

ON THE NUMBER OF VERTICES WITH A GIVEN DEGREE IN A GALTON–WATSON TREE

NARIYUKI MINAMI,* *University of Tsukuba*

Abstract

Let $Y_k(\omega)$ ($k \geq 0$) be the number of vertices of a Galton–Watson tree ω that have k children, so that $Z(\omega) := \sum_{k \geq 0} Y_k(\omega)$ is the total progeny of ω . In this paper, we will prove various statistical properties of Z and Y_k . We first show, under a mild condition, an asymptotic expansion of $P(Z = n)$ as $n \rightarrow \infty$, improving the theorem of Otter (1949). Next, we show that $\mathcal{Y}_k(\omega) := \sum_{j=0}^k Y_j(\omega)$ is the total progeny of a new Galton–Watson tree that is hidden in the original tree ω . We then proceed to study the joint probability distribution of Z and $\{Y_k\}_k$, and show that, as $n \rightarrow \infty$, $\{Y_k/n\}_k$ is asymptotically Gaussian under the conditional distribution $P(\cdot \mid Z = n)$.

Keywords: Galton–Watson tree; total progeny; Lagrange inversion; central limit theorem

2000 Mathematics Subject Classification: Primary 60J80

Secondary 60C05; 05A15; 60F05

1. Introduction

Let $\Pi = \{p_n\}_{n=0}^\infty$ be a probability distribution on nonnegative integers and let $\{X_n\}_{n=0}^\infty$ ($X_0 = 0$) be the Galton–Watson process (or the discrete-time branching process) with the offspring distribution Π . Various long-time behaviors of the integer-valued Markov chain $\{X_n\}$ form the main subject matter of classical treatises of Galton–Watson processes (e.g. [2], [14], [17], [27]). There, the random tree structure, which is obvious in the intuitive description of the process, is rather implicit or even entirely omitted from the exposition. In this paper, we are interested in the typical shape of the random tree, which we will call the Galton–Watson tree, obtained as the ‘trajectory’ of the branching mechanism that gives rise to the process $\{X_n\}$. Namely, a Galton–Watson tree is the random graph each of whose vertices gives birth to a random number of children, according to the probability distribution Π and independently of each other. Two vertices are adjacent if and only if one is the parent of the other. Let Z be the total number of vertices, or the total progeny, of the Galton–Watson tree and let Y_k be the number of vertices having k children. Motivated by the classical pioneering work by Otter [24], which does not yet seem to have been fully appreciated, we will investigate in some detail the asymptotic behavior of the probability distribution of Z , and of $\mathcal{Y}_k := \sum_{j=0}^k Y_j$, $k = 0, 1, 2, \dots$, the number of vertices having at most k children. It will be seen that \mathcal{Y}_k is the total progeny of a Galton–Watson tree that is hidden in the original tree. We will also prove a central limit theorem for the distribution of (Y_0, Y_1, \dots) conditioned on Z . As a corollary of this theorem, the central limit theorem due to Mahmoud [22] for vertices of uniform binary trees will be reproduced. Our proofs are based on the analysis of generating functions, and we rely entirely

Received 25 May 2004; revision received 9 October 2004.

* Postal address: Institute of Mathematics, University of Tsukuba, Tsukuba 305-8571, Japan.

Email address: minami@math.tsukuba.ac.jp

upon the Lagrange inversion formula in obtaining explicit formulae for various probabilities, and upon the classical saddle point method in obtaining their asymptotic expressions.

In order to prepare a solid basis for our work, it is necessary to give a precise description of ‘trees’ and to introduce a suitable σ -field \mathcal{F} as well as a probability measure P on the space Ω of such trees. In this work we follow the convenient and elegant construction of the probability space (Ω, \mathcal{F}, P) due to Neveu [23]. (It should be remarked, however, that an alternative construction of the probability space of a Galton–Watson tree had already been given by Otter himself [24]. This turns out to be equivalent to that of Neveu (see [21]), and even has some advantages when one takes Ω to be topological space.)

Now, let U be the totality of finite sequences $u = j_1 j_2 \cdots j_n$ of strictly positive integers. We can write $U = \bigcup_{n \geq 0} \mathbb{N}^n$, where $\mathbb{N} = \{1, 2, \dots\}$ and $\mathbb{N}^0 = \{\phi\}$, ϕ standing for the empty sequence. A tree ω is by definition a subset of U satisfying the following three conditions: (a) $\phi \in \omega$; (b) if $uj \in \omega$ then $u \in \omega$; and (c) for each $u \in \omega$, there is a nonnegative integer $\nu_u(\omega)$ such that $uj \in \omega \Leftrightarrow 1 \leq j \leq \nu_u(\omega)$. Here, for two sequences $u = j_1 j_2 \cdots j_n$ and $v = k_1 k_2 \cdots k_p$, we write uv for their conjunction: $uv = j_1 j_2 \cdots j_n k_1 k_2 \cdots k_p$. In particular, $uj = j_1 j_2 \cdots j_n j$ for $j \in \mathbb{N}$. Thus, ω is a graph with vertices $u \in \omega$ and edges of the form (u, uj) . Condition (a) says that ω always contains the special vertex ϕ , the root of ω , while condition (b) expresses the essential feature of a tree. Condition (c) says that our ω is what is called an ordered tree in combinatorics. The integer $\nu_u(\omega)$ is the number of children of the vertex $u \in \omega$, and we follow Otter in calling it the ‘type’ of the vertex u . For notational convenience, we will sometimes write $\nu_\omega(u)$ instead of $\nu_u(\omega)$.

For each $u \in U$, let $\Omega_u = \{\omega \in \Omega : \omega \ni u\}$ be the set of trees containing u as their common vertex, and let \mathcal{F} be the σ -field generated by the family $\{\Omega_u, u \in U\}$. For $n = 1, 2, \dots$, we let \mathcal{F}_n be the σ -field generated by Ω_u s with $|u| \leq n$, $|u|$ being the length of the sequence u . For each $u \in U$, we also define the translation $T_u : \Omega_u \rightarrow \Omega$ by $T_u(\omega) = \{v \in U : uv \in \omega\}$.

Now, let $(\Omega^*, \mathcal{F}^*, P^*) = \prod_{u \in U} (\mathbb{Z}_+, \Pi)$ be the product over U of the discrete probability space (\mathbb{Z}_+, Π) , the σ -field \mathcal{F}^* being generated by the coordinate maps $v_u^*, u \in U$, and let the measurable map $\psi : (\Omega^*, \mathcal{F}^*) \rightarrow (\Omega, \mathcal{F})$ be defined by

$$\psi(\omega^*) = \{u = j_1 \cdots j_p : j_{k+1} \leq v_{j_1 \cdots j_k}^*(\omega^*), 0 \leq k < p\}.$$

If we define $P := P^* \circ \psi^{-1}$ then, for each n , $\{T_u(\omega), u \in \omega, |u| = n\}$ are independent under the conditional distribution $P(\cdot | \mathcal{F}_n)$, and obey the same law P . Moreover, if we set $X_n(\omega) := \sharp\{u \in \omega : |u| = n\}$, $n = 0, 1, 2, \dots$, then $\{X_n\}_{n \geq 0}$ forms the Galton–Watson process with offspring distribution Π such that $X_0 = 1$. See [23] for the precise statement of this proposition. Here, $\sharp B$ means the cardinality of a finite set B .

In this setting, our Z and Y_k can be expressed as

$$Z(\omega) = \sum_{u \in U} \mathbf{1}_{\Omega_u}(\omega) \quad \text{and} \quad Y_k(\omega) = \sharp\{u \in \omega : \nu_u(\omega) = k\} = \sum_{u \in U} \mathbf{1}_{\Omega_{uk} \setminus \Omega_{u(k+1)}}(\omega),$$

so that they are certainly \mathcal{F} -measurable. For $k = 0$, uk in the second formula is understood to mean $u\phi = u$.

Since we are interested in the probability distribution of Z and Y_k , we will assume throughout this work that $p_0 > 0$, in which case we have $Z < \infty$ with positive probability. We also assume that $p_0 + p_1 < 1$ in order to avoid the trivial case, which can be treated separately if necessary.

The outline of the paper is as follows. In Section 2, we consider the asymptotic behavior of $P(Z = n)$ as $n \rightarrow \infty$. Under a certain condition on the generating function of Π , we will

show that $P(Z = n)$ has an asymptotic expansion as $n \rightarrow \infty$. If we take its main term, we can reproduce Otter’s result (Theorem 4 of [24]). We also give some partial results when the above-mentioned condition fails to hold. In Section 3, we show that $\mathcal{Y}_k(\omega)$ can be viewed as the total progeny of a random tree structure, which we will call an abstract Galton–Watson tree and which is hidden in the original tree ω . We then investigate the asymptotic behavior of $P(\mathcal{Y}_k = n)$ as $n \rightarrow \infty$, for $k = 0, 1, 2, \dots$. In Section 4, we give some explicit formulae for the joint distribution of $Y_k, k \geq 0$, conditioned on the event $\{Z = n\}$. We then consider its limiting behavior as $n \rightarrow \infty$, in Section 5, and prove a central limit theorem for the Y_k .

