ to our estimate of $\mathbf{Pr}(d(v,t)=a+l \mid d(v,s)=a)$ to account for the event $\mathcal{E}(\tau)$ at some step $s\leq\tau\leq t$. This explains the final term in (27).

We note that $\Delta(\tau)/\tau^{3/2}$ is monotone decreasing, and so

$$\sum_{\tau=s}^{t} \frac{\Delta(\tau)}{\tau^{3/2}} = O\left(\min\left(\frac{a+l}{\sqrt{s}}, \frac{aK(\log t)^4}{\sqrt{s}}\right)\right) = \frac{\theta(t)}{\sqrt{s}}.$$

Thus

$$\begin{aligned} b_s+\cdots+b_{t-1} &= \frac{O(1)}{2\sqrt{ts}} + \int_s^t \frac{1}{2\sqrt{tx}}\left(1+O\left(\tfrac{\Delta(x)}{x}\right)\right)dx \\ &= 1-\left(\frac{s}{t}\right)^{\frac{1}{2}}\left(1+O\left(\tfrac{\theta}{s}\right)\right), \end{aligned}$$

and

$$\begin{aligned} {b_s}^2+\cdots+{b_{t-1}}^2 &= \frac{O(1)}{4ts} + \int_s^t \frac{1}{4tx}\left(1+O\left(\tfrac{\Delta(x)}{x}\right)\right)dx \\ &= \frac{1}{4t}\left(\log\frac{t}{s}+O\left(\frac{\theta}{s}\right)\right). \end{aligned}$$

The expression for (27)-(28) follows from (30) using the estimates given above. □

**Lemma 9.** *Let $\boldsymbol{\tau} = (s, \tau_1, \ldots, \tau_l, t)$, let $\tau_0 = s$ and*

$$\Phi(\boldsymbol{\tau}) = \left(1 - \tfrac{a}{2s}\right) \cdots \left(1 - \tfrac{a}{2(\tau_1 - 1)}\right) \left(1 - \tfrac{a+1}{2(\tau_1+1)}\right) \cdots \left(1 - \tfrac{a+j}{2(\tau_j+1)}\right) \cdots \left(1 - \tfrac{a+j}{2(\tau_{j+1}-1)}\right) \cdots \left(1 - \tfrac{a+l}{2(t-1)}\right).$$

*If $(a+j)/\tau_j \le 1$ for $j = 0, 1, \ldots, l$, and $a^2 = O(s)$ then*

$$\Phi(\boldsymbol{\tau}) = \left(1 + O\left(\tfrac{a^2}{s}\right)\right) \left(\frac{s}{t}\right)^{\frac{a}{2}} \prod_{j=1}^{l} \left(\frac{\tau_j}{t}\right)^{\frac{1}{2}} \left(1 + O\left(\tfrac{a+j}{\tau_j}\right)\right). \tag{31}$$

**Proof** We note that for $0 \le x < 1$

$$\begin{aligned} \log(1-x) &= -x - \int_0^x \frac{y}{1-y}\, dy \\ &= -x - \frac{x^2}{2} R_x, \end{aligned}$$

where $1 \le R_x \le 1/(1-x)$. Thus

$$-\sum_{x \in \chi(\boldsymbol{\tau})} x - \sum_{x \in \chi(\boldsymbol{\tau})} x^2 \le \log \Phi(\tau) \le -\sum_{x \in \chi(\boldsymbol{\tau})} x - \sum_{x \in \chi(\boldsymbol{\tau})} \frac{x^2}{2}, \tag{32}$$

where

$$\chi(\boldsymbol{\tau}) = \left\{ \frac{a}{2s}, \ldots, \frac{a}{2(\tau_1 - 1)}, \frac{a+1}{2(\tau_1+1)}, \ldots, \frac{a+j}{2(\tau_j+1)}, \ldots, \frac{a+j}{2(\tau_{j+1}-1)}, \ldots, \frac{a+l}{2(t-1)} \right\}.$$

The value of $\max_{x \in \chi(\boldsymbol{\tau})} \{1/(1-x)\} \le 2$ because, by assumption $(a+j)/\tau_j \le 1$ and from then on, the sequence $\frac{a+j}{\tau_j+1}, \ldots, \frac{a+j}{\tau_{j+1}-1}$ is monotone decreasing.

Now

$$\begin{aligned} &\frac{a}{s} + \cdots + \frac{a}{\tau_1 - 1} + \frac{a+1}{\tau_1} + \frac{a+1}{\tau_1+1} + \cdots + \frac{a+j-1}{\tau_{j-1}-1} + \frac{a+j}{\tau_j} + \frac{a+j}{\tau_j+1} + \cdots + \frac{a+l}{t-1} \\ &\quad = a \sum_{T=s}^{t-1} \frac{1}{T} + \sum_{T=\tau_1}^{t-1} \frac{1}{T} + \cdots + \sum_{T=\tau_l}^{t-1} \frac{1}{T} \\ &\quad = a \log \frac{t}{s} + \log \frac{t}{\tau_1} + \cdots + \log \frac{t}{\tau_j} + \cdots + \log \frac{t}{\tau_l} \\ &\qquad + O\left(\frac{a}{s}\right) + O\left(\frac{1}{\tau_1}\right) + \cdots + O\left(\frac{1}{\tau_j}\right) + \cdots + O\left(\frac{1}{\tau_l}\right). \end{aligned}$$

Thus, as $(a+j)/\tau_j \le 1$,

$$\exp\left(-\sum_{x \in \chi(\boldsymbol{\tau})} x\right) = \left(\frac{s}{t}\right)^{\frac{a}{2}} \left(1 + O\left(\tfrac{a}{s}\right)\right) \prod_{j=1}^{l} \left(\frac{\tau_j}{t}\right)^{\frac{1}{2}} \left(1 + O\left(\tfrac{a+j}{\tau_j}\right)\right). \tag{33}$$

We next consider the term $\sum x^2$. For $j = 1, \ldots, l$ we have $(a+j)^2 - (a+j-1)^2 = 2(a+j) - 1$ and thus

$$\begin{aligned}
&\left(\tfrac{a}{s}\right)^2 + \cdots + \left(\tfrac{a}{\tau_1 - 1}\right)^2 + \left(\tfrac{a+1}{\tau_1}\right)^2 + \left(\tfrac{a+1}{\tau_1+1}\right)^2 + \cdots + \left(\tfrac{a+j-1}{\tau_j - 1}\right)^2 + \left(\tfrac{a+j}{\tau_j}\right)^2 + \left(\tfrac{a+j}{\tau_j+1}\right)^2 + \cdots + \left(\tfrac{a+l}{t-1}\right)^2 \\
&\quad = a^2 \sum_{T=s}^{t-1} \tfrac{1}{T^2} + (2(a+1)-1) \sum_{T=\tau_1}^{t-1} \tfrac{1}{T^2} + \cdots + (2(a+j)-1) \sum_{T=\tau_j}^{t-1} \tfrac{1}{T^2} + \cdots + (2(a+l)-1) \sum_{T=\tau_l}^{t-1} \tfrac{1}{T^2} \\
&\quad = a^2 \left(\tfrac{1}{s} - \tfrac{1}{t}\right) + (2(a+1)-1)\left(\tfrac{1}{\tau_1} - \tfrac{1}{t}\right) + \cdots + (2(a+l)-1)\left(\tfrac{1}{\tau_l} - \tfrac{1}{t}\right) + c\left(\left(\tfrac{a}{s}\right)^2 + \cdots + \left(\tfrac{a+l}{\tau_l^2}\right)\right) \\
&\quad = O\left(\frac{a^2}{s}\right) + O\left(\frac{a+1}{\tau_1}\right) \cdots + O\left(\frac{a+l}{\tau_l}\right).
\end{aligned}$$

Thus for $c > 0$ constant,

$$\exp\left(-c \sum_{x \in \chi(\boldsymbol{\tau})} x^2\right) = \left(1 + O\left(\frac{a^2}{s}\right)\right) \prod_{j=1,\ldots,l} \left(1 + O\left(\frac{a+j}{\tau_j}\right)\right). \tag{34}$$

The expression (31) follows from (32), (33), (34) and the definition of $\Phi(\boldsymbol{\tau})$. □

## 4.2 Proof of Lemma 10

**Lemma 10.** *Let $f(t) = [t^c]$ and let $F(t) = \sum_{s=1}^{t} f(s)$ then $F(t) = \frac{t^{1+c}}{1+c}(1 + O(t^{-c}))$.*

**Proof**

$$\begin{aligned}
\sum_{i=1}^{t} [i^c] &= \sum_{i=1}^{t} (i^c + O(1)) \\
&= \sum_{i=1}^{t} i^c + O(t) \\
&= \frac{t^{1+c}}{1+c} + O(1) + O(t) \\
&= \frac{t^{1+c}}{1+c}(1 + O(t^{-c}))\,.
\end{aligned}$$

□