Proc. Indian Acad. Sci. (Math. Sci.) Vol. 118, No. 3, August 2008, pp. 473–494.

Growth of preferential attachment random graphs via continuous-time branching processes

KRISHNA B ATHREYA[1], ARKA P GHOSH[2] and SUNDER SETHURAMAN[3]

[1]Departments of Mathematics and Statistics, Iowa State University, Ames, IA 50011, USA
[2]Department of Statistics, Iowa State University, Ames, IA 50011, USA
[3]Department of Mathematics, Iowa State University, Ames, IA 50011, USA
E-mail: kba@iastate.edu; apghosh@iastate.edu; sethuram@iastate.edu

MS received 10 April 2007; revised 5 November 2007

Abstract. Some growth asymptotics of a version of 'preferential attachment' random graphs are studied through an embedding into a continuous-time branching scheme. These results complement and extend previous work in the literature.

Keywords. Branching processes; preferential attachment; embedding; random graph; scale-free.

1. Introduction and results

Preferential attachment processes have a long history dating back at least to Yule [20] and Simon [19] (cf. [12] for an interesting survey). Recently, Barabasi and Albert [7] proposed a random graph version of these processes as a model for several real-world networks, such as the internet and various communication structures, on which there has been much renewed study (see [1], [9], [11], [15] and references therein). To summarize, the basic idea is that, starting from a small number of nodes, or vertices, one builds an evolving graph by 'preferential attachment', that is by attaching new vertices to existing nodes with probabilities proportional to their 'weight'. When the weights are increasing functions of the 'connectivity', already well connected vertices tend to become even more connected as time progresses, and so, these graphs can be viewed as types of 'reinforcement' schemes (cf. [17]). A key point, which makes these graph models 'practical', is that, when the weights are linear, the long term degree proportions are often in the form of a 'power-law' distribution whose exponent, by varying parameters, can be matched to empirical network data.

The purpose of this note is to understand a general form of the linear weights model with certain random 'edge additions' (described below in subsection 1.1) in terms of an embedding in continuous-time branching processes which allows for extensions of law of large numbers and maximal degree growth asymptotics, first approached by difference equations and martingale methods, in [8], [10], [13], [14].

We remark that some connections to branching and continuous-time Markov processes have also been studied in two recent papers. In [18], certain laws of large numbers for the degree distributions of the whole tree, and as seen from a randomly selected vertex are proved for a class of 'non-explosive' weights including linear weights. In [16], asymptotic degree distributions under super-linear weights are considered. In this context, the embedding given here is of a different character with respect to Markov branching systems with immigration, and the contributions made concentrate on detailed investigations of a generalized linear weights degree landscape.

1.1 *Model*

Start with two vertices v_1, v_2 and one edge joining them – denote this graph as G_0. To obtain G_1, create a new vertex v_3, and join it a random number X_1 times to one of v_1 and v_2 of G_0 with equal probability. For any finite graph $G_n = \{v_1, v_2, \ldots, v_{n+2}\}$, let the degree of each vertex be defined as the number of edges emanating from that vertex, and the degree of the jth vertex, $v_j \in G_n$ be denoted by $d_j(n)$ for $j = 1, \ldots, n+2$ and $n \geq 0$ (note that in our notation, G_n has $n+2$ vertices at step $n \geq 0$). After $n+2$ vertices are created, to obtain G_{n+1} from G_n, create an $(n+3)$rd vertex v_{n+3}, and connect it a random number X_{n+1} times to one of the $n+2$ existing vertices $v_1, \ldots, v_{n+2}$ with probability

$$\frac{d_i(n) + \beta}{\sum_{j=1}^{n+2}(d_j(n) + \beta)} \tag{1.1}$$

of being joined to vertex v_i for $1 \leq i \leq n+2$ where $\beta \geq 0$ is a parameter. We will also assume throughout that $\{X_i\}_{i\geq 1}$ are independently and identically distributed (i.i.d.) positive integer valued random variables with distribution $\{p_j\}_{j\geq 1}$ with finite mean. The 'weight' then of the ith vertex at the nth step is proportional to $d_i(n) + \beta$, and linear in the degree.

We remark that this basic model creates a growing graph (which is a tree when $X_i \equiv 1$) with undirected edges. As the referee remarked, one can alternatively think of this model as a tree with each edge having a 'count factor weight' which corresponds to the number of times a connection was made between the two associated nodes. Our model includes the 'one-edge' case of the original Barabasi–Albert process, made precise in [8], by setting $X_i \equiv 1$ and $\beta = 0$, as well as the '$\beta \geq 0$' scheme considered in [13] and [14], by taking $X_i \equiv 1$. Also, the '$\beta \geq 1$' linear case considered in [18] is recovered by taking $X_i \equiv 1$.

The aspect of adding a random number of edges $\{X_i\}_{i\geq 1}$ at each step to vertices chosen preferentially seems to be a new twist on the standard model which can be interpreted in various ways. The results, as will be seen, involve the mean number $\sum jp_j$ of added edges, indicating a sort of 'averaging' effect in the asymptotics.

We also note, in the case $\beta = 0$, a more general graph process, allowing cycles and self-loops, can be formed in terms of the 'tree' model above (cf. [8] and Ch. 4 [11]) where several sets of edges are added to possibly different existing vertices at each step preferentially. Namely, let $\{L_i\}_{i\geq 1}$ be independent and identically distributed positive integer valued random variables with distribution $\{q_j\}_{j\geq 1}$ with finite mean, and let $\bar{L}_i = \sum_{k=1}^{i} L_k$ for $i \geq 1$. As before, initially, we start with two vertices, $v_1^{(L)}$ and $v_2^{(L)}$ and one edge between them. Run the 'tree' model now to obtain vertices $\{w_i\}_{i\geq 3}$ and identify sets

$$\{w_3, \ldots, w_{2+L_1}\},\ \{w_{3+L_1}, \ldots, w_{2+\bar{L}_2}\},\ \ldots,\ \{w_{3+\bar{L}_{k-1}}, \ldots, w_{2+\bar{L}_k}\},\ \ldots$$

as vertices $v_3^{(L)}, v_4^{(L)}, \ldots, v_{k+2}^{(L)}, \ldots$. One interprets the sequence of graphs $G_n^{(L)} = \{v_1^{(L)}, \ldots, v_{n+2}^{(L)}\}$ for $n \geq 0$ as a more general graph process where L_i sets of edges are added at the ith step preferentially for $i \geq 1$. This model has some overlap with the very general model given in [10] where vertices can be selected preferentially or at random;

in [10], when only 'new' vertices are selected preferentially, their assumptions become $X_i \equiv 1$ and $\{q_j\}_{j\geq 1}$ has bounded support (as well as $\beta = 0$).

For the remainder of the article, we will focus, for simplicity, on the basic 'tree' model given through (1.1), although extensions to the other case ($L_i \geq 1$, $\beta = 0$) under various conditions on $\{q_j\}_{j\geq 1}$ are possible.

1.2 *Results*

For $n \geq 0$ and $j \geq 1$, let

$$R_j(n) = \sum_{i=1}^{n+2} I(d_i(n) = j)$$

be the number of vertices in G_n with degree j. Also, define the maximum degree in G_n by

$$M_n = \max_{1\leq i\leq n+2} d_i(n).$$

In addition, denote the mean

$$m = \sum_{j\geq 1} j p_j.$$

Our first result is on the growth rates of individual degree sequences $\{d_i(n)\}_{n\geq 0}$ and the maximal one M_n. It also describes the asymptotic behavior of the index where the maximal degree is attained.

Theorem 1.1. *Suppose* $\sum (j \log j) p_j < \infty$, *and let* $\theta = m/(2m + \beta)$.

(i) *For each* $i \geq 1$, *there exists a random variable* γ_i *on* $(0, \infty)$ *such that*

$$\lim_{n\to\infty} \frac{d_i(n)}{n^\theta} = \gamma_i \text{ exists a.s.}.$$

(ii) *Further, there exist positive absolutely continuous independent random variables* $\{\xi_i\}_{i\geq 1}$ *with* $E[\xi_i] < \infty$, *and a random variable* V *on* $(0, \infty)$ *such that* $\gamma_i = \xi_i V$ *for* $i \geq 1$. *In particular, for all* $i, j \geq 1$,

$$\lim_{n\to\infty} \frac{d_i(n)}{d_j(n)} = \frac{\xi_i}{\xi_j} \text{ exists a.s.}.$$

(iii) *Also, when* $\sum j^r p_j < \infty$ *for an* $r > \theta^{-1} = 2 + \beta/m$, *then*

$$\lim_{n\to\infty} \frac{M_n}{n^\theta} = \max_{i\geq 1} \gamma_i < \infty \quad a.s.$$

(iv) *Moreover, in this case* ($\sum j^r p_j < \infty$ *for* $r > \theta^{-1}$), *if* I_n *is the index where*

$$d_{I_n}(n) = M_n,$$

then $\lim_{n\to\infty} I_n = I < \infty$ *exists a.s.*

Remark 1.1. Note that Theorem 1.1 asserts that the individual degrees $d_i(n)$ and the maximal degree M_n grow at the same rate n^θ, and also the vertex with maximal degree freezes eventually, that is it does not change for large n.

The next result is on the convergence of the empirical distribution of the degrees $\{d_i(n)\colon 1 \le i \le n+2\}$. Let $\{D(y)\colon y \ge 0\}$ be a Markov branching process with exponential (1) lifetime distribution, offspring distribution $\{p'_j = p_{j-1}\}_{j\ge 2}$, immigration rate $\beta \ge 0$, immigration size distribution $\{p_j\}_{j\ge 1}$, and initial value $D(0)$ distributed according to $\{p_j\}_{j\ge 1}$ (see Definition 2.2 in §2 for the full statement). Also, for $y \ge 0$ and $j \ge 1$, let

$$p_j(y) = P(D(y) = j). \tag{1.2}$$

Theorem 1.2. *Suppose* $\sum (j \log j) p_j < \infty$, *and define the probability distribution* $\{\pi_j\}_{j\ge 1}$ *by*

$$\pi_j = (2m+\beta) \int_0^\infty p_j(y) \mathrm{e}^{-(2m+\beta)y} \mathrm{d}y.$$

Then, for $j \ge 1$, *we have*

$$\frac{R_j(n)}{n} \to \pi_j, \quad \textit{in probability, as } n \to \infty.$$

Remark 1.2. As a direct consequence, for bounded functions $f\colon \mathbb{N} \to \mathbb{R}$,

$$\frac{1}{n} \sum_{j=1}^\infty f(j) R_j(n) \to \sum_{j=1}^\infty f(j) \pi_j, \quad \text{in probability, as } n \to \infty.$$

We now consider the 'power-law' behavior of the limit degree distribution $\{\pi_j\}_{j\ge 1}$.

Theorem 1.3. *Suppose* $\sum_{j\ge 1} j^{2+\beta/m} p_j < \infty$. *Then, for* $s \ge 0$, *we have*

$$\sum_{j\ge 1} j^s \pi_j < \infty \quad \text{if and only if} \quad s < 2 + \beta/m.$$

Remark 1.3. Heuristically, the last result suggests $\pi_j = O(j^{-[3+\beta/m]})$ as $j \uparrow \infty$. In the case $X_i \equiv x_0$ for $x_0 \ge 1$, (1.2) can be explicitly evaluated (Proposition 3.2) to get $\pi_j = O(j^{-[3+\beta/x_0]})$ when j is a multiple of x_0.

The next section discusses the embedding method and auxiliary estimates. In the third section, the proofs of Theorems 1.1, 1.2, and 1.3 are given.

2. Embedding and some estimates

We start with the following definitions, and then describe in the following subsections the embedding and various estimates.

DEFINITION 2.1

A Markov branching process with offspring distribution $\{p'_j\}_{j\ge 0}$ and lifetime parameter $0 < \lambda < \infty$ is a continuous-time Markov chain $\{Z(t)\colon t \ge 0\}$ with state space $\mathbb{S} = \{0, 1, 2, \dots\}$ and waiting time parameters $\lambda_i \equiv i\lambda$ for $i \ge 0$, and jump probabilities $p(i, j) = p'_{j-i+1}$ for $j \ge i-1 \ge 0$ and $i \ge 1$, $p(0, 0) = 1$, and $p(i, j) = 0$ otherwise (cf. Chapter III of [5]).

DEFINITION 2.2

A Markov branching process with offspring distribution $\{p'_j\}_{j\geq 0}$ and lifetime parameter $0 < \lambda < \infty$, immigration parameter $0 \leq \beta < \infty$ and immigration size distribution $\{p_j\}_{j\geq 0}$ is a continuous-time Markov chain $\{D(t): t \geq 0\}$ such that $D(t) = Z(t)$ as in Definition 2.1 when $\beta = 0$, and when $\beta > 0$,

$$D(t) = \sum_{i=0}^{\infty} Z_i(t - T_i)\, I(T_i \leq t),$$

where $\{T_i\}_{i\geq 1}$ are the jump times of a Poisson process $\{N(t): t \geq 0\}$ with parameter β, $T_0 = 0$, and $\{Z_i(\cdot)\}_{i\geq 0}$ are independent copies of $\{Z(t): t \geq 0\}$ as in Definition 2.1, with $Z_0(0) = D(0)$ and $Z_i(0)$ distributed according to $\{p_j\}_{j\geq 0}$ for $i \geq 1$ and also independent of $\{N(t): t \geq 0\}$.

Remark 2.1. The condition that the mean number of offspring is finite, $\sum jp'_j < \infty$, is sufficient to ensure that $P(Z(t) < \infty) = 1$ and $P(D(t) < \infty) = 1$ for all $t \geq 0$, that is no explosion occurs in finite time (cf. p. 105 of [5]).

2.1 *Embedding process*

We now construct a Markov branching process through which a certain ‘embedding’ is accomplished. Recall $\{p_j\}_{j\geq 1}$ is a probability distribution on the positive integers. Consider an infinite sequence of independent processes $\{D_i(t): t \geq 0\}_{i\geq 1}$ where each $\{D_i(t): t \geq 0\}$ is a Markov branching process with immigration as in Definition 2.2, corresponding to exponential ($\lambda = 1$) lifetimes, offspring distribution $\{p'_j = p_{j-1}\}_{j\geq 2}$ (with $p'_0 = p'_1 = 0$), and immigration parameter $\beta \geq 0$ and immigration size distribution $\{p_j\}_{j\geq 1}$. The distributions of $\{D_i(0)\}_{i\geq 1}$ will be specified later.

Now, define recursively the following processes:

- At time 0, the first two processes $\{D_i(t): t \geq 0\}_{i=1,2}$ are started with $D_1(0) = D_2(0) = 1$. Let $\tau_{-1} = \tau_0 = 0$, and τ_1 be the first time an ‘event’ occurs in any one of the two processes.
- Now add a random X'_1 of new particles to the process in which the event occurred: (i) If the event is ‘immigration’, then $P(X'_1 = j) = p_j$ for $j \geq 1$. (ii) If the event is the death of a particle, then $P(X'_1 = j) = p_{j-1}$ for $j \geq 2$. Denote X_1 as the net addition; then $P(X_1 = j) = p_j$ for $j \geq 1$.
- At time τ_1, start a new Markov branching process $\{D_3(t): t \geq 0\}$ with $D_3(0) = X_1$.
- Let τ_2 be the first time after τ_1 that an event occurs in any of the processes $\{D_i(t): t \geq \tau_1\}_{i=1,2}$ and $\{D_3(t - \tau_1): t \geq \tau_1\}$. Add a random (net) number X_2, following the scheme above for X_1, of particles with distribution $\{p_j\}_{j\geq 1}$ to the process in which the event occurred. At time τ_2, start a new Markov branching process $\{D_4(t): t \geq 0\}$ with $D_4(0) = X_2$.
- Suppose that n processes have been started with the first two at $\tau_0 = 0$, the third at time τ_1, the fourth at time τ_2, and so on with the nth at time τ_{n-2}, and with (net) additions $X_1, X_2, \ldots, X_{n-2}$ at these times. Now, let τ_{n-1} be the first time after τ_{n-2} that an event occurs in one of the processes $\{D_i(t): t \geq 0\}_{i=1,2}$, $\{D_3(t - \tau_1): t \geq \tau_1\}$, $\{D_4(t - \tau_2): t \geq \tau_2\}, \ldots, \{D_n(t - \tau_{n-2}): t \geq \tau_{n-2}\}$. Add a (net) random number X_{n-1} of new particles with distribution $\{p_j\}_{j\geq 1}$ (following the scheme above) to the process

in which the event happened. Now start the $(n+1)$st process $\{D_{n+1}(t)\colon t \geq 0\}$ with $D_{n+1}(0) = X_{n-1}$.

Theorem 2.1 (Embedding theorem). *Recall the degree sequence $d_j(n)$ defined for the graphs $\{G_n\}$ near* (1.1). *For $n \geq 0$, let*

$$Z_n \equiv \{D_j(\tau_n - \tau_{j-2})\colon 1 \leq j \leq n+2\}, \quad \text{and}$$

$$\tilde{Z}_n \equiv \{d_j(n)\colon 1 \leq j \leq n+2\}.$$

Then, the two collections $\{Z_n\}_{n\geq 0}$ and $\{\tilde{Z}_n\}_{n\geq 0}$ have the same distribution.

Proof. First note that both sequences $\{Z_n\}_{n\geq 0}$ and $\{\tilde{Z}_n\}_{n\geq 0}$ have the Markov property and $Z_0 = \tilde{Z}_0 = \{1, 1\}$. Next, it will be shown below that the transition probability mechanism from Z_n to Z_{n+1} is the same as that from $\tilde{Z}_n$ to $\tilde{Z}_{n+1}$. To see this note that, at time 0, both $D_1(\cdot)$ and $D_2(\cdot)$ are 'turned on', and, at time τ_1, $D_3(\cdot)$ is 'turned on', and more generally, at τ_j, $D_{j+2}(\cdot)$ is 'turned on'. At time τ_{n+1}, the 'event' could be in $D_i(\cdot)$ for $1 \leq i \leq n+2$ with probability

$$\frac{D_i(\tau_n - \tau_{i-2}) + \beta}{\sum_{j=1}^{n+2}(D_j(\tau_n - \tau_{j-2}) + \beta)}$$

in view of the fact that the minimum of $n+2$ independent exponential random variables $\{\eta_i\}_{1\leq i\leq n+2}$ with means $\{\mu_i^{-1}\}_{1\leq i\leq n+2}$ is an exponential random variable with mean $\left(\sum_{i=1}^{n+2}\mu_i\right)^{-1}$, and coincides with η_i with probability $\mu_i\left(\sum_{i=1}^{n+2}\mu_i\right)^{-1}$ for $1 \leq i \leq n+2$. At that event time τ_{n+1}, $D_{n+3}(\cdot)$ is 'turned on', that is a new $(n+3)$rd vertex is created and connected to the chosen vertex v_i with X_{n+1} edges between them. Hence both the degree of the new vertex and increment in the degree of the chosen vertex (among the existing ones) is X_{n+1}. This shows that the conditional distribution of Z_{n+1} given $Z_n = z$ is the same as that of $\tilde{Z}_{n+1}$ given $\tilde{Z}_n = z$. $\square$

2.2 *Estimates on branching times*

We now develop some properties of the branching times $\{\tau_n\}_{n\geq 1}$, used in the embedding in subsection 2.1, which have some analogy to results in section III.9 of [5] (cf. [4]). Define $S_0 = 2 + 2\beta$ and, for $n \geq 1$,

$$S_n = 2 + 2\beta + \sum_{j=1}^{n} 2X_j + n\beta,$$

where as before $X_1, \ldots, X_n$ are the net independent additions, distributed according to $\{p_j\}_{j\geq 1}$, at event times $\tau_1, \ldots, \tau_n$. Define also, for $n \geq 1$, $\mathcal{F}_n$ as the σ-algebra,

$$\mathcal{F}_n = \sigma\{\{D_j(t - \tau_{j-2})\colon \tau_{j-2} \leq t \leq \tau_n\}_{1\leq j\leq n+2}, \{X_k\}_{1\leq k\leq n}\}. \tag{2.1}$$

PROPOSITION 2.1

The random variable τ_1 is exponential with mean S_0^{-1}. Also, for $n \geq 1$, conditioned on $\mathcal{F}_n$, the random variable $\tau_{n+1} - \tau_n$ is exponential with mean S_n^{-1}.

Proof. Follows from the construction of the $\{\tau_i\}_{i\geq 1}$. □

PROPOSITION 2.2

Suppose $m = \sum jp_j < \infty$. *Then,*

$$\left\{\tau_n - \sum_{j=1}^{n} \frac{1}{S_{j-1}}; \mathcal{F}_n\right\}_{n\geq 1}$$

is an L^2 *bounded martingale and hence converges a.s. as well as in* L^2.

Proof. The martingale property follows from the fact

$$\tau_n = \sum_{j=1}^{n}(\tau_j - \tau_{j-1})$$

and Proposition 2.1.

Next, with $\phi(a) = E[\mathrm{e}^{-aX_1}]$ for $a \geq 0$, we have the uniform bound in $n \geq 1$,

$$\begin{aligned}
\mathrm{Var}\left(\tau_n - \sum_{j=1}^{n}\frac{1}{S_{j-1}}\right) &= \mathrm{Var}\left(\sum_{j=1}^{n}\left(\tau_j - \tau_{j-1} - \frac{1}{S_{j-1}}\right)\right)\\
&= \sum_{j=1}^{n}\mathrm{Var}\left(\tau_j - \tau_{j-1} - \frac{1}{S_{j-1}}\right) \quad \text{(by martingale property)}\\
&= \sum_{j=1}^{n} E\left[\frac{1}{S_{j-1}^2}\right]\\
&= \sum_{j=1}^{n} E\left[\int_0^\infty x\mathrm{e}^{-S_{j-1}x}\mathrm{d}x\right]\\
&\leq \sum_{j=1}^{\infty}\int_0^\infty (\phi(2x)\mathrm{e}^{-x\beta})^{j-1}x\mathrm{e}^{-(2+2\beta)x}\mathrm{d}x\\
&\leq \int_0^\infty \frac{x\mathrm{e}^{-(2+2\beta)x}\mathrm{d}x}{1-\phi(2x)\mathrm{e}^{-x\beta}} < \infty,
\end{aligned}$$

where the finiteness in the last bound follows from the fact that

$$\lim_{x\downarrow 0}\frac{x}{1-\phi(2x)\mathrm{e}^{-x\beta}} = \frac{1}{2m+\beta} < \infty.$$

The a.s. and L^2-convergence follows from Doob's martingale convergence theorem (c.f. Theorem 13.3.9 of [6]). □

PROPOSITION 2.3

Suppose $\sum(j\log j)p_j < \infty$, *and recall* $m = \sum jp_j$. *Let also* $\alpha = (2m+\beta)^{-1}$. *Then, there exists a real random variable* Y *so that a.s.,*

$$\lim_{n\to\infty}\tau_n - \sum_{j=1}^{n}\frac{\alpha}{j} = Y.$$

Proof. By Proposition 2.2, there is a real random variable Y' such that

$$\tau_n - \sum_{j=1}^{n} \frac{1}{S_{j-1}} \;\to\; Y' \quad \text{a.s.}$$

To complete the proof, we note, as $E[X_1 \log X_1] = \sum (j \log j) p_j < \infty$, by Theorem III.9.4 of [5] on reciprocal sums, that $\sum_{j=1}^{\infty}(1/S_j - \alpha/j)$ converges a.s. □

COROLLARY 2.1

Suppose $m = \sum j p_j < \infty$. *Then,*

(i) $\tau_n \uparrow \infty$ *a.s., as* $n \to \infty$.
Also, when $\sum (j \log j) p_j < \infty$, *we have, with* $\alpha = (2m + \beta)^{-1}$, *that*
(ii) $\tau_n - \alpha \log n \to \tilde{Y} := Y - \alpha\gamma$ *a.s., as* $n \to \infty$, *where* γ *is Euler's constant.*
(iii) *For each fixed* $\epsilon > 0$, $\sup_{n\epsilon \le k \le n}(\tau_n - \tau_k - \alpha \log(n/k)) \to 0$ *a.s., as* $n \to \infty$.

Proof. The first claim follows from Proposition 2.2 and the fact that $\sum 1/S_j = \infty$, since by strong law of large numbers, we have a.s. that $S_j \le j(1/\alpha + 1)$ for large j. The last two claims, as $\sum_{j=1}^{n} 1/j - \log n \to \gamma$, Euler's constant, are direct consequences of Proposition 2.3. □

2.3 *Estimates on Markov branching processes*

As in Definition 2.2, let $\{D(t)\colon t \ge 0\}$ be a Markov branching process with offspring distribution $\{p'_j = p_{j-1}\}_{j \ge 2}$, lifetime $\lambda = 1$ and immigration $\beta \ge 0$ parameters, and immigration distribution $\{p_j\}_{j\ge 1}$.

PROPOSITION 2.4

Suppose $\sum (j \log j) p_j < \infty$, *and* $D(0) \ge 1$, $E[D(0)] < \infty$. *Recall* $m = \sum j p_j$. *Then,*

$$\lim_{t\to\infty} D(t)\mathrm{e}^{-mt} = \zeta$$

converges a.s. and in L^1, *and* ζ *is supported on* $(0, \infty)$ *and has an absolutely continuous distribution.*

Proof. Let $\beta > 0$; when $\beta = 0$ the argument is easier and a special case of the following development. Let $0 = T_0 < T_1 < \cdots < T_n < \cdots$ be the times at which immigration occurs, and let $\eta_1, \eta_2, \ldots$ be the respective number of immigrating individuals (distributed according to $\{p_j\}_{j\ge 1}$). From Definition 2.2, $D(t)$ has representation

$$D(t) = \sum_{i=0}^{\infty} Z_i(t - T_i) I(T_i \le t), \tag{2.2}$$

where $\{Z_i(t)\colon t \ge 0\}_{i\ge 0}$ are independent Markov branching processes with offspring distribution $\{p'_j = p_{j-1}\}_{j\ge 2}$, with exponential ($\lambda = 1$) lifetime distributions, with no immigration, with $Z_0(0) = D(0)$ and $Z_i(0) = \eta_i$ for $i \ge 1$, and also independent of

$\{T_i\}_{i\geq 0}$. Under the hypothesis $\sum(j\log j)p_j < \infty$, it is known (Theorem III.7.2 of [5]; with rate $\lambda\big(\sum_{j\geq 2} jp'_j - 1\big) = \sum_{j\geq 1}(j+1)p_j - 1 = m$), for $i \geq 0$, that

$$\lim_{t\to\infty} Z_i(t)\mathrm{e}^{-mt} = W_i \tag{2.3}$$

converges in $(0,\infty)$ a.s. and W_i has a continuous distribution on $(0,\infty)$. Also under the hypothesis that $\sum(j\log j)p_j < \infty$, it can be shown (Proposition 2.5) that

$$E[\widetilde{W}_i] < \infty, \quad \text{where } \widetilde{W}_i = \sup_{t\geq 0} Z_i(t)\mathrm{e}^{-mt}, \tag{2.4}$$

and hence convergence in (2.3) holds in L^1 as well.

Since $\{T_i\}_{i\geq 0}$ is a Poisson process with rate β, and independent of $\{Z_i(t)\}_{t\geq 0}$,

$$E\left[\sum_{i=0}^{\infty} \widetilde{W}_i \mathrm{e}^{-mT_i}\right] \leq E[\widetilde{W}_1]\left(E[D(0)] + \sum_{i=1}^{\infty}\left(\frac{\beta}{m+\beta}\right)^i\right) < \infty, \tag{2.5}$$

yielding

$$\sum_{i=0}^{\infty} \widetilde{W}_i \mathrm{e}^{-mT_i} < \infty \quad \text{a.s..} \tag{2.6}$$

Hence, noting (2.3), (2.4) and (2.6), by dominated convergence,

$$\begin{aligned}\lim_{t\to\infty} D(t)\mathrm{e}^{-mt} &= \sum_{i=0}^{\infty} \lim_{t\to\infty}[Z_i(t-T_i)I(T_i\leq t)\mathrm{e}^{-m(t-T_i)}]\mathrm{e}^{-mT_i} \\ &= \sum_{i=0}^{\infty} W_i \mathrm{e}^{-mT_i} := \zeta \end{aligned}\tag{2.7}$$

converges in $(0,\infty)$ a.s.. Also,

$$\sup_{t\geq 0} D(t)\mathrm{e}^{-mt} \leq \sum_{i=0}^{\infty} \widetilde{W}_i \mathrm{e}^{-mT_i} \tag{2.8}$$

and hence by (2.5) and (2.7), we get that

$$\lim_{t\to\infty} D(t)\mathrm{e}^{-mt} = \zeta \quad \text{in } L^1.$$

Finally, since $\{W_i\}_{i\geq 0}$, $\{T_i\}_{i\geq 1}$ are independent, absolutely continuous random variables, ζ is absolutely continuous as well. □

2.4 *Suprema estimates*

We now give some moment estimates which follow by combination of results in the literature.

Let $\{Z(t)\colon t\geq 0\}$ be a Markov branching process with offspring distribution $\{p'_j = p_{j-1}\}_{j\geq 2}$ and lifetime parameter $\lambda = 1$ as in Definition 2.1 with independent initial population $Z(0)$ distributed according to $\{p_j\}_{j\geq 1}$. Recall $m = \sum jp_j$, and, from (2.3) and (2.4),

$$W = \lim_{t\to\infty} Z(t)\mathrm{e}^{-mt} \quad \text{and} \quad \widetilde{W} = \sup_{t\geq 0} Z(t)\mathrm{e}^{-mt}.$$

PROPOSITION 2.5

The following implications hold: *If* $\sum_{j\geq 1}(j\log j)p_j < \infty$, *then* $E[\widetilde{W}] < \infty$.

Also, for $s > 1$, *if* $\sum_{j\geq 1} j^s p_j < \infty$, *then* $E[\widetilde{W}^s] < \infty$.

The proof of the above proposition involves a basic lemma about sums of independent non-negative random variables, which we state below.

Lemma 2.1. *Let* $f\colon [0,\infty)\mapsto[0,\infty)$ *be a concave function and* $S = X_1+\cdots+X_N$ *be a sum of* N *independent nonnegative random variables* $\{X_i\colon i\geq 1\}$. *Then*

$$E[S^\nu f(S)]$$
$$\leq E[S^\nu]f(ES) + \sum_{k=1}^{\nu}\binom{\nu}{k}E[S^{\nu-k}]\sum_{i=1}^{N}E[X_i^k f(X_i)], \quad \textit{for } \nu\geq 1.$$

The proof of this lemma is similar to that of Lemma I.4.5 of [2], where the proof is given for $\nu = 1$, and the general case is stated without proof (see (4.15) in page 42 of [2]). For convenience, we provide a short proof of Lemma 2.1 in the general form in the Appendix.

Proof of Proposition 2.5. First note that, without loss of generality, we can assume $Z(0) = 1$, since the initial value $Z(0)$ in both statements of the proposition is assumed to have enough integrability. To see this more clearly, observe that for $s\geq 1$,

$$E(\widetilde{W}^s) = E[E(\widetilde{W}^s|Z(0))] = \sum_{j\geq 1} p_j E(\widetilde{W}^s|Z(0)=j). \tag{2.9}$$

Let $\{\bar{W}_k\}_{k\geq 1}$ denote a sequence of i.i.d. random variables with distributions same as that of $\sup_{t\geq 0} Z(t)\mathrm{e}^{-mt}$ conditioned on the event $\{Z(0)=1\}$. Using Jensen's inequality, we get for $s\geq 1$,

$$E(\widetilde{W}^s|Z(0)=j) \leq j^s E\left(\frac{1}{j}\sum_{k=1}^{j}\bar{W}_k\right)^s \leq j^s E\left(\frac{1}{j}\sum_{k=1}^{j}\bar{W}_k^s\right) = j^s E\bar{W}_1^s. \tag{2.10}$$

Hence, from (2.9) and (2.10), we get

$$E(\widetilde{W}^s) \leq \left(\sum_{j\geq 1} j^s p_j\right) E(\bar{W}_1^s). \tag{2.11}$$

Hence, it is enough to prove the result for $\bar{W}_1$ (instead of $\widetilde{W}$), or alternatively we can assume that $Z(0) = 1$ for the proof.

Then, first, as $Z(\cdot)$ is increasing, we have

$$\widetilde{W} = \sup_{t\geq 0} Z(t)\mathrm{e}^{-mt} \leq \mathrm{e}^m \sup_{n\geq 0} Z(n)\mathrm{e}^{-mn} := \widetilde{W}_0. \tag{2.12}$$

The process $\{Z(n)\}_{n\geq 0}$ is a discrete-time branching process with

$$W = \lim_{n\to\infty} Z(n)\mathrm{e}^{-mn},$$

$P(Z(1)=0)=0$, $P(Z(1)=j)<1$ for all $j\geq 1$, and $\sum jP(Z(1)=j) = EZ(1) = \mathrm{e}^{-m}$.

From Lemma I.2.6 in [2], for $r \geq 1$, we have, when $P(W > 0) > 0$, that

$$E[\widetilde{W}_0^r] \leq C_0(1 + E[W^r]) \tag{2.13}$$

for a constant C_0.

From Theorem I.10.1 of [5] or Theorem I.2.1 of [2],

$$P(W > 0) > 0,\ E[W] < \infty \text{ if and only if } \sum_{j\geq 1}(j \log j)P(Z(1) = j) < \infty. \tag{2.14}$$

In particular, when $\sum j^r P(Z(1) = j) < \infty$ for $r > 1$, $P(W > 0) > 0$.

Also, from Theorem I.4.4 of [2], and the discussion on p. 41–42 of [2] (cf. eq. (4.15) of [2]), we can derive (see after eq. (2.16) for the argument) that for $r > 1$ there exists a constant $C_1 > 0$ such that

$$E[W^r] \leq C_1\left(1 + \sum_{j\geq 1} j^r P(Z(1) = j)\right). \tag{2.15}$$

From (2.15), we get $E[W^r] < \infty$ when $\sum j^r P(Z(1) = j) < \infty$. From Corollary III.6.2 of [5] (cf. [3]), for $a \geq 1$, $b \geq 0$,

$$\sum_{j\geq 1} j^a |\log j|^b P(Z(1) = j) < \infty \text{ if and only if } \sum_{j\geq 1} j^a |\log j|^b p_j < \infty. \tag{2.16}$$

Then, combining (2.12)–(2.16), we conclude the proof of Proposition 2.5.

We now give the argument for bound (2.15). Since the proof of (2.15) is given only for $1 < r < 2$ in [2] (see Theorem I.4.4, pages 41–42 of [2]), we provide a proof for $r \geq 2$ below. Note that the proof for $1 < r < 2$ given in [2] works for $r = 2$ (see pages 41–42 of [2]), without any modification. So we assume that (2.15) is true for all $1 < r \leq \nu$ for some integer $\nu \geq 2$, and prove that the bound holds for all $\nu < r \leq \nu + 1$ as well. We will use Lemma 2.1 with $f(x) = x^{r-\nu}$ which is concave and nonnegative on $[0, \infty]$. Define the martingale $W'_n = Z(n)e^{-mn}$ and $\mathcal{F}'_n = \sigma\{W'_m : m \leq n\}$ for $n \geq 1$. Note that, conditional on $Z(n)$, $Z(n+1) \stackrel{d}{=} \sum_{i=1}^{Z(n)} X_{n,i}$, where $\{X_{n,i}\}_{i\geq 1}$ are i.i.d. with distribution given by $\{P(Z(1) = j)\}_{j\geq 1}$. Hence, $W'_{n+1} \stackrel{d}{=} \sum_{i=1}^{Z(n)} X_{n,i}e^{-m(n+1)}$, conditional on $\mathcal{F}'_n$. Using Lemma 2.1, noting $\nu \geq 1$, we get

$$\begin{aligned}
&E[(W'_{n+1})^\nu f(W'_{n+1})|\mathcal{F}'_n] \\
&\quad\leq E[(W'_{n+1})^\nu|\mathcal{F}'_n] f(E[W'_{n+1}|\mathcal{F}'_n]) \\
&\qquad+ \sum_{k=1}^{\nu}\binom{\nu}{k} E[(W'_{n+1})^{\nu-k}|\mathcal{F}'_n] \sum_{i=1}^{Z(n)} E[X_{n,i}^k e^{-m(n+1)k} f(X_{n,i}e^{-m(n+1)})] \\
&\quad\leq (W'_n)^\nu f(W'_n) + \sum_{k=1}^{\nu}\binom{\nu}{k}[(W'_n)^{\nu-k}]Z(n)e^{-m(n+1)(k+r-\nu)} \sum_{j\geq 1} j^{k+r-\nu} P(Z(1) = j) \\
&\quad= (W'_n)^\nu f(W'_n) + \sum_{k=1}^{\nu}\binom{\nu}{k}[(W'_n)^{\nu-k+1}]c(n,k) \sum_{j\geq 1} j^{k+r-\nu} P(Z(1) = j),
\end{aligned} \tag{2.17}$$

where $c(n,k) = \exp\{mn - m(n+1)(k+r-\nu)\} = (\mathrm{e}^{-m\alpha_k})^n \mathrm{e}^{-m(\alpha_k+1)}$, and $\alpha_k = (k+r-\nu-1) \geq r-\nu > 0$ for all $k \geq 1$. Hence, we have for all $k \geq 1$ that $c(n,k) \leq (\mathrm{e}^{-m\alpha_k})^n$ and

$$\sum_{n\geq 1} (\mathrm{e}^{-m\alpha_k})^n < \infty. \tag{2.18}$$

Taking expectation and rearranging terms in (2.17), one gets

$$\begin{aligned} &E[(W'_{n+1})^\nu f(W'_{n+1})] - E[(W'_n)^\nu f(W'_n)] \\ &\quad \leq \sum_{k=1}^{\nu} \binom{\nu}{k} E[(W'_n)^{\nu-k+1}](\mathrm{e}^{-m\alpha_k})^n \sum_{j\geq 1} j^{k+r-\nu} P(Z(1)=j). \end{aligned} \tag{2.19}$$

Now, since, $W'_0 = 1$ and $f(x) = x^{r-\nu}$ we have from (2.19) that

$$\begin{aligned} E[W^r] &\leq \underline{\lim}_{N\to\infty} E[(W'_{N+1})^\nu f(W'_{N+1})] \\ &= \underline{\lim}_{N\to\infty} E\left[f(1) + \sum_{n=0}^{N} [(W'_{n+1})^\nu f(W'_{n+1}) - (W'_n)^\nu f(W'_n)] \right] \\ &\leq 1 + \sum_{k=1}^{\nu} \binom{\nu}{k} \sum_{n\geq 0} E[(W'_n)^{\nu-k+1}](\mathrm{e}^{-m\alpha_k})^n \sum_{j\geq 1} j^r P(Z(1)=j), \end{aligned} \tag{2.20}$$

observing $j^{k+r-\nu} \leq j^r$ for $j \geq 1$ when $1 \leq k \leq \nu$. Notice that $\lim_{n\to\infty} E[(W'_n)^{\nu-k+1}] = EW^{\nu-k+1} < \infty$, by the induction hypothesis, (2.13), $W'_n \to W$ a.s., and dominated convergence. As $1 \leq \nu - k + 1 \leq \nu$, the proof of (2.15) is complete using (2.18). □

Now let $\{D(t): t \geq 0\}$ be a Markov branching process with offspring distribution $\{p'_j = p_{j-1}\}_{j\geq 2}$, lifetime $\lambda = 1$ and immigration $\beta \geq 0$ parameters, and immigration distribution $\{p_j\}_{j\geq 1}$ as in Proposition 2.4 with also $D(0)$ distributed as $\{p_j\}_{j\geq 1}$. Let also

$$\widetilde{D} := \sup_{t\geq 0} D(t)\mathrm{e}^{-mt}.$$

PROPOSITION 2.6

For $r > 1$, we have

$$\textit{if} \sum_{j\geq 1} j^r p_j < \infty, \quad \textit{then} \quad E[\widetilde{D}^r] < \infty.$$

Proof. When $\beta = 0$, the statement is the same as Proposition 2.5. When $\beta > 0$, as in the proof of Proposition 2.4, let $\{T_i\}_{i\geq 1}$ be the times of immigration, and $T_0 = 0$. Note that $\sum_{i\geq 0} \mathrm{e}^{-mT_i} < \infty$ a.s. as the expected value $\sum_{i\geq 0} (\beta/(m+\beta))^i$ is finite. From (2.8), and Jensen's inequality, we have

$$\widetilde{D}^r \leq \left(\sum_{i\geq 0} \widetilde{W}_i \mathrm{e}^{-mT_i} \right)^r$$

$$\leq \left(\left[\sum_{j \geq 0} \mathrm{e}^{-mT_j} \right]^{-1} \sum_{i \geq 0} \widetilde{W}_i^r \mathrm{e}^{-mT_i} \right) \left(\sum_{j \geq 0} \mathrm{e}^{-mT_j} \right)^r$$

$$= \left(\sum_{i \geq 0} \widetilde{W}_i^r \mathrm{e}^{-mT_i} \right) \left(\sum_{j \geq 0} \mathrm{e}^{-mT_j} \right)^{r-1}.$$

Hence, by independence of $\{\widetilde{W_i}\}_{i \geq 0}$ and $\{T_i\}_{i \geq 0}$, for any integer $K \geq r - 1$, we have

$$E[\widetilde{D}^r] \leq E[\widetilde{W}_1^r] \sum_{i \geq 0} E \left[\mathrm{e}^{-mT_i} \left(\sum_{j \geq 0} \mathrm{e}^{-mT_j} \right)^K \right].$$

From Proposition 2.5, $E[\widetilde{W}_1^r] < \infty$. Also,

$$E \left[\mathrm{e}^{-mT_i} \left(\sum_{j \geq 0} \mathrm{e}^{-mT_j} \right)^K \right] \leq E[\mathrm{e}^{-2mT_i}]^{1/2} E \left[\left(\sum_{j \geq 0} \mathrm{e}^{-mT_j} \right)^{2K} \right]^{1/2}$$

$$= \left(\sqrt{\frac{\beta}{2m + \beta}} \right)^i E \left[\left(\sum_{j \geq 0} \mathrm{e}^{-mT_j} \right)^{2K} \right]^{1/2}.$$

Given T_j is the sum of j independent exponential random variables with parameter β for $j \geq 1$, we now bound

$$\begin{aligned}
&E \left[\left(\sum_{j \geq 0} \mathrm{e}^{-mT_j} \right)^{2K} \right] \\
&\quad \leq (2K)! \sum_{0 \leq j_1 \leq \cdots \leq j_{2K}} E \left[\prod_{l=1}^{2K} \mathrm{e}^{-mT_{j_l}} \right] \\
&\quad = (2K)! \sum_{0 \leq j_1 \leq \cdots \leq j_{2K}} E \left[\prod_{l=1}^{2K-2} \mathrm{e}^{-mT_{j_l}} \mathrm{e}^{-2mT_{j_{2K-1}}} \right] \left(\frac{\beta}{m + \beta} \right)^{j_{2K} - j_{2K-1}} \\
&\quad \leq (2K)! \left(\frac{m + \beta}{m} \right) \sum_{0 \leq j_1 \leq \cdots \leq j_{2K-1}} E \left[\prod_{l=1}^{2K-1} \mathrm{e}^{-mT_{j_l}} \right] \\
&\quad \leq (2K)! \left(\frac{m + \beta}{m} \right)^{2K}
\end{aligned}$$

is finite for fixed K. $\square$

3. Proof of main results

We give the proofs of the three main results in successive subsections.

3.1 *Growth rates for degrees and the maximal degree*

We first begin with a basic analysis result.

PROPOSITION 3.1

Let $\{a_{n,i}\colon 1 \le i \le n\}_{n\ge 1}$ *be a double array of nonnegative numbers such that*

(1) *For all* $i \ge 1$, $\lim_{n\to\infty} a_{n,i} = a_i < \infty$,
(2) $\sup_{n\ge 1} a_{n,i} \le b_i < \infty$,
(3) $\lim_{i\to\infty} b_i = 0$,
(4) *For* $i \ne j$, $a_i \ne a_j$.

Then,

(a) $\max_{1\le i\le n} a_{n,i} \to \max_{i\ge 1} a_i$, *as* $n \to \infty$.
(b) *In addition*, *there exists* I_0 *and* N_0 *such that* $\max_{1\le i\le n} a_{n,i} = a_{n,I_0}$ *for* $n \ge N_0$.

Proof. For each $k \ge 1$,

$$\lim_{n\to\infty} \max_{1\le i\le k} a_{n,i} = \max_{1\le i\le k} a_i.$$

Hence,

$$\underline{\lim}_{n\to\infty} \max_{i\ge 1} a_{n,i} \ge \underline{\lim}_{n\to\infty} \max_{1\le i\le k} a_{n,i} = \max_{1\le i\le k} a_i$$

which gives

$$\underline{\lim}_{n\to\infty} \max_{i\ge 1} a_{n,i} \ge \max_{i\ge 1} a_i. \tag{3.1}$$

Also, for each $k \ge 1$,

$$\max_{i\ge 1} a_{n,i} \le \max_{1\le i\le k} a_{n,i} + \max_{i>k} b_i.$$

Then,

$$\overline{\lim}_{n\to\infty} \max_{i\ge 1} a_{n,i} \le \max_{1\le i\le k} a_i + \max_{i>k} b_i \le \max_{i\ge 1} a_i + \max_{i\ge k} b_i.$$

Since $\lim_{i\to\infty} b_i = \overline{\lim}_{i\to\infty} b_i = \lim_{k\to\infty} \max_{i\ge k} b_i = 0$, we have

$$\overline{\lim}_{n\to\infty} \max_{i\ge 1} a_{n,i} \le \max_{i\ge 1} a_i. \tag{3.2}$$

Now, (3.1) and (3.2) yield part (a). By assumptions (3) and (4), $\max_{i\ge 1} a_i$ is attained at some finite index I_0, and this index is unique, giving part (b). □

Proof of Theorem 1.1. By the embedding theorem (Theorem 2.1), to establish Theorem 1.1 for the sequence $\{\widetilde{Z}_n\}_{n\ge 0}$, it suffices to prove the corresponding results for the $\{Z_n\}_{n\ge 0}$ sequence.

By Proposition 2.4 and Corollary 2.1(i),

$$\lim_{n\to\infty} D_i(\tau_n - \tau_{i-2})\mathrm{e}^{-m(\tau_n - \tau_{i-2})} = \zeta_i$$

converges a.s. in $(0, \infty)$ for $i \geq 1$. By Proposition 2.3, a.s. as $n \uparrow \infty$,

$$\exp\left\{\frac{-m}{2m+\beta}\sum_{j=1}^{n}\frac{1}{j}\right\}\exp\{m\tau_n\} = \exp\left\{m\left(\tau_n - \sum_{j=1}^{n}\frac{1}{j(2m+\beta)}\right)\right\}$$
$$\to \mathrm{e}^{mY}.$$

Further, $\sum_{j=1}^{n}(1/j) - \log n \to \gamma$, a Euler's constant. Thus, a.s. as $n \uparrow \infty$,

$$\mathrm{e}^{m\tau_n} n^{-m/(2m+\beta)} \to \mathrm{e}^{mY}\mathrm{e}^{m\gamma/(2m+\beta)} := V, \tag{3.3}$$

where V is a positive real random variable. Hence,

$$\begin{aligned} &D_i(\tau_n - \tau_{i-2})n^{-m/(2m+\beta)} \\ &\quad = D_i(\tau_n - \tau_{i-2})\mathrm{e}^{-m(\tau_n-\tau_{i-2})}\mathrm{e}^{-m\tau_{i-2}}\mathrm{e}^{m\tau_n}n^{-m/(2m+\beta)} \\ &\quad \to \zeta_i \mathrm{e}^{-m\tau_{i-2}}V := \xi_i V, \end{aligned}$$

a.s. as $n \uparrow \infty$, where $\xi_i = \zeta_i \mathrm{e}^{-m\tau_{i-2}}$ is a positive real random variable. This proves part (i) with $\gamma_i = \xi_i V$.

By independence of τ_{i-2} and $\{D_i(t)\}_{t\geq 0}$, absolute continuity of τ_{i-2} for $i \geq 3$ ($\tau_0 = \tau_{-1} = 0$), and Proposition 2.4, it follows that ξ_i has an absolutely continuous distribution with finite mean, proving part (ii).

To prove part (iii) and (iv), we first note, for each $i \geq 1$, that

$$D_i(\tau_n - \tau_{i-2})\mathrm{e}^{-m(\tau_n-\tau_{i-2})} \leq \sup_{t\geq 0} D_i(t)\mathrm{e}^{-mt} := \widetilde{D}_i.$$

Let

$$\begin{aligned} a_{n,i} &= D_i(\tau_n - \tau_{i-2})\mathrm{e}^{-m\tau_n} \quad \text{for } 1 \leq i \leq n, \quad \text{and} \\ b_i &= \widetilde{D}_i \mathrm{e}^{-m\tau_{i-2}} \quad \text{for } i \geq 1. \end{aligned}$$

For each $i \geq 1$, $\sup_{n\geq 1} a_{n,i} \leq b_i$ and $a_{n,i} \to \zeta_i \mathrm{e}^{-m\tau_{i-2}} := a_i$ say. Since $\sum_{j=1}^{\infty} j^r p_j < \infty$ for some $r > 1$ (satisfying $rm/(2m+\beta) = r\theta > 1$), we have that $E(\widetilde{D}_i^r) < \infty$ (Proposition 2.6). By Markov's inequality, for all $\epsilon > 0$,

$$P(\widetilde{D}_i > \epsilon i^{m/(2m+\beta)}) \leq E[\widetilde{D}_1^r]/(\epsilon^r i^{rm/(2m+\beta)}).$$

Hence, by Borel–Cantelli, we have a.s.

$$\widetilde{D}_i \leq \epsilon i^{m/(2m+\beta)} \text{ for all large } i.$$

Note, by Corollary 2.1(ii), we have $m\tau_{i-2} - [m/(2m+\beta)]\log(i-2) \to \tilde{Y}$ a.s., for $i \to \infty$ for some finite random variable $\tilde{Y}$. Hence, as $\epsilon > 0$ is arbitrary, it follows that $b_i = \widetilde{D}_i \mathrm{e}^{-m\tau_{i-2}} \to 0$ a.s. as $i \uparrow \infty$. By Proposition 3.1 and (3.3), this implies that a.s.,

$$\lim_{n\to\infty} \max_{1\leq i\leq n} D_i(\tau_n - \tau_{i-2})n^{-m/(2m+\beta)} = V \max_{i\geq 1} \zeta_i \mathrm{e}^{-m\tau_{i-2}}.$$

Now we claim $\{\zeta_i \mathrm{e}^{-m\tau_{i-2}}\}_{i\geq 1}$ are all distinct, that is $P(\zeta_j \mathrm{e}^{-m\tau_{j-2}} = \zeta_i \mathrm{e}^{-m\tau_{i-2}}) = P(\zeta_j \mathrm{e}^{-m(\tau_{j-2}-\tau_{i-2})} = \zeta_i) = 0$ for any $1 \leq i < j$. This follows from the fact that ζ_i is measurable with respect to $\mathcal{F}_{i-2}$ (cf. (2.1)), and conditioned on $\mathcal{F}_{i-2}$, the random variables ζ_j and $\tau_{j-2} - \tau_{i-2}$ are independent with absolutely continuous distributions for $j \geq 3$; when $j = 2, i = 1$, note that ζ_2 is absolutely continuous and $\tau_{j-2} - \tau_{i-2} = \tau_0 - \tau_{-1} = 0$. Hence, $V \max_{i\geq 1} \zeta_i \mathrm{e}^{-m\tau_{i-2}}$ is attained at a unique index I_0. Also, as $\{I_n\}_{n\geq 1}$ are integer-valued random variables, I_n will equal I_0 a.s. for all large n. □

3.2 *Convergence of the empirical distribution of degrees*

The following lemma will be helpful in the proof of Theorem 1.2.

Lemma 3.1. *Let* $\{X(t)\colon t \geq 0\}$ *be a continuous-time*, *discrete state-space*, *Markov chain which is non-explosive*, *that is the number of jumps of* $\{X(t)\colon t \geq 0\}$ *in any finite time-interval* $[0, K]$ *is finite a.s.. For* $K > 0, \delta > 0$, *let*

$$p_K(\delta) \equiv \sup_{0\leq t\leq K} P(|X(t+\delta) - X((t-\delta)\vee 0)| \geq 1).$$

Then, *for all* $K > 0$,

$$\lim_{\delta\downarrow 0} p_K(\delta) = 0.$$

Proof. Since $\{X(t)\colon t \geq 0\}$ is non-explosive, for any $0 < K < \infty$, the number of jumps $N(K)$ of $\{X(t)\colon 0 \leq t \leq K\}$ is a finite-valued random variable a.s.. Also for any $j < \infty$, the jump times $(T_1, \ldots, T_j)$ of $\{X(t)\colon t \geq 0\}$ have a continuous joint distribution. These two facts together yield the lemma. □

The following result follows from Remark 2.1 and the above lemma.

COROLLARY 3.1

Let $\{D(t)\colon t \geq 0\}$ *be as in Definition* 2.2. *Define*, *for* $0 \leq s \leq t$, $D(s,t] = D(t) - D(s)$ *and*, *for* $K > 0, \delta > 0$,

$$p_K^D(\delta) = \sup_{0\leq t\leq K} P(D((t-\delta)\vee 0, t+\delta] \geq 1).$$

Then for $K > 0$,

$$\lim_{\delta\downarrow 0} p_K^D(\delta) = 0.$$

Proof of Theorem 1.2. Recall $\alpha = (2m+\beta)^{-1}$ (from subsection 2.2). For $n \geq 1$, note that

$$\frac{R_j(n)}{n} \overset{d}{=} \frac{1}{n}\sum_{i=1}^{n+2} I(D_i(\tau_n - \tau_{i-2}) = j)$$

$$= \frac{1}{n}\sum_{i=1}^{n+2}\{I(D_i(\tau_n - \tau_{i-2}) = j) - I(D_i(\alpha\log(n/(i-2))) = j)\}$$

$$+ \frac{1}{n}\sum_{i=1}^{n+2}\{I(D_i(\alpha\log(n/(i-2))) = j) - p_j(\alpha\log(n/(i-2)))\}$$

$$+ \frac{1}{n}\sum_{i=1}^{n+2}\left\{p_j(\alpha\log(n/(i-2))) - \int_0^1 p_j(-\alpha\log(x))\right\}$$

$$+ \int_0^1 p_j(-\alpha\log(x))\mathrm{d}x$$

$$= J_1(n) + J_2(n) + J_3(n) + J_4(n), \quad \text{say.}$$

For notational convenience, we use the convention $\log(n/(i-2)) \equiv \log(n)$ for $i = 1, 2$ here. The proof is now obtained by showing $J_i(n)$ vanishes in probability for $i = 1, 2, 3$ and observing, after change of variables, that

$$J_4(n) \equiv \frac{1}{\alpha}\int_0^\infty p_j(y)\exp\{-y/\alpha\}\mathrm{d}y.$$

To show that the first term $J_1(n)$ goes to 0 in probability, fix $\epsilon > 0$. Note, from Corollary 2.1(iii), for $\delta > 0$, if

$$\mathcal{A}_n(\delta) \equiv \left\{\sup_{n\epsilon+2\le i\le n+2} |\tau_n - \tau_{i-2} - \alpha\log(n/(i-2))| > \delta\right\},$$

then

$$\overline{\lim}_{\delta\downarrow 0}\overline{\lim}_{n\to\infty} P(\mathcal{A}_n(\delta)) = 0. \tag{3.4}$$

Now for $n\epsilon + 2 \le i \le n+2$, we have $0 \le \alpha\log(n/(i-2)) \le -\alpha\log\epsilon$. Hence, from the definition of $\mathcal{A}_n(\delta)$, we get the following bound on the expectation of a typical summand in J_1 in this range. Using notation from Corollary 3.1, we have

$$\begin{aligned}
&E(|I(D_i(\tau_n - \tau_{i-2}) = j) - I(D_i(\alpha\log(n/(i-2))) = j)|)\\
&\quad = P(|I(D_i(\tau_n - \tau_{i-2}) = j) - I(D_i(\alpha\log(n/(i-2))) = j)| = 1)\\
&\quad \le P(|D_i(\tau_n - \tau_{i-2}) - D_i(\alpha\log(n/(i-2))| \ge 1))\\
&\quad \le P(\{|D_i(\tau_n - \tau_{i-2}) - D_i(\alpha\log(n/(i-2)))| \ge 1\} \cap \mathcal{A}_n^c(\delta)) + P(\mathcal{A}_n(\delta))\\
&\quad \le P(\{D_i((\alpha\log(n/(i-2)) - \delta)\vee 0, \alpha\log(n/(i-2)) + \delta] \ge 1\}\\
&\qquad \cap \mathcal{A}_n^c(\delta)) + P(\mathcal{A}_n(\delta))\\
&\quad \le \sup_{a\in[0,-\alpha\log\epsilon]} P(\{D_i((a-\delta)\vee 0, a+\delta] \ge 1\} \cap \mathcal{A}_n^c(\delta)) + P(\mathcal{A}_n(\delta))\\
&\quad = p_{K(\epsilon)}^{D_3}(\delta) + P(\mathcal{A}_n(\delta)),
\end{aligned}$$

where $K(\epsilon) = -\alpha \log \epsilon$, and we recall $D_3(\cdot)$ is a Markov branching process with immigration with $D_3(0)$ distributed according to $\{p_j\}_{j\geq 1}$. Since the absolute value of each summand in J_1 is bounded (by 1), we have by splitting the sum over indices $1 \leq i < n\epsilon + 2$ and $n\epsilon + 2 \leq i \leq n + 2$, the following bound:

$$E|J_1(n)| \leq \frac{1}{n}(n\epsilon + 2) + \frac{1}{n}(n - n\epsilon)[p^{D_3}_{K(\epsilon)}(\delta) + P(\mathcal{A}_n(\delta))].$$

Now for fixed $\epsilon > 0$, taking limit as $n \to \infty$ first and then over $\delta \to 0$, we get from Corollary 3.1 and (3.4) that

$$\overline{\lim}_{n\to\infty} E|J_1(n)| < \epsilon$$

and, as $\epsilon > 0$ is arbitrary, that $\lim_{n\to\infty} E|J_1(n)| = 0$. Hence $J_1(n) \to 0$ in probability, as $n \to \infty$.

For the second term $J_2(n)$, we have from Markov's inequality that for any $\epsilon > 0$,

$$\begin{aligned} &P(|J_2(n)| > \epsilon) \\ &\quad \leq \frac{1}{n^4\epsilon^4} E\left(\sum_{i=1}^{n+2}(I(D_i(\alpha\log(n/(i-2))) = j) - p_j(\alpha\log(n/(i-2)))\right)^4 \\ &\quad \leq \frac{6}{n^2\epsilon^4}, \end{aligned}$$

using independence of $\{D_i(\cdot)\}_{i\geq 1}$, and hence of the summands above. Now, by Borel–Cantelli arguments and the method of fourth moments (cf. Theorem 8.2.1 of [6]), we get $J_2(n) \to 0$ a.s., as $n \to \infty$.

Finally, by simple estimates, and Riemann integrability of $p_j(-\alpha \log(x))$ (as $p_j(\cdot)$ is bounded, continuous), the third term vanishes as $n \to \infty$. □

3.3 *Power-laws for limiting empirical degree distribution*

Recall, with respect to the definition of π_j (1.2), that $\{D(y)\colon y \geq 0\}$ is a Markov branching process with exponential ($\lambda = 1$) lifetime distribution, offspring distribution $\{p'_j = p_{j-1}\}_{j\geq 2}$, immigration rate $\beta \geq 0$, immigration size distribution $\{p_j\}_{j\geq 1}$, and initial value $D(0)$ distributed according to $\{p_j\}_{j\geq 1}$.

Proof of Theorem 1.3. First note, for $s \geq 0$, that

$$\begin{aligned} \sum_{j\geq 1} j^s \pi_j &= (2m + \beta) \int_0^\infty \mathrm{e}^{-(2m+\beta)y} \sum_{j\geq 1} j^s p_j(y)\mathrm{d}y \\ &= (2m + \beta) \int_0^\infty \mathrm{e}^{-(2m+\beta)y} E[(D(y))^s]\mathrm{d}y. \end{aligned} \tag{3.5}$$

By Proposition 2.4, $D(y)\mathrm{e}^{-my} \to \zeta$ a.s., and when $\sum_{j\geq 1} j^{2+\beta/m} p_j < \infty$, we have by Proposition 2.6 that $E[(\sup_{y\geq 0} D(y)\mathrm{e}^{-my})^{2+\beta/m}] < \infty$. Hence, by dominated convergence, for $s \leq 2 + \beta/m$, and a constant C,

$$C\mathrm{e}^{msy} \leq E[(D(y))^s] \leq C^{-1}\mathrm{e}^{msy}.$$

Plugging into (3.5), we prove the theorem for all $0 \le s \le 2+\beta/m$. Noting $\sum_{j\ge1} j^s\pi_j \ge \sum_{j\ge1} j^{2+\beta/m}\pi_j$ for $s \ge 2+\beta/m$ finishes the proof. □

We now evaluate π_j in a special case. The formula, similar to that in [13] and §4.2 of [18], gives the asymptotics mentioned in Remark 1.3.

PROPOSITION 3.2

When $X_i \equiv x_0$ for an integer $x_0 \ge 1$, we have for $j \ge 1$ that

$$\pi_j = \frac{2x_0+\beta}{j+2x_0+2\beta}\prod_{k=1}^{l-1}\frac{kx_0+\beta}{(k+2)x_0+2\beta}$$

when $j = lx_0$ for $l \ge 1$ (where the product is set equal to 1 when $l = 1$), and $\pi_j = 0$ otherwise. Hence, for large j, $\pi_j = O(j^{-[3+\beta/x_0]})$ when $j = lx_0$ for $l \ge 1$, and $\pi_j = 0$ otherwise.

Proof. First note as $X_i \equiv x_0$ that $m = D(0) = x_0$, and the process $D(\cdot)$ moves in steps of x_0. Clearly, then $\pi_j = 0$ when j is not a multiple of x_0. When $j = lx_0$ for $l \ge 1$, let A_j be the first time the process $D(\cdot)$ equals j,

$$A_j = \inf\{y \ge 0\colon D(y) = j\} < \infty \quad \text{a.s.}$$

Let B_j be the time, after A_j, that the process $D(\cdot)$ spends at j; note that conditioned on A_j, B_j is an exponential $(j+\beta)$ variable. Then, we write

$$\begin{aligned}
\pi_j &= (2x_0+\beta)\int_0^\infty \mathrm{e}^{-y(2x_0+\beta)}p_j(y)\mathrm{d}y\\
&= (2x_0+\beta)E\left[\int_{A_j}^{A_j+B_j}\mathrm{e}^{-y(2x_0+\beta)}\mathrm{d}y\right]\\
&= E[\mathrm{e}^{-A_j(2x_0+\beta)}[1-\mathrm{e}^{-B_j(2x_0+\beta)}]]\\
&= E[\mathrm{e}^{-A_j(2x_0+\beta)}]\left[1-\frac{j+\beta}{j+2x_0+2\beta}\right]\\
&= \frac{2x_0+\beta}{j+2x_0+2\beta}E[\mathrm{e}^{-A_j(2x_0+\beta)}].
\end{aligned}$$

As $X_i \equiv x_0$, for $j = lx_0$ and $l \ge 2$, we have A_j is the sum of $l-1$ independent exponential random variables with parameters $x_0+\beta, \ldots, (l-1)x_0+\beta$, and so

$$E[\mathrm{e}^{-A_j(2x_0+\beta)}] = \prod_{k=1}^{l-1}\frac{kx_0+\beta}{(k+2)x_0+2\beta}.$$

When $l = 1$, then $A_j = 0$, giving $\pi_{x_0} = (2x_0+\beta)/(3x_0+2\beta)$. □

4. Appendix

Proof of Lemma 2.1. For $q \ge 1$ and a finite set $D' \subset \mathbb{N}$, define

$$\mathcal{A}^q_{D'} = \{\langle i_1, \ldots, i_q\rangle\colon i_k \in D', k = 1, \ldots, q\}.$$

For $A \in \mathcal{A}^q_{D'}$, let $m_i = \sum_{k=1}^q I(i_k = i)$ for $i \in D'$, and $B_A = \{i \in D' \colon m_i > 0\}$. Note that for all $i \in D'$, $m_i \geq 0$ and $\sum_{i \in D'} m_i = q$. For any such finite set of indices D' and any integer $q \geq 1$, one has the following identity for any sequence of reals $\{x_i\}_{i \geq 1}$,

$$\left(\sum_{i \in D'} x_i\right)^q = \sum_{A \in \mathcal{A}^q_{D'}} \prod_{i \in B_A} x_i^{m_i}. \tag{4.1}$$

Also, we remark, as noted in Lemma I.4.5 of [2] that the concavity and nonnegativity of the function f on $[0, \infty)$ implies that f is also increasing, and subadditive. Indeed, for the specific function $f(x) = x^{r-\nu}$ for $0 < r - \nu < 1$, used in the argument of (2.15), by a simple inspection, this holds.

Now, fix an integer $N \geq 1$ and let $D = \{1, \ldots, N\}$. Using the above identity, the independence of $\{X_i\}$ and the subadditivity of f, we have

$$\begin{aligned} E[S^\nu f(S)] \leq{} & \sum_{A \in \mathcal{A}^\nu_D} E\left[\left(\prod_{i \in B_A} X_i^{m_i}\right) f\left(\sum_{i \in B_A} X_i\right)\right] \\ & + \sum_{A \in \mathcal{A}^\nu_D} E\left(\prod_{i \in B_A} X_i^{m_i}\right) E\left[f\left(\sum_{i \in B_A^c} X_i\right)\right]. \end{aligned} \tag{4.2}$$

Using Jensen's inequality, the fact that f is increasing, $\{X_i\}$ are nonnegative, and the identity (4.1) above, we get the following bound on the second term in (4.2):

$$\begin{aligned} \sum_{A \in \mathcal{A}^\nu_D} E\left(\prod_{i \in B_A} X_i^{m_i}\right) E\left[f\left(\sum_{i \in B_A^c} X_i\right)\right] &\leq \sum_{A \in \mathcal{A}^\nu_D} E\left(\prod_{i \in B_A} X_i^{m_i}\right)\left[f\left(\sum_{i \in B_A^c} EX_i\right)\right] \\ &\leq \sum_{A \in \mathcal{A}^\nu_D} E\left(\prod_{i \in B_A} X_i^{m_i}\right) f(ES) \\ &= E[S^\nu] f(ES). \end{aligned} \tag{4.3}$$

For the first term in (4.2), we use the subadditivity of f, and the independence of $\{X_i\}$ to get

$$\begin{aligned} & \sum_{A \in \mathcal{A}^\nu_D} E\left[\left(\prod_{i \in B_A} X_i^{m_i}\right) f\left(\sum_{i \in B_A} X_i\right)\right] \\ & \quad \leq \sum_{A \in \mathcal{A}^\nu_D} E\left[\sum_{i \in B_A} f(X_i)\left(\prod_{i' \in B_A} X_{i'}^{m_{i'}}\right)\right] \\ & \quad \leq \sum_{A \in \mathcal{A}^\nu_D} \left[\sum_{i \in B_A} E(X_i^{m_i} f(X_i)) E\left(\prod_{i' \in B_A \setminus \{i\}} X_{i'}^{m_{i'}}\right)\right]. \end{aligned} \tag{4.4}$$

Now the collection $\{A \in A^\nu_D \colon m_i = m\}$ are those indices $\langle i_1, \ldots, i_\nu \rangle$ where components $i_{q_1} = \cdots = i_{q_m} = i$ for m distinct locations $q_1, \ldots, q_m \in \{1, \ldots, \nu\}$, and the other components, put together as $\langle j_1, \dots j_{\nu-m}\rangle$, belongs to $A^{\nu-m}_{D\backslash\{i\}}$. Then, (4.4) is bounded above by

$$\sum_{m=1}^{\nu}\binom{\nu}{m}\sum_{i=1}^{N}E(X_i^m f(X_i))E\left\{\sum_{A\in\mathcal{A}^{\nu-m}_{D\backslash\{i\}}}\left(\prod_{i'\in B_A}X_{i'}^{m_{i'}}\right)\right]$$

$$\leq \sum_{m=1}^{\nu}\binom{\nu}{m}\sum_{i=1}^{N}E(X_i^m f(X_i))E[S^{\nu-m}] \tag{4.5}$$

noting (4.1) and $E[(S-X_i)^{\nu-m}]\leq E[S^{\nu-m}]$. From (4.2)–(4.5), the proof of the lemma is complete. □

Acknowledgement

The authors would like to thank the referee for helpful suggestions and comments, which improved the exposition of the material presented in this paper. Research supported in part by NSA-H982300510041, NSF-DMS-0504193 and NSF-DMS-0608634.

References

[1] Albert R and Barabasi A-L, Statistical mechanics of complex networks, *Rev. Mod. Phys.* **74** (2002) 47–97
[2] Asmussen S and Hering H, Branching Processes, Progress in Probability and Statistics, 3 (Boston: Birkhäuser) (1983)
[3] Athreya K B, On the equivalence of conditions on a branching process in continuous time and on its offspring distribution, *J. Math. Kyoto Univ.* **9** (1969) 41–53
[4] Athreya K B and Karlin S, Limit theorems for the split times of branching processes, *J. Math. Mech.* **17** (1967) 257–277
[5] Athreya K B and Ney P, Branching Processes (New York: Dover) (2004)
[6] Athreya K B and Lahiri S N, Measure Theory and Probability Theory (New York: Springer) (2006)
[7] Barabasi A-L and Albert R, Emergence of scaling in random networks, *Science* **286** (1999) 509–512
[8] Bollobas B, Riordan O, Spencer J and Tusnady G, The degree sequence of a scale-free random graph process, *Random Structures and Algorithms* **18** (2001) 279–290
[9] Chung F and Lu L, Complex Graphs and Networks, CBMS, 107, American Mathematical Society, Providence (2006)
[10] Cooper C and Frieze A, A general model of web graphs, *Random Structures and Algorithms* **22** (2003) 311–335
[11] Durrett R, Random Graph Dynamics (Boston: Cambridge University Press) (2006)
[12] Mitzenmacher M, A brief history of generative models for power law and lognormal distributions, *Internet Math.* **1** (2006) 226–251
[13] Mori T, On random trees, *Studia Sci. Math. Hungar.* **39** (2002) 143–155
[14] Mori T, The maximum degree of the Barabasi-Albert random tree, *Comb. Probab. Computing* **14** (2005) 339–348
[15] Newman M E J, The structure and function of complex networks, *SIAM Review* **45** (2003) 167–256
[16] Oliveira R and Spencer J, Connectivity transitions in networks with super-linear preferential attachment, *Internet Math.* **2** (2006) 121–163

[17] Pemantle R, A survey of random processes with reinforcement, *Probability Surveys* **4** (2007) 1–79

[18] Rudas A, Toth B and Valko B, Random trees and general branching processes, *Random Structures and Algorithms* **31** (2007) 186–202

[19] Simon H A, On a class of skew distribution functions, *Biometrika* **42** (1955) 425–440

[20] Yule G, A mathematical theory of evolution based on the conclusions of Dr J. C. Willis, *Phil. Trans. Roy. Soc. London* **B213** (1925) 21–87