2. Asymptotic results for $P(Z = n)$

2.1. Preliminaries

Let $f(z) = \sum_{n=0}^{\infty} p_n z^n$ be the generating function of the offspring distribution $\Pi = \{p_n\}_{n=0}^{\infty}$, and let $\rho \geq 1$ be its radius of convergence. As announced in the Introduction, we assume that $p_0 > 0$ and $p_0 + p_1 < 1$, so that $f(0) > 0$ and $f(z)$ is strictly convex on $[0, \rho)$. Furthermore, let $\mathcal{P}(z) = \sum_{n=1}^{\infty} P(Z = n)z^n = E[z^Z]$ be the generating function of the total progeny Z . Note that $Z \geq 1$, since a tree ω always contains the root ϕ . From the relation

$$Z(\omega) = 1 + \sum_{j=1}^{v_\phi(\omega)} \sum_{u \in U} \mathbf{1}_{\Omega_{ju}}(\omega) = 1 + \sum_{j=1}^{v_\phi(\omega)} Z \circ T_j(\omega) \tag{1}$$

and the conditional independence, given \mathcal{F}_1 , of $Z \circ T_j, j = 1, \dots, v_\phi(\omega)$, we obtain

$$\begin{aligned} \mathcal{P}(z) &= \sum_{k=0}^{\infty} P(v_\phi = k) E[z^{1+Z \circ T_1 + \dots + Z \circ T_k} \mid v_\phi = k] \\ &= z \sum_{k=0}^{\infty} p_k (E[z^Z])^k = z f(\mathcal{P}(z)). \end{aligned}$$

Thus, $w = \mathcal{P}(z)$ satisfies the functional equation $w = z f(w)$. Since we are assuming that $f(0) \neq 0$, we can apply the Lagrange inversion formula (see, e.g. [16]) to show that $P(Z = n)$, the n th coefficient of the power series $\mathcal{P}(z)$, is given by

$$P(Z = n) = \frac{1}{n} \operatorname{Res} \left[\left(\frac{f(z)}{z} \right)^n \right] = \frac{1}{2\pi i n} \oint \left(\frac{f(z)}{z} \right)^n dz, \tag{2}$$

where \oint denotes contour integration along a circle surrounding the origin. Equation (2) shows that $P(Z = n)$ is the coefficient of z^{n-1} of the power series $f(z)^n$, divided by n . However, $f(z)^n$ is the generating function of the n th convolution power Π^{*n} of Π and, hence, $n P(Z = n)$ is equal to the probability that the first generation of our Galton–Watson process consists of $n - 1$ individuals when there are n individuals in the zeroth generation. In other words, if we let $\{q(n, m)\}_{n,m \geq 0}$ be the transition probability of the Galton–Watson process with offspring distribution Π , we then have

$$P(Z = n) = \frac{1}{n} q(n, n - 1). \tag{3}$$

By a purely probabilistic argument, Dwass [9] (see also [17]) obtained this seemingly nontrivial relation. The basic functional equation $\mathcal{P}(z) = z f(\mathcal{P}(z))$ seems to date back to

the work of Hawkins and Ulam (see the footnote on page 32 of [14]), and was independently rediscovered by Good [10] and Otter [24], the latter noting (2) at the same time. This was also rediscovered by Boyd [3], who used it to give an alternative proof of (3).

The class of distributions on natural numbers that are obtained as distributions of the total progeny of Galton–Watson trees coincides with the class of ‘Lagrange distributions’ treated systematically by Consul and Shenton [4]. Good [11] generalized the Lagrange inversion formula to multivariable functions, and discussed its relation to multitype branching processes. (See also [12].)

Now, let $\alpha \geq 1$ be the radius of convergence of the power series $\mathcal{P}(z)$, and let $a = \mathcal{P}(\alpha-)$. We then have the following relations.

Proposition 1. *Under the conditions $p_0 > 0$ and $p_0 + p_1 < 1$, we have*

- (i) $a \leq \rho$,
- (ii) α and a are finite,
- (iii) $f(a)/a = 1/\alpha$,
- (iv) $f'(a) \leq 1/\alpha$,
- (v) $f'(a) = 1/\alpha$ if $a < \rho$, and
- (vi) if, conversely, $f'(\zeta) = f(\zeta)/\zeta$ for some $0 < \zeta < \rho$, then $a < \rho$ and $\zeta = a$.

Proof. All the statements except (vi) are proved in Lemma 4 and Theorem 3 of [24]. To show (vi), suppose that $a = \rho$, in which case ρ is finite. Then, as z increases from 0 to α , $\mathcal{P}(z)$ strictly increases from 0 to ρ . Hence, there exists a $z_0 \in (0, \alpha)$ such that $\zeta = \mathcal{P}(z_0)$. If we put $z = z_0$ in the formulae $zf(\mathcal{P}(z)) = \mathcal{P}(z)$ and $zf'(\mathcal{P}(z))\mathcal{P}'(z) + f(\mathcal{P}(z)) = \mathcal{P}'(z)$, which is obtained by differentiating the former with respect to z , then, by $f'(\zeta) = f(\zeta)/\zeta$, we obtain $\mathcal{P}'(z_0) + f(\zeta) = \mathcal{P}'(z_0)$, whence $f(\zeta) = 0$. However, this is a contradiction, because $f(z)$ is increasing and $f(0) = p_0 > 0$. Thus, we must have $a < \rho$ and, by (iii) and (v), $f'(z) = f(z)/z$ is satisfied also at $z = a$. However, since $(zf'(z) - f(z))' = zf''(z) > 0$ on $(0, \rho)$, this z is unique, so that $\zeta = a$.

From now on, we will refer to the condition contained in statement (vi) as *Condition A*. Since $(zf'(z) - f(z))|_{z=0} = -p_0 < 0$, it is obviously equivalent to the condition that

$$\lim_{z \uparrow \rho} (zf'(z) - f(z)) > 0. \tag{4}$$

It is also easy to see that Condition A holds if and only if the function $f(z)/z$ attains a unique minimum in $(0, \rho)$, which is equal to $f(a)/a$.

When, on the other hand, Condition A fails, we have $\rho = a < \infty$, and $f(z)/z$ is decreasing on $(0, a]$. Hence, the number a is always characterized by the relation

$$\frac{f(a)}{a} = \inf_{0 < z < \rho} \frac{f(z)}{z}, \tag{5}$$

whether or not Condition A holds.

The following proposition shows some trivial a-priori estimates for $P(Z = n)$.

Proposition 2. (i) Let $d = d(\Pi)$ be the greatest common divisor of all $n \geq 0$ such that $p_n > 0$. Then $P(Z = n) = 0$ if $n \not\equiv 1 \pmod d$.

(ii) We have

$$\limsup_{n \rightarrow \infty} P(Z = n)^{1/n} = \frac{f(a)}{a},$$

where a is given by (5).

Proof. (i) This is part of Theorem 4 of [24]. Here we give an alternative proof, to emphasize that it has nothing to do with Condition A.

In the expansion

$$\left(\frac{f(z)}{z}\right)^n = \sum_{j_1, j_2, \dots, j_n \geq 0} p_{j_1} p_{j_2} \cdots p_{j_n} z^{j_1 + j_2 + \dots + j_n - n},$$

we have $p_{j_1} p_{j_2} \cdots p_{j_n} > 0$ only when all of j_1, j_2, \dots, j_n are integer multiples of d . Hence, if $n \not\equiv 1 \pmod d$ we then have $j_1 + j_2 + \dots + j_n - n \neq -1$ for any such j_1, j_2, \dots, j_n , so that the residue at the origin of $(f(z)/z)^n$ vanishes and so, by (2), does $P(Z = n)$. This proves part (i). The assertion (ii) is obvious from the Cauchy–Hadamard formula for the radius of convergence of power series and from part (iii) of Proposition 1. Note that we have $0 < f(a)/a \leq 1$, since $1 \leq \alpha < \infty$.

2.2. Asymptotic expansion of $P(Z = n)$ as $n \rightarrow \infty$

The following is an improvement of Theorem 4 of [24].

Theorem 1. Under Condition A, $P(Z = n)$ has an asymptotic expansion of the form

$$P(Z = n) \sim \sum_{k=1}^{\infty} c_k \left(\frac{f(a)}{a}\right)^n n^{-k-1/2}$$

as we let $n \rightarrow \infty$ keeping $n \equiv 1 \pmod d$. Here, a and d are the numbers discussed in Propositions 1 and 2, respectively, and, in particular,

$$c_1 = d \sqrt{\frac{f(a)}{2\pi f''(a)}}.$$

Proof. Throughout the proof, we assume that $n \equiv 1 \pmod d$. Condition A means that we can take the contour $|z| = a$ in the integral of (2). Noting that $f(z)$ depends only on z^d , we can rewrite (2) as

$$P(Z = n) = \frac{ad}{2\pi n} \int_{-\pi/d}^{\pi/d} \exp(n\psi(\theta)) e^{i\theta} d\theta, \tag{6}$$

where we let $\psi(\theta) = \log[f(a e^{i\theta})/a e^{i\theta}]$. From $f'(a) = f(a)/a$, it is easy to verify that $\psi(0) = \log[f(a)/a]$, $\psi'(0) = 0$, and $\psi''(0) = -a^2 f''(a)/f(a) < 0$. Thus, we are in a typical situation in which the Laplace method yields the asymptotic expansion (see [5] and [6]). Since the technique is standard, we will only sketch the outline, omitting the details.

Pick a $\delta \in (0, \pi/d)$ sufficiently small that, for $|\theta| \leq \delta$, we have the inequality $\operatorname{Re} \psi(\theta) \leq -\eta\theta^2$, for some $\eta > 0$, and the expansion

$$\psi(\theta) = \log \frac{f(a)}{a} - K\theta^2 + \sum_{k=3}^{\infty} B_k \theta^k,$$

where we have let $\psi''(0) = -2K$. In fact, if we let $\rho_k := p_k a^k / f(a)$ then $\Pi_a := \{\rho_k\}_{k \geq 0}$ is a probability distribution on \mathbb{Z}_+ whose characteristic function is $f(a e^{i\theta}) / f(a)$. Hence, if κ_k is the k th cumulant of the distribution Π_a , we then have the relation $B_k = i^k \kappa_k / k!$.

It is easy to show that $|f(a e^{i\theta})| < f(a)$ for $\theta \in [-\pi/d, \pi/d] \setminus \{0\}$ (see the lemma below), so that

$$\Delta_\delta := \sup_{\delta \leq |\theta| \leq \pi/d} \left| \frac{f(a e^{i\theta})}{a e^{i\theta}} \right| < \frac{f(a)}{a}. \tag{7}$$

Hence,

$$\begin{aligned} P(Z = n) &= \frac{ad}{2\pi n} \int_{-\delta}^{\delta} \exp\{n\psi(\theta)\} e^{i\theta} d\theta + \mathcal{O}(n^{-1} \Delta_\delta^n) \\ &= \frac{ad}{2\pi n} \left(\frac{f(a)}{a} \right)^n \int_{-\delta}^{\delta} \exp(-nK\theta^2) \exp(i\theta + n\theta^3 \beta(\theta)) d\theta + \mathcal{O}(n^{-1} \Delta_\delta^n) \end{aligned} \tag{8}$$

for large n , where we have set $\beta(\theta) := \sum_{k \geq 3} B_k \theta^{k-3}$.

Following [6], we introduce the power series

$$P(\omega, \theta) = \exp(i\theta + \omega\beta(\theta)) = \sum_{\ell, m \geq 0} c_{\ell m} \omega^\ell \theta^m$$

and the polynomial

$$P_A(\omega, \theta) = \sum_{\ell+m \leq A} c_{\ell m} \omega^\ell \theta^m.$$

We then have

$$P(\omega, \theta) - P_A(\omega, \theta) = \mathcal{O}(\omega^{A+1}) + \mathcal{O}(\theta^{A+1})$$

as $|\omega| + |\theta| \rightarrow 0$.

We now replace δ by $\tau_n := n^{-1/3}$ in the integral in (8). The resulting error can be estimated to be $\mathcal{O}(n^{-2/3} e^{-\eta n^{1/3}})$. Then we insert $P_A(n\theta^3, \theta)$ in place of $\exp(i\theta + n\theta^3 \beta(\theta)) = P(n\theta^3, \theta)$. The bound of the error introduced by this replacement is given by

$$\begin{aligned} &\frac{ad}{2\pi n} \left(\frac{f(a)}{a} \right)^n \times \mathcal{O} \left(\int_0^{\tau_n} \exp(-nK\theta^2) \{(n\theta^3)^{A+1} + \theta^{A+1}\} d\theta \right) \\ &= \mathcal{O} \left(\frac{1}{n} \left(\frac{f(a)}{a} \right)^n n^{-1-A/2} \right), \end{aligned}$$

which absorbs the other, smaller error terms. Thus, for any $A \geq 0$, we have

$$P(Z = n) = \frac{ad}{2\pi n} \left(\frac{f(a)}{a} \right)^n \left[\int_{-\tau_n}^{\tau_n} \exp(-nK\theta^2) P_A(n\theta^3, \theta) d\theta + \mathcal{O}(n^{-1-A/2}) \right].$$

Finally, we extend the integral to the whole real line with the error, which can be estimated to be $\mathcal{O}(e^{-\alpha n^{1/3}})$ for some $\alpha > 0$. Thus, we arrive at

$$\begin{aligned}
 P(Z = n) &= \frac{ad}{2\pi n} \left(\frac{f(a)}{a}\right)^n \left[\sum_{\ell+m \leq A} c_{\ell m} n^\ell \int_{-\infty}^{\infty} \exp(-nK\theta^2) \theta^{3\ell+m} d\theta + \mathcal{O}(n^{-1-A/2}) \right] \\
 &= \frac{ad}{2\pi} \left(\frac{f(a)}{a}\right)^n \left[\sum_{\ell+m \leq A} \gamma_{\ell m} K^{-(3\ell+m+1)/2} n^{-1-(\ell+m+1)/2} + \mathcal{O}(n^{-2-A/2}) \right], \tag{9}
 \end{aligned}$$

where $\gamma_{\ell m} = 0$ if $\ell + m$ is odd, and $\gamma_{\ell m} = \Gamma((3\ell + m + 1)/2)c_{\ell m}$ if $\ell + m$ is even. If we take $A = 2k - 1, k = 1, 2, \dots$, then the error term becomes $\mathcal{O}(n^{-k-3/2})$, while the final term in the sum is a constant multiple of $n^{-k-1/2}$. The validity of (9) for all $A > 0$ thus implies that we have the asymptotic expansion

$$P(Z = n) \sim \frac{ad}{2\pi} \sum_{k=1}^{\infty} \left(\sum_{\ell+m=2(k-1)} \gamma_{\ell m} K^{-(3\ell+m+1)/2} \right) \left(\frac{f(a)}{a}\right)^n n^{-k-1/2}$$

as $n \rightarrow \infty$. The coefficient for $k = 1$ is

$$\frac{ad}{2\pi} \gamma_{00} K^{-1/2} = \frac{ad}{2\pi} c_{00} \Gamma\left(\frac{1}{2}\right) \left(\frac{a^2 f''(a)}{2f(a)}\right)^{-1/2} = d \sqrt{\frac{f(a)}{2\pi f''(a)}}.$$

Lemma 1. *If $0 < p_0 < 1$ we then have $|f(r e^{i\theta})| < f(r)$, given that $0 < r < \rho$ and $0 < |\theta| \leq \pi/d$.*

Proof. Let $\mathcal{A} := \{k \geq 1 : p_k > 0\}$. Since we have, for any $k \in \mathcal{A}$, the inequality

$$|f(r e^{i\theta})| \leq |p_0 + p_k r^k e^{ik\theta}| + \sum_{n \neq 0, k} p_n r^n,$$

where

$$|p_0 + p_k r^k e^{ik\theta}|^2 = p_0^2 + p_k^2 r^{2k} + 2p_0 p_k r^k \cos k\theta,$$

all we have to do is to verify that $\cos k\theta < 1$ for some $k \in \mathcal{A}$. When $\mathcal{A} = \{d\}$, this is obvious from the assumption on θ . When \mathcal{A} contains at least two numbers, take distinct $k, k' \in \mathcal{A}$ such that $\gcd(k, k') = d$. Then, there are integers m and m' such that $mk + m'k' = d$. Now, if we had $\cos k\theta = \cos k'\theta = 1$, then $k\theta = 2\pi\ell$ and $k'\theta = 2\pi\ell'$ for some integers ℓ and ℓ' . Hence, we would have $\theta d = \theta(mk + m'k') \in 2\pi\mathbb{Z}$, contradicting the assumption on θ . Thus, either $\cos k\theta < 1$ or $\cos k'\theta < 1$.

2.3. Some partial results when Condition A fails

When Condition A fails to hold, we have, in view of Proposition 1, $\rho = a < \infty$ and $f(\rho) = f(a) = a/\alpha < \infty$. We also have $f'(\rho) = f'(a) \leq 1/\alpha < \infty$; in fact, $f'(\rho) \leq f(\rho)/\rho$. On the other hand, $f(\rho)/\rho = f(a)/a \leq 1/\alpha \leq 1$ since $\alpha \geq 1$. In particular, $f(z) = f(x + iy)$ is C^1 with respect to (x, y) on the closed disk $\{|z| \leq \rho\}$ and, hence, we can take the circle $\{|z| = \rho\}$ as the contour of integration in (2), so that

$$P(Z = n) = \frac{\rho d}{2\pi n} \int_{-\pi/d}^{\pi/d} \exp(n\psi(\theta)) e^{i\theta} d\theta,$$

where we let $\psi(\theta) = \log[f(\rho e^{-i\theta})/\rho e^{i\theta}]$. We will consider two cases separately.

2.3.1. *The case $f'(\rho) = f(\rho)/\rho$.* In this case, we will assume that $f^{(4)}(\rho) < \infty$. As before, we have $\psi'(0) = 0$ and

$$-2K := \psi''(0) = \rho^2 \frac{f''(\rho)}{f(\rho)} < 0,$$

so that

$$\psi(\theta) = \psi(0) - K\theta^2 + i\beta\theta^3 + \phi(\theta), \tag{10}$$

where $i\beta = \psi^{(3)}(0)/3!$, $\beta \in \mathbb{R}$, and $\phi(\theta) = \mathcal{O}(\theta^4)$ as $\theta \rightarrow 0$. (The function $\phi(\cdot)$ should not be confused with the symbol for the empty sequence.) Moreover, $\psi(0) = \log[f(\rho)/\rho]$ is the maximum value of $\text{Re } \psi(\theta)$. Now we can proceed as in Section 2.2, to obtain

$$\begin{aligned} P(Z = n) &= \frac{\rho d}{2\pi n} \left\{ \int_{-\delta}^{\delta} \exp(n\psi(\theta)) e^{i\theta} d\theta + \mathcal{O}(\Delta_{\delta}^n) \right\} \\ &= \frac{\rho d}{2\pi n} \left\{ \int_{-\tau_n}^{\tau_n} \exp(n\psi(\theta)) e^{i\theta} d\theta + \mathcal{O}\left(\left(\frac{f(\rho)}{\rho} \right)^n \int_{\tau_n}^{\delta} e^{-\eta n\theta^2} d\theta \right) + \mathcal{O}(\Delta_{\delta}^n) \right\}, \end{aligned}$$

where $\delta, \Delta_{\delta}, \tau_n$, and η have the same meaning as before. Since

$$\int_{\tau_n}^{\delta} e^{-\eta n\theta^2} d\theta = \mathcal{O}(n^{-2/3} e^{-\eta n^{1/3}}),$$

it absorbs the error term $\mathcal{O}(\Delta_{\delta}^n)$. Hence, taking (10) and $\tau_n = n^{-1/3}$ into account, we can write

$$\begin{aligned} P(Z = n) &= \frac{\rho d}{2\pi n} \left(\frac{f(\rho)}{\rho} \right)^n \left\{ \int_{-\tau_n}^{\tau_n} \exp(-Kn\theta^2 + i\beta n\theta^3 + n\phi(\theta)) e^{i\theta} d\theta + \mathcal{O}(n^{-2/3} e^{-\eta n^{1/3}}) \right\} \\ &= \frac{\rho d}{2\pi n} \left(\frac{f(\rho)}{\rho} \right)^n \left\{ \int_{-\tau_n}^{\tau_n} \exp(-Kn\theta^2 + i\beta n\theta^3) e^{i\theta} d\theta \right. \\ &\quad \left. + \mathcal{O}\left(\int_0^{\tau_n} e^{-Kn\theta^2} n\phi(\theta) d\theta \right) + \mathcal{O}(n^{-2/3} e^{-\eta n^{1/3}}) \right\}. \end{aligned}$$

From $\phi(\theta) = \mathcal{O}(\theta^4)$, we obtain the estimate

$$\int_0^{\tau_n} e^{-Kn\theta^2} n\phi(\theta) d\theta = \mathcal{O}(n^{-3/2}).$$

On the other hand,

$$\begin{aligned} &\int_{-\tau_n}^{\tau_n} \exp(-Kn\theta^2 + i\beta n\theta^3 + i\theta) d\theta \\ &= \int_{-\tau_n}^{\tau_n} e^{-Kn\theta^2} \cos(\beta n\theta^3 + \theta) d\theta \\ &= \int_{-\tau_n}^{\tau_n} e^{-Kn\theta^2} d\theta + \mathcal{O}\left(\int_{-\tau_n}^{\tau_n} e^{-Kn\theta^2} (n^2\theta^6 + \theta^2) d\theta \right) \\ &= \frac{1}{\sqrt{n}} \int_{-\infty}^{\infty} e^{-Kz^2} dz + \mathcal{O}(n^{-3/2}) \end{aligned}$$

and, thus,

$$P(Z = n) = \sqrt{\frac{f(\rho)}{2\pi f''(\rho)}} \left(\frac{f(\rho)}{\rho} \right)^n n^{-3/2} (1 + \mathcal{O}(n^{-1})).$$

2.3.2. *The case $f'(\rho) < f(\rho)/\rho$.* In this case, we will assume that $f^{(k)}(\rho) < \infty$ for some $k \geq 3$. As before, $\psi(0) = \log[f(\rho)/\rho]$ is the maximum value of $\operatorname{Re} \psi(\theta)$. Moreover, we have

$$\psi'(0) = i \frac{\rho f'(\rho) - f(\rho)}{f(\rho)} =: -i\beta,$$

with $\beta > 0$, and

$$\psi''(0) = \frac{\rho f'(\rho)}{f(\rho)^2} (\rho f'(\rho) - f(\rho)) - \rho^2 \frac{f''(\rho)}{f(\rho)} < 0.$$

Thus, letting $-2K = \psi''(0)$, we can write

$$\psi(\theta) = \psi(0) - i\beta\theta - K\theta^2 + \phi(\theta),$$

with ϕ a C^k function satisfying $\phi(\theta) = \mathcal{O}(\theta^3)$ and $\phi(0) = \phi'(0) = \phi''(0) = 0$.

We now estimate, with the same notation as before,

$$P(Z = n) = \frac{\rho d}{2\pi n} \left\{ \int_{-\tau_n}^{\tau_n} \exp(n\psi(\theta)) e^{i\theta} d\theta + \mathcal{O}\left(\left(\frac{f(\rho)}{\rho}\right)^n n^{-2/3} e^{-\eta n^{1/3}}\right)\right\}.$$

If we let

$$J_n := \int_{-\tau_n}^{\tau_n} \exp(n\psi(\theta)) e^{i\theta} d\theta$$

then

$$J_n = \left(\frac{f(\rho)}{\rho}\right)^n \frac{1}{\sqrt{n}} \int_{-n^{1/6}}^{n^{1/6}} e^{-i\beta\sqrt{nz}} F_n(z) dz,$$

where

$$F_n(z) := \exp\left[-Kz^2 + n\phi\left(\frac{z}{\sqrt{n}}\right) + i\frac{z}{\sqrt{n}}\right].$$

Now, it is easy to see that we can find constants $A, C > 0$, independent of n , so that, for $j = 0, 1, \dots, k$ and $|z| \leq n^{1/6}$, we have

$$\left| \frac{d^j}{dz^j} F_n(z) \right| \leq A e^{-Cz^2}.$$

Integrating by parts k times, we obtain

$$\begin{aligned} J_n &= \left(\frac{f(\rho)}{\rho}\right)^n \frac{1}{\sqrt{n}} \left\{ \sum_{j=1}^k (-1)^{j-1} \left[\frac{e^{-i\beta\sqrt{nz}}}{(-i\beta\sqrt{n})^j} \frac{d^{j-1}}{dz^{j-1}} F_n(z) \right]_{z=-n^{1/6}}^{z=n^{1/6}} \right. \\ &\quad \left. + (-1)^k \int_{-n^{1/6}}^{n^{1/6}} \frac{e^{-i\beta\sqrt{nz}}}{(-i\beta\sqrt{n})^k} \frac{d^k}{dz^k} F_n(z) dz \right\} \\ &= \left(\frac{f(\rho)}{\rho}\right)^n \frac{1}{\sqrt{n}} \times \mathcal{O}(n^{-k/2}) \end{aligned}$$

and, consequently,

$$P(Z = n) = \mathcal{O}\left(\left(\frac{f(\rho)}{\rho}\right)^n n^{-3/2-k/2}\right)$$

as $n \rightarrow \infty$.

For illustration, consider the case $\rho = 1$, so that $f(\rho)/\rho = 1$. The condition $f'(\rho) < f(\rho)/\rho$ means that our Galton–Watson tree is subcritical. What has just been proved shows that

$$P(Z = n) = \mathcal{O}(n^{-3/2-k/2})$$

if $f^{(k)}(1) < \infty$ ($k \geq 3$). If, for example, $p_n \sim e^{-\delta n^\gamma}$ with $\delta > 0$ and $0 < \gamma < 1$, then $f^{(k)}(1) < \infty$ for all $k \geq 3$ and, consequently, $P(Z = n)$ decays faster than any power of n . However, one still has

$$\limsup_{n \rightarrow \infty} P(Z = n)^{1/n} = 1,$$

by Proposition 2.

3. The distribution of \mathcal{Y}_k

3.1. \mathcal{Y}_k as the total progeny of a hidden Galton–Watson tree

In this section, we consider $\mathcal{Y}_k(\omega) := \sum_{j=0}^k Y_j(\omega)$, the number of those vertices of the Galton–Watson tree ω having at most k children, $\mathcal{Y}_0(\omega)$ being in particular the number of ‘leaves’ of ω .

Let

$$\mathcal{Q}(z) \equiv \mathcal{Q}_k(z) := E[\mathbf{1}_{\{\mathcal{Y}_k < \infty\}} z^{\mathcal{Y}_k}] = \sum_{n=1}^{\infty} P(\mathcal{Y}_k = n) z^n$$

be the generating function of the distribution of \mathcal{Y}_k . (Note that, from our assumption that $p_0 > 0$, we have $P(Y_0 > 0) = 1$. Thus, $P(\mathcal{Y}_k = 0) = 0$ for all $k \geq 0$.)

Since we have

$$\mathcal{Y}_k(\omega) = \mathbf{1}_{\{v_\phi \leq k\}}(\omega) + \sum_{j=1}^{v_\phi(\omega)} (\mathcal{Y}_k \circ T_j)(\omega),$$

we can proceed, just as in Section 2.1, to write

$$\begin{aligned} \mathcal{Q}(z) &= E \left[z^{\mathbf{1}_{\{v_\phi \leq k\}}} \prod_{j=1}^{v_\phi} (\mathbf{1}_{\{\mathcal{Y}_k < \infty\}} \circ T_j)(z^{\mathcal{Y}_k \circ T_j}) \right] \\ &= E \left[z^{\mathbf{1}_{\{v_\phi \leq k\}}} E \left[\prod_{j=1}^{v_\phi} (\mathbf{1}_{\{\mathcal{Y}_k < \infty\}} z^{\mathcal{Y}_k \circ T_j} \mid \mathcal{F}_1) \right] \right] \\ &= E \left[z^{\mathbf{1}_{\{v_\phi \leq k\}}} \prod_{j=1}^{v_\phi} E[\mathbf{1}_{\{\mathcal{Y}_k < \infty\}} z^{\mathcal{Y}_k}] \right] \\ &= z E[\mathbf{1}_{\{v_\phi \leq k\}} \mathcal{Q}(z)^{v_\phi}] + E[\mathbf{1}_{\{v > k\}} \mathcal{Q}(z)^{v_\phi}] \\ &= z \sum_{j=0}^k p_j \mathcal{Q}(z)^j + f(\mathcal{Q}(z)) - \sum_{j=0}^k p_j \mathcal{Q}(z)^j. \end{aligned}$$

Thus, $w = \mathcal{Q}(z)$ solves the equation

$$z \sum_{j=0}^k p_j w^j = w - f(w) + \sum_{j=0}^k p_j w^j,$$

or

$$z = \frac{w}{g_k(w)},$$

where we have set

$$g_k(w) = \frac{w \sum_{j=0}^k p_j w^j}{w - f(w) + \sum_{j=0}^k p_j w^j}.$$

For later use, let us further define

$$h_k(w) = \sum_{j=0}^k p_j w^j \quad \text{and} \quad \varphi_k(w) = \sum_{j>k} p_j w^{j-1}.$$

Then

$$g_k(w) = \frac{h_k(w)}{1 - \varphi_k(w)}. \tag{11}$$

We will now prove that $g_k(w)$ is the generating function of a probability distribution $\Pi^{(k)} := \{p_n^{(k)}\}_{n=0}^\infty$ that satisfies our basic assumptions, namely that $p_0^{(k)} > 0$ and $p_0^{(k)} + p_1^{(k)} < 1$.

To this end, let us write

$$\begin{aligned} g_k(w) &= h_k(w) \sum_{n=0}^\infty \varphi_k(w)^n \\ &= \left(\frac{h_k(w)}{h_k(1)} \right) h_k(1) \sum_{n=0}^\infty \varphi_k(1)^n \left(\frac{\varphi_k(w)}{\varphi_k(1)} \right)^n. \end{aligned}$$

If we set

$$\begin{aligned} H(w) &:= \frac{h_k(w)}{h_k(1)}, \\ G(\zeta) &:= h_k(1) \sum_{n=0}^\infty \varphi_k(1)^n \zeta^n, \\ \psi(w) &:= \frac{\varphi_k(w)}{\varphi_k(1)}, \end{aligned}$$

then H , G , and ψ are the generating functions, respectively, of $P(v_\phi \in \cdot \mid v_\phi \leq k)$, the geometric distribution with parameter $\varphi_k(1) = P(v_\phi > k)$, and $P(v_\phi - 1 \in \cdot \mid v_\phi > k)$. Consequently, if X, N, S_1, S_2, \dots are independent, integer-valued random variables, and if X, N , and S_j have H, G , and ψ as their respective generating functions, then $g_k(w) = G(\psi(w))H(w)$ is the generating function of the random variable

$$\sum_{j=1}^N S_j + X.$$

Obviously, $g_k(0) = p_0$ for $k \geq 1$ and $g_k(0) = p_0/(1 - p_1)$ for $k = 0$, whence $p_0^{(k)} = g_k(0) > 0$. From the facts that

$$g'_k(w) = \frac{h'_k(w)(1 - \varphi_k(w)) + h_k(w)\varphi'_k(w)}{(1 - \varphi_k(w))^2}$$

and $\varphi_k(1) + h_k(1) = 1$, we obtain

$$g'_k(1) = \frac{h'_k(1) + \varphi'_k(1)}{h_k(1)} = 1 + \frac{f'(1) - 1}{h_k(1)}.$$

This shows that $g'_k(1)$ is less than 1, equals 1, or is greater than 1 if and only if $f'(1)$ satisfies the same condition. On the other hand, we have

$$g'_k(0) = \begin{cases} p_0 p_2 / (1 - p_1)^2 & (k = 0), \\ p_1 + p_0 p_2 & (k = 1), \\ p_1 & (k > 1). \end{cases}$$

Hence, for $k > 1$, $p_0^{(k)} + p_1^{(k)} = g_k(0) + g'_k(0) = p_0 + p_1 < 1$, by one of the assumptions. Similarly, for $k = 1$,

$$g_1(0) + g'_1(0) = p_0 + p_1 + p_0 p_2 < 1.$$

To consider the case $k = 0$, let $q_2 = \sum_{n \geq 2} p_n > 0$. Then,

$$\begin{aligned} p_0(1 - p_1 + p_2) &= (1 - p_1 - q_2)(1 - p_1 + p_2) \\ &\leq (1 - p_1 - q_2)(1 - p_1 + q_2) \\ &= (1 - p_1)^2 - q_2^2 \\ &< (1 - p_1)^2 \end{aligned}$$

and, hence,

$$g_0(0) + g'_0(0) = \frac{p_0}{1 - p_1} + \frac{p_0 p_2}{(1 - p_1)^2} = \frac{p_0(1 - p_1 + p_2)}{(1 - p_1)^2} < 1.$$

Finally, if $t > 0$ satisfies $f(t) = t$ then $h_k(t) + t\varphi_k(t) = t$ for any $k \geq 0$, that is,

$$g_k(t) = \frac{h_k(t)}{1 - \varphi_k(t)} = t.$$

These considerations can be summarized in the following theorem.

Theorem 2. *For any $k \geq 0$, the distribution of \mathcal{Y}_k is equal to that of the total progeny of the Galton–Watson tree with offspring distribution $\Pi^{(k)} = \{p_n^{(k)}\}_{n=0}^\infty$. Like the original $\Pi = \{p_n\}_{n=0}^\infty$, $\Pi^{(k)}$ satisfies the conditions $p_0^{(k)} > 0$ and $p_0^{(k)} + p_1^{(k)} < 1$. The new Galton–Watson tree is subcritical, critical, or supercritical if and only if the original one is so. Moreover, the extinction probabilities of both Galton–Watson trees, which are solutions, respectively, of $f(t) = t$ and $g_k(t) = t$, coincide.*

Corollary 1. $P(Z = \infty, \mathcal{Y}_k < \infty) = 0$ for any $k \geq 0$.

Proof. Obviously $\{Z < \infty\} \subset \{\mathcal{Y}_k < \infty\}$, but the probabilities of the events, being the extinction probability of the old and that of the new Galton–Watson tree, respectively, are equal.

We have thus shown the existence of a Galton–Watson tree whose total progeny has the same distribution as \mathcal{Y}_k . This would be sufficient for most applications, but it is natural to ask if the new Galton–Watson tree can be constructed as a functional of the original tree ω , so that its total progeny coincides with $\mathcal{Y}_k(\omega)$. We will now show that this is actually the case, but before doing so we need to formulate an abstract notion of Galton–Watson tree. To begin with, let us generalize the notion of tree itself.

Definition 1. Let $G = (V, E)$ be a directed graph, where V is a finite or countable vertex set and $E \subset V \times V$. We will call G an abstract tree if the following five conditions hold.

1. For any $u \in V$, $(u, u) \notin E$. That is, G has no loops.
2. There is a special vertex $u_0 \in V$ such that $(u, u_0) \notin E$ for all $u \in V$. We call u_0 the ‘root’ of G .
3. For any $v \in V \setminus \{u_0\}$, there is one and only one $u \in V$ such that $(u, v) \in E$. In particular, G is an oriented graph.
4. The (nondirected) graph $G^s = (V, E^s)$, obtained by letting $E^s := \{(u, v) \in E \text{ or } (v, u) \in E\}$, is connected.
5. For any $u \in V$, the set $W_G(u) = \{v \in V : (u, v) \in E\}$ is finite and totally ordered.

Let us first show that an abstract tree $G = (V, E)$ is isomorphic to a tree $\omega \in \Omega$ in the sense of Neveu (and, equivalently, of Otter). From part 4, we see that for an arbitrary $u \in V$, there is a path connecting u_0 and u in the graph G^s . Let $u_0u_1 \cdots u_{n-1}u_n$, where we let $u_n = u$, be such a path that has minimum length. By part 1, we have $u_{i-1} \neq u_i$ for $i = 1, \dots, n$ and, by part 2, we must have $(u_0, u_1) \in E$. Let k be the maximum of those $j = 1, \dots, n$ for which $(u_{i-1}, u_i) \in E$ for all $i < j$. If $k < n$ we then have $(u_{k-1}, u_k) \in E$ and $(u_{k+1}, u_k) \in E$, yielding $u_{k-1} = u_{k+1}$ by virtue of part 3. Then $u_0u_1 \cdots u_{k-1}u_{k+2} \cdots u_n$ would also be a path in G^s of length $n - 2$ connecting u_0 and u , contradicting the minimality of n . Hence, $k = n$ and we must have $(u_{i-1}, u_i) \in E$ for $i = 1, \dots, n$. If there existed another path $u_0v_1 \cdots v_{n-1}v_n$, where $v_n = u$, connecting u_0 and u in G^s then, by the same argument, we would have $(u_0, v_1) \in E, (v_1, v_2) \in E, \dots, (v_{n-1}, v_n) \in E$. Hence, we must have $v_j = u_j$ for $j = n - 1, n - 2, \dots, 1$, by part 3.

Thus, the shortest path connecting u_0 and u in G^s is unique, and connects u_0 and u also in G . By part 3, it is actually the unique path connecting u_0 and u in G . Thus, for each $u \in V$, there corresponds a unique path $u_0u_1 \cdots u_n$ ($u_n = u$) in G . Since $u_i \in W_G(u_{i-1})$ and since $W_G(u_{i-1})$ is in one-to-one correspondence with the set $\{1, 2, \dots, v_G(u_{i-1})\}$, where we set $v_G(v) = \sharp W_G(v)$ for $v \in V$, this path can be represented without ambiguity by a sequence of integers $j_1j_2 \cdots j_n \in U$. If we write $\tau_G(u) = j_1j_2 \cdots j_n$ for $u \in V$, then $\omega = \tau_G(V) \subset U$ obviously satisfies the conditions (a), (b), and (c) stated in the Introduction, and so $\omega \in \Omega$ and τ_G is a graph isomorphism between two trees G and ω .

In order to randomize the notion of abstract tree, we fix a nonempty countable set S and let \mathbf{G}_S be the totality of abstract trees $G = (V, E)$ such that $V \subset S$. Define a mapping $\tau : \mathbf{G}_S \rightarrow \Omega$ by letting $\tau(G) = \tau_G(V)$ for $G = (V, E) \in \mathbf{G}_S$, and let \mathcal{G} be the σ -field on \mathbf{G}_S generated by the mapping τ .

Now let $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ be a probability space. A random S -tree is by definition a mapping $G : \tilde{\Omega} \rightarrow \mathbf{G}_S$ which is $\tilde{\mathcal{F}}/\mathcal{G}$ -measurable. This condition is obviously equivalent to saying that the composition $\tau \circ G : \tilde{\Omega} \rightarrow \Omega$ is $\tilde{\mathcal{F}}/\mathcal{F}$ -measurable. When the image $P = \tilde{P} \circ (\tau \circ G)^{-1}$ is the distribution of a Galton–Watson tree, we will say that our random S -tree $G(\cdot)$ is an abstract Galton–Watson tree.

Returning to our Galton–Watson tree (Ω, \mathcal{F}, P) with offspring distribution $\Pi = \{p_n\}_{n=0}^\infty$, we first note that our assumption $p_0 > 0$ implies that $P(\bigcap_{n=1}^\infty \Omega_{u1_n}) = 0$ for each $u \in U$, where we denote by $1_n = 11 \cdots 1$ the sequence consisting of n consecutive 1s. Hence, if we let $\tilde{\Omega} = \bigcap_{u \in U} \bigcup_{n=1}^\infty (\Omega_{u1_n})^c$ we then have $\tilde{\Omega} \in \mathcal{F}$ and $P(\tilde{\Omega}) = 1$. Restricting \mathcal{F} and P to $\tilde{\Omega}$, we obtain a new probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$.

For $k \geq 0$ and for each $\omega \in \tilde{\Omega}$, set

$$V_k(\omega) = \{u \in U : u \in \omega \text{ and } v_\omega(u) \leq k\} \subset U.$$

Then $\mathcal{Y}_k(\omega) = \sharp V_k(\omega)$. We are now going to show that, by suitably defining the adjacency relation on $V_k(\omega)$, this can be turned into an abstract Galton–Watson tree whose offspring distribution is $\Pi^{(k)}$. To make the idea transparent, while avoiding notational complications, we will give the full details of its construction only for $k = 0$, in which case $V_0(\omega)$ is the set of end points, or ‘leaves’, of ω , and give a brief sketch for the general case $k \geq 1$.

To begin with, let us make $V_0(\omega)$ into an abstract tree, which we will denote $G_0(\omega)$, in the following way. Since $\omega \in \tilde{\Omega}$, we have

$$n_0(\omega) := \max\{p \geq 0 : 1_p \in \omega\} < \infty.$$

We write $u_0(\omega) = 1_{n_0(\omega)}$, which will play the role of the root of $G_0(\omega)$. Now, if $n_0(\omega) = 0$ or $v_\omega(1_p) = 1$ for all $p < n_0(\omega)$, then $u_0(\omega)$ has no children and $V_0(\omega) = \{u_0(\omega)\}$. If $v_\omega(1_p) > 1$ for some $0 \leq p < n_0(\omega)$, then those vertices $u \in V_0(\omega)$ of the form $u = 1_p a 1_m$, where $0 \leq p < n_0(\omega)$, $1 < a \leq v_\omega(1_p)$, and $m = \max\{j \geq 0 : 1_p a 1_j \in \omega\}$, are by definition children of $u_0(\omega)$. In general, $u \in V_0(\omega)$ has the form

$$u = 1_{p_1} a_1 1_{p_2} a_2 \cdots 1_{p_N} a_N 1_m =: \tilde{u} 1_m, \tag{12}$$

where $p_i \geq 0$, $i = 1, \dots, N$, $1 < a_i \leq v_\omega(1_{p_1} a_1 \cdots 1_{p_i})$, and $m = \max\{j \geq 0 : \tilde{u} 1_j \in \omega\}$. When $m = 0$ or $v_\omega(\tilde{u} 1_j) = 1$ for all $0 \leq j < m$, the vertex u has no children, while in the case $v_\omega(\tilde{u} 1_j) > 1$ for some $0 \leq j < m$, it has children of the form $v = \tilde{u} 1_i b 1_\ell \in V_0(\omega)$, where $0 \leq i < m$, $1 < b \leq v_\omega(\tilde{u} 1_i)$, and $\ell = \max\{p \geq 0 : \tilde{u} 1_i b 1_p \in \omega\}$. If we let $E_0(\omega)$ be the totality of such parent–child pairs (u, v) , then the resulting directed graph $G_0(\omega) = (V_0(\omega), E_0(\omega))$ obviously satisfies parts 1 and 2 of Definition 1. To verify part 3, consider

$$v = 1_{p_1} a_1 \cdots 1_{p_N} a_N 1_i b 1_\ell =: \tilde{v} 1_i b 1_\ell \in V(\omega) \tag{13}$$

with $N \geq 0$. Since $v \in \omega$ and since ω is a tree in the sense of Neveu, we have $\tilde{v} 1_i b \in \omega$. Replacing b by 1, we also have $\tilde{v} 1_{i+1} \in \omega$. Hence, if we let $m = \max\{p \geq 0 : \tilde{v} 1_p \in \omega\}$ then $0 < m < \infty$ and $u := \tilde{v} 1_m \in V_0(\omega)$ is the uniquely determined parent of v , i.e. $(u, v) \in E_0(\omega)$.

If $N > 0$ in (13) then, by the same procedure, we can in turn find the parent of u , and so on. Thus, starting from a vertex $v \in V_0(\omega)$ of the form (13), we can find u_N, u_{N-1}, \dots, u_1 such that $(u_{i-1}, u_i) \in E_0(\omega)$ for $i = 1, \dots, N + 1$, where we set $u_{N+1} = v$. Since this is true for all $v \in V_0(\omega)$, we see that part 4 of Definition 1 holds, too. Finally, for each $u \in V_0(\omega)$, let $W_\omega(u) \subset V_0(\omega)$ be the totality of its children. If u is of the form (12), i.e. $u = \tilde{u} 1_m$, then from the consideration above, we have

$$\sharp W_\omega(u) = \sum_{i=0}^m (v_\omega(\tilde{u} 1_i) - 1)_+ < \infty. \tag{14}$$

Moreover, we can arrange the members of $W_\omega(u)$ into natural lexicographic order, under which $W_\omega(u)$ is totally ordered. Namely, given two typical vertices $v = \tilde{u} 1_i b 1_\ell$ and $v' = \tilde{u}' 1_{i'} b' 1_{\ell'}$ of $W_\omega(u)$, we define $v < v'$ if $i < i'$ or if $i = i'$ and $b < b'$. Thus we have shown that $G_0(\omega) = (V_0(\omega), E_0(\omega))$ is an abstract tree for each $\omega \in \tilde{\Omega}$.

For example, consider the tree $\omega \in \Omega$ shown in Figure 1, for which $\mathcal{Y}_0(\omega) = 4$. The corresponding tree $G_0(\omega)$ is isomorphic to the one shown in Figure 2. In Figure 3, it is shown how the tree $G_0(\omega)$ grows in the original tree ω .

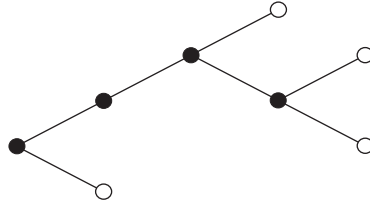


FIGURE 1: The tree ω .

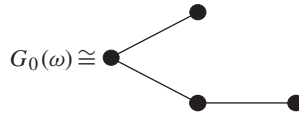


FIGURE 2: The tree $G_0(\omega)$.

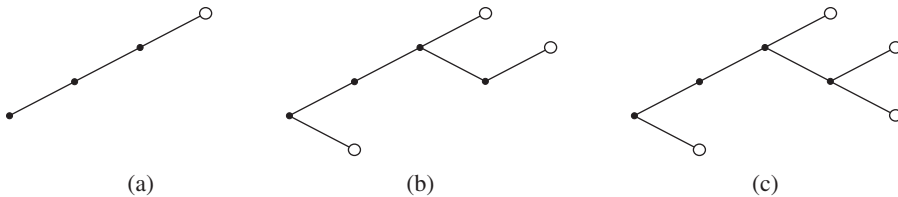


FIGURE 3: How $G_0(\omega)$ grows in ω . Diagram (a) shows the zeroth generation; diagram (b) the zeroth and first generations; and diagram (c) the zeroth, first, and second generations.

Thus, we have constructed a mapping $G_0(\omega)$ defined for $\omega \in \tilde{\Omega}$ and taking its values in \mathbf{G}_U , the totality of abstract trees whose vertex sets are subsets of U . We will now show that $G_0: \tilde{\Omega} \rightarrow \mathbf{G}_U$ is a random U -tree in the sense previously defined. This amounts to showing that $(\tau \circ G_0)^{-1}(\Omega_u) \in \tilde{\mathcal{F}}$ for each $u \in U$. This we will do by induction on the length of the sequence u .

First consider the case $|u| = 1$, i.e. suppose that $u = j$ for some $j \in \mathbb{N}$. Then,

$$(\tau \circ G_0)^{-1}(\Omega_j) = \{\omega \in \tilde{\Omega}: \sharp W_\omega(u_0(\omega)) \geq j\}.$$

By letting $m = n_0(\omega)$ and $\tilde{u} = \phi$ in (14), we can write, for each $j \geq 1$,

$$\{\omega \in \tilde{\Omega}: \sharp W_\omega(u_0(\omega)) \geq j\} = \bigcup_{n=1}^{\infty} \bigcup_{k_0+\dots+k_{n-1} \geq j, k_i \geq 0} C_{n,k_0,\dots,k_{n-1}}, \tag{15}$$

where, for $n \geq 1$ and $k_0, k_1, \dots, k_{n-1} \geq 0$, we write

$$\begin{aligned} C_{n,k_0,\dots,k_{n-1}} &= \{\omega \in \tilde{\Omega}: n_0(\omega) = n, v_\omega(1_i) = 1 + k_i, i = 0, 1, \dots, n - 1\} \\ &= \{\omega \in \tilde{\Omega}: v_\omega(1_i) = 1 + k_i, i = 0, 1, \dots, n - 1, v_\omega(1_n) = 0\}, \end{aligned} \tag{16}$$

which obviously belongs to $\tilde{\mathcal{F}}$, proving that $(\tau \circ G_0)^{-1}(\Omega_j) \in \tilde{\mathcal{F}}$.

Suppose that we have shown that $(\tau \circ G_0)^{-1}(\Omega_v) \in \tilde{\mathcal{F}}$ for $v \in U$ with $|v| = n - 1$. Then, picking a sequence $u = j_1 j_2 \cdots j_n \in U$ with $j_1 \geq 1$, since we can write $\Omega_u = \{v_\phi \geq j_1\} \cap T_{j_1}^{-1}(\Omega_{j_2 \cdots j_n})$ we have, noting (15), that

$$(\tau \circ G_0)^{-1}(\Omega_u) = \bigcup_{n=1}^{\infty} \bigcup_{k_0 + \cdots + k_{n-1} \geq j_1} C_{n, k_0, \dots, k_{n-1}} \cap (T_{j_1} \circ \tau \circ G_0)^{-1}(\Omega_{j_2 \cdots j_n}).$$

Given $n \geq 1$ and $k_0, k_1, \dots, k_{n-1} \geq 0$ with $\sum_{i=0}^{n-1} k_i \geq j_1$, choose $\ell \geq 1$ and $1 < a \leq 1 + k_\ell$ so that $j_1 = k_0 + \cdots + k_{\ell-1} + (a - 1)$. Then, on the set $C_{n, k_0, \dots, k_{n-1}}$, we obviously have $T_{j_1} \circ \tau \circ G_0 = \tau \circ G_0 \circ T_{1_\ell a}$. Hence, we obtain

$$(\tau \circ G_0)^{-1}(\Omega_u) = \bigcup_{n=1}^{\infty} \bigcup_{k_0 + \cdots + k_{n-1} \geq j_1} C_{n, k_0, \dots, k_{n-1}} \cap \{(T_{1_\ell a})^{-1}((\tau \circ G_0)^{-1}(\Omega_{j_2 \cdots j_n}))\}.$$

However, since $C_{n, k_0, \dots, k_{n-1}} \in \tilde{\mathcal{F}}$, as was already shown; since $(\tau \circ G_0)^{-1}(\Omega_{j_2 \cdots j_n}) \in \tilde{\mathcal{F}}$, from the assumption; and since the translation $T_v : \tilde{\Omega} \cap \Omega_v \rightarrow \tilde{\Omega}$ is measurable, we must have $(\tau \circ G_0)^{-1}(\Omega_u) \in \tilde{\mathcal{F}}$.

Finally, let us prove that the probability measure $Q = \tilde{P} \circ (\tau \circ G_0)^{-1}$ induced on (Ω, \mathcal{F}) is the distribution of the Galton–Watson tree with offspring distribution $\Pi^{(0)} = \{p_n^{(0)}\}_{n=0}^{\infty}$. For this purpose, we note that the following general criterion holds.

Proposition 3. *Let P be a probability distribution on (Ω, \mathcal{F}) and let $\Pi = \{p_n\}_{n=0}^{\infty}$ be that on \mathbb{Z}_+ . Then, for P to be the distribution of the Galton–Watson tree with offspring distribution Π , it is necessary and sufficient that $P(v_\phi = k) = p_k$ for $k \geq 0$ and that*

$$P\left(\bigcap_{j=1}^k T_j^{-1}(B_j) \mid v_\phi = k\right) = \prod_{j=1}^k P(B_j)$$

for any $k \geq 1$ and $B_1, \dots, B_k \in \mathcal{F}$.

Proof. For the proof of necessity, see Section 3 of [23]. To show that the condition is sufficient, let us define, for any finite tree $t \in \Omega$ and for any $\Lambda = (\lambda_e : e \in G_0(t)) \in \mathbb{Z}_+^{G_0(t)}$,

$$[t; \Lambda] := \{\omega \in \Omega : v_u(\omega) = v_u(t) \text{ for } u \in \mathcal{I}(t) \text{ and } v_e(\omega) = \lambda_e \text{ for } e \in G_0(t)\},$$

where $\mathcal{I}(t) := \{u \in t : v_u(t) > 0\}$ is the totality of the ‘inner points’ of the tree t . Subsets of Ω of this form were called ‘neighborhoods’ by Otter [24]. By the equivalence of the two constructions of the probability space for Galton–Watson trees due to Otter and Neveu (see [21]), it is enough to verify that

$$P([t; \Lambda]) = \left(\prod_{u \in \mathcal{I}(t)} p_{v_u(t)}\right) \left(\prod_{e \in G_0(t)} p_{\lambda_e}\right). \tag{17}$$

Given t and $\Lambda = (\lambda_e : e \in G_0(t))$ as above, let $\Lambda_j = (\lambda_e : e \in G_0(T_j t))$ for $j = 1, 2, \dots, v_\phi(t)$, which together form a partition of Λ . Then, it is easily seen that

$$[t; \Lambda] = \{\omega \in \Omega : v_\phi(\omega) = v_\phi(t)\} \cap \bigcap_{j=1}^{v_\phi(t)} T_j^{-1}([T_j t; \Lambda_j])$$

holds. Thus, from the assumption,

$$\begin{aligned} P([t; \Lambda]) &= P(v_\phi = v_\phi(t)) P\left(\bigcap_{j=1}^{v_\phi(t)} T_j^{-1}([T_j t; \Lambda_j]) \mid v_\phi = v_\phi(t)\right) \\ &= p_{v_\phi(t)} \prod_{j=1}^{v_\phi(t)} P([T_j t; \Lambda_j]). \end{aligned}$$

Repeating this argument to rewrite $P([T_j t; \Lambda_j])$, and so on, we arrive at (17).

Let us apply this criterion to our measure Q . From (14) and (16), we see that

$$Q(v_\phi = k) = \sum_{n=0}^{\infty} \sum_{k_0 + \dots + k_{n-1} = k, k_i \geq 0} P(v_\omega(1_i) = 1 + k_i, 0 \leq i < n, \text{ and } v_\omega(1_n) = 0).$$

But our P was defined to be the image, under the mapping ψ , of the product measure $P^* = \Pi^U$. Moreover, on the set

$$\{\omega^* \in \Omega^*: v_{1_i}^*(\omega^*) = 1 + k_i, 0 \leq i < n, v_{1_n}^*(\omega^*) = 0\},$$

which is a subset of $\psi^{-1}(\Omega_{1_n})$, we have $v_{1_i}(\psi(\omega^*)) = v_{1_i}^*(\omega^*)$, $i = 0, 1, \dots, n$. Thus we have, from independence,

$$\begin{aligned} P(v_\omega(1_i) = 1 + k_i, 0 \leq i < n, v_\omega(1_n) = 0) &= P^*(v_{1_i}^* = 1 + k_i, 0 \leq i < n, v_{1_n}^* = 0) \\ &= p_0 \prod_{i=0}^{n-1} p_{1+k_i} \\ &= p_0(1 - p_0)^n \prod_{i=0}^{n-1} P(v_\phi - 1 = k_i \mid v_\phi > 0) \end{aligned}$$

and, hence,

$$Q(v_\phi = k) = \sum_{n=0}^{\infty} p_0(1 - p_0)^n \sum_{k_0 + \dots + k_{n-1} = k} \prod_{i=0}^{n-1} P(v_\phi - 1 = k_i \mid v_\phi > 0) = p_k^{(0)},$$

verifying the first part of the criterion. Next, for $k \geq 1$ and $B_1, \dots, B_k \in \mathcal{F}$, we write

$$\begin{aligned} Q\left(\{v_\phi = k\} \cap \bigcap_{i=1}^k T_i^{-1}(B_i)\right) &= P\left(\{v_\phi \circ \tau \circ G_0 = k\} \cap \bigcap_{i=1}^k (T_i \circ \tau \circ G_0)^{-1}(B_i)\right) \\ &= \sum_{n=1}^{\infty} \sum_{k_0 + \dots + k_{n-1} = k} P\left(C_{n, k_0, \dots, k_{n-1}} \cap \bigcap_{\ell=0}^k \bigcap_{a=2}^{1+k_i} (T_{1_\ell a})^{-1}((\tau \circ G_0)^{-1}(B_{\ell, a}))\right), \end{aligned}$$

where we have written $B_{\ell,a}$ instead of B_j when $j = k_0 + \dots + k_{\ell-1} + (a - 1)$ holds. By the same reasoning as above, we have

$$\begin{aligned} & P\left(C_{n,k_0,\dots,k_{n-1}} \cap \bigcap_{\ell=0}^{n-1} \bigcap_{a=2}^{1+k_i} (T_{1_{\ell a}})^{-1}((\tau \circ G_0)^{-1}(B_{\ell,a}))\right) \\ &= P^*\left(\{v_{1_i}^* = 1 + k_i, 0 \leq i < n, v_{1_n}^* = 0\} \cap \bigcap_{\ell=0}^{n-1} \bigcap_{a=2}^{1+k_i} (T_{1_{\ell a}}^*)^{-1}((\tau \circ G_0 \circ \psi)^{-1}(B_{\ell,a}))\right) \\ &= \left(p_0 \prod_{i=0}^{n-1} p_{1+k_i}\right) \prod_{\ell=0}^{n-1} \prod_{a=2}^{1+k_i} P^*((\tau \circ G_0 \circ \psi)^{-1}(B_{\ell,a})) \\ &= P(C_{n,k_0,\dots,k_{n-1}}) \prod_{j=1}^k Q(B_j). \end{aligned}$$

Here $T_v^*: \Omega^* \rightarrow \Omega$ is the translation on Ω^* defined by $v_v^* \circ T_u^* = v_{uv}^*$ (see [23]). Hence, summing over n and the k_i , we finally obtain

$$Q\left(\{v_\phi = k\} \cap \bigcap_{j=1}^k T_j^{-1}(B_j)\right) = Q(v_\phi = k) \prod_{j=1}^k Q(B_j).$$

This completes the proof that our $G_0(\omega)$ is an abstract Galton–Watson tree defined on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ and having $\Pi^{(0)}$ as its offspring distribution.

Having given a full account of the hidden Galton–Watson tree corresponding to \mathcal{Y}_0 , let us now give a brief sketch of the construction of the abstract Galton–Watson tree corresponding to \mathcal{Y}_k for $k \geq 1$. Again by the definition of $\tilde{\Omega}$, for each $\omega \in \tilde{\Omega}$ there is a finite $n_0(\omega)$ such that $v_\omega(1_i) > k$ for $i < n_0(\omega)$ and $v_\omega(1_{n_0(\omega)}) \leq k$. Then $u_0(\omega) := 1_{n_0(\omega)} \in V_k(\omega)$, and this vertex will play the role of the root of the tree $G_k(\omega)$ in which we are interested. In contrast to the case of $k = 0$, $u_0(\omega)$ can bear two types of children, described as follows.

Type I. When $n_0(\omega) \geq 1$, vertices of the form $v = 1_\ell b 1_{\ell'}$ are children of $u_0(\omega)$, where $0 \leq \ell < n_0(\omega)$, $1 < b \leq v_\omega(1_\ell)$, and $\ell' = 0$ if $v_\omega(1_\ell b) \leq k$, while

$$\ell' = 1 + \max\{i \geq 0: v_\omega(1_\ell b 1_i) > k\}$$

if $v_\omega(1_\ell b) > k$.

Type II. When $v_\omega(1_{n_0(\omega)}) > 0$, vertices of the form $v = 1_{n_0(\omega)} a 1_m$ are children of $u_0(\omega)$, where $1 \leq a \leq v_\omega(1_{n_0(\omega)})$ and $m = 0$ if $v_\omega(1_{n_0(\omega)} a) \leq k$, while

$$m = 1 + \max\{i \geq 0: v_\omega(1_{n_0(\omega)} a 1_i) > k\}$$

if $v_\omega(1_{n_0(\omega)} a) > k$.

The parent–child relationship between general vertices in $V_k(\omega)$ is defined similarly, giving $V_k(\omega)$ the structure of an abstract tree $G_k(\omega)$. It is actually an abstract Galton–Watson tree whose offspring distribution is the distribution of the independent sum of the random numbers of the two types of children, namely $\Pi^{(k)}$.

For illustration, consider again the tree ω drawn in Figure 1, for which $\mathcal{Y}_1(\omega) = 5$. $G_1(\omega)$ is seen to be isomorphic to the tree depicted in Figure 4. In Figure 5, it is shown how the tree $G_1(\omega)$ grows in ω from the zeroth generation (the root) to the third generation.

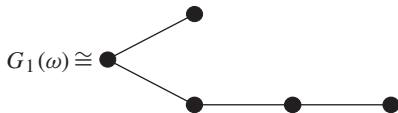


FIGURE 4: The tree $G_1(\omega)$.

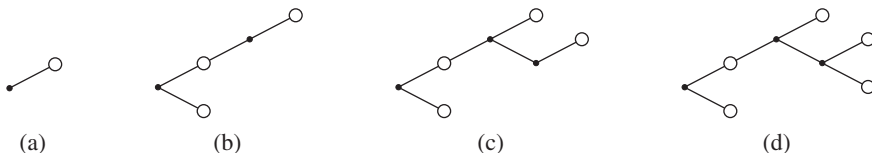


FIGURE 5: How $G_1(\omega)$ grows in ω . Diagram (a) shows the zeroth generation; diagram (b) the zeroth and first generations; diagram (c) the zeroth, first, and second generations; and diagram (d) the zeroth, first, second, and third generations.

3.2. Asymptotic behavior of the distribution of \mathcal{Y}_k

By Theorem 2, the analysis of the asymptotic behavior of $P(\mathcal{Y}_k = n)$ as $n \rightarrow \infty$ is essentially reduced to that of $P(Z = n)$ with a proper choice of offspring distribution. In fact, if we let σ_k be the radius of convergence of the power series $g_k(w)$ and if we define $b_k \in (0, \sigma_k]$ by (5), that is, by the relation

$$\inf_{0 < w < \sigma_k} \frac{g_k(w)}{w} = \frac{g_k(b_k)}{b_k},$$

then we immediately obtain the following result.

Proposition 4. (i) Let $d(\Pi^{(k)})$ be the greatest common divisor of all those $n \geq 0$ such that $p_n^{(k)} > 0$. Then $P(\mathcal{Y}_k = n) > 0$ only for all those n such that $n = 1 \pmod{d(\Pi^{(k)})}$.

(ii) $\limsup_{n \rightarrow \infty} P(\mathcal{Y}_k = n)^{1/n} = g_k(b_k)/b_k$.

(iii) If $0 < b_k < \sigma_k$ then the following asymptotic expansion holds as we let $n \rightarrow \infty$, keeping $n = 1 \pmod{d(\Pi^{(k)})}$:

$$P(\mathcal{Y}_k = n) \sim \sum_{j=1}^{\infty} c_j^{(k)} \left(\frac{g_k(b_k)}{b_k} \right)^n n^{-j-1/2},$$

where

$$c_1^{(k)} = d(\Pi^{(k)}) \sqrt{\frac{g_k(b_k)}{2\pi g_k''(b_k)}}.$$

There is no simple relation between $d(\Pi)$ and $d(\Pi^{(k)})$. In fact, $d(\Pi^{(k)})$ is not determined solely by $d(\Pi)$, but we can prove the following proposition.

Proposition 5. *The greatest common divisor $d(\Pi^{(k)})$ is given by*

$$d(\Pi^{(k)}) = \gcd[\{0 \leq j \leq k : p_j > 0\} \cup \{j \geq k : p_{j+1} > 0\}].$$

Proof. Note that

$$d(\Pi^{(k)}) = \max\{\ell \geq 1 : g_k(e^{2\pi i/\ell}) = 1\},$$

while $g_k(w) = 1$ is equivalent to $h_k(w) = 1 - \varphi_k(w)$, i.e. to the condition

$$\Gamma_k(w) := \sum_{j=0}^{k-1} p_j w^j + (p_k + p_{k+1})w^k + \sum_{j=k+1}^{\infty} p_{j+1} w^{j+1} = 1.$$

Therefore,

$$d(\Pi^{(k)}) = \max\{\ell \geq 1 : \Gamma_k(e^{2\pi i/\ell}) = 1\} \\ = \begin{cases} \gcd[\{0 \leq j \leq k : p_j > 0\} \cup \{k\} \cup \{j \geq k : p_{j+1} > 0\}] & \text{if } p_k + p_{k+1} > 0, \\ \gcd[\{0 \leq j \leq k : p_j > 0\} \cup \{j \geq k : p_{j+1} > 0\}] & \text{if } p_k + p_{k+1} = 0, \end{cases}$$

which gives the desired result.

Now we will give a condition, in terms of $\Pi = \{p_n\}_{n=0}^{\infty}$ and ρ , to ensure that $0 < b_k < \sigma_k$. Due to (11), our consideration is divided into four cases.

Case 1. $\varphi_k(\rho-) > 1$. In this case, there is a $\sigma_k \in (0, \rho)$ such that $\varphi_k(\sigma_k) = 1$. This σ_k is the radius of convergence of the power series $g_k(w)$, and we have $g_k(\sigma_k-) = \infty$. Consequently, $b_k \in (0, \sigma_k)$.

Case 2. If $\varphi_k(\rho-) = 1$ then $\sigma_k = \rho < \infty$ and $g_k(\sigma_k) = \infty$. In this case, also, we have $b_k \in (0, \sigma_k)$.

Case 3. If $0 < \varphi_k(\rho-) < 1$ then we have $\sigma_k = \rho < \infty$ and $g_k(\sigma_k-) < \infty$.

Case 4. Finally, if $\varphi_k(\rho-) = 0$ then $g_k(w) = f(w)$ and we trivially have $b_k = a$.

In the last two cases, both $b_k < \sigma_k$ and $b_k = \sigma_k$ are possible.

Proposition 6. *The inequalities $0 < b_k < \sigma_k$ hold if and only if*

$$\lim_{z \uparrow \rho} \{(z - f(z))h'_k(z) - (1 - f'(z))h_k(z)\} > 0.$$

Proof. Define

$$\Delta_k(z) := (z - f(z))h'_k(z) - (1 - f'(z))h_k(z).$$

Then

$$zg'_k(z) - g_k(z) = \frac{\Delta_k(z)}{(1 - \varphi_k(z))^2}$$

and, after some computation, we get

$$\begin{aligned} \Delta_k(z) &= \sum_{\ell=0}^k (\ell - 1) p_\ell z^\ell + \sum_{j=k+1}^\infty \sum_{\ell=0}^k (j - \ell) p_j p_\ell z^{j+\ell-1} \\ &= \sum_{\ell=1}^k (\ell - 1) p_\ell z^\ell + \sum_{\ell=1}^k \sum_{j=k+1}^\infty (j - \ell) p_j p_\ell z^{j+\ell-1} + p_0 \left(\sum_{j=k+1}^\infty j p_j z^{j-1} - 1 \right), \end{aligned} \tag{18}$$

which shows, in particular, that $\Delta'_k(z) > 0$ for $z > 0$.

From the preliminary consideration preceding the proposition and the condition (4) applied to $g_k(z)$, we first see that $0 < b_k < \sigma_k$ if and only if $\varphi_k(\rho-) \geq 1$, or $\varphi_k(\rho-) < 1$ and $\lim_{z \uparrow \sigma_k} \Delta_k(z) > 0$.

Hence, if $0 < b_k < \sigma_k$ and $\varphi_k(\rho-) < 1$, we have that $\sigma_k = \rho$ and

$$\lim_{z \uparrow \rho} \Delta_k(z) = \lim_{z \uparrow \sigma_k} \Delta_k(z) > 0.$$

On the other hand, in the case $\varphi_k(\rho-) \geq 1$, we have that $\varphi_k(\sigma_k) = 1$, i.e.

$$\sum_{j=k+1}^\infty p_j \sigma_k^{j-1} = 1.$$

Since $p_0 + p_1 < 1$, this implies that

$$\sum_{j=k+1}^\infty j p_j \sigma_k^{j-1} > 1,$$

so that $\Delta_k(\sigma_k) > 0$, by (18). Since $\rho \geq \sigma_k$, this implies that $\Delta_k(\rho-) \geq \Delta_k(\sigma_k) > 0$.

Conversely, suppose that $\Delta_k(\rho-) > 0$. If, in addition, $\varphi_k(\rho-) \geq 1$ then we have $0 < b_k < \sigma_k$ as we already saw, and if $\varphi_k(\rho-) < 1$ then, since $\sigma_k = \rho$,

$$\lim_{z \uparrow \sigma_k} \{z g'_k(z) - g_k(z)\} = \frac{1}{(1 - \varphi_k(\rho-))^2} \Delta_k(\rho-) > 0,$$

that is, $0 < b_k < \sigma_k$, by (4).

Since $\mathcal{Y}_k \uparrow Z$ as $k \rightarrow \infty$, the number $g_k(b_k)/b_k$ governing the exponential behavior of $P(\mathcal{Y}_k = n)$ is expected to increase with k . This is actually the case, as the following proposition shows.

Proposition 7. *For each $k \geq 0$, we have*

$$\frac{g_k(b_k)}{b_k} \leq \frac{g_{k+1}(b_{k+1})}{b_{k+1}} \leq \frac{f(a)}{a}. \tag{19}$$

If $f(z)$ is critical (i.e. $f'(1) = 1$) or subcritical ($f'(1) < 1$) with $\rho = 1$, then each member is equal to 1. If $p_{k+1} = 0$ then $g_k(z) = g_{k+1}(z)$ and the first inequality is actually an equality. If, further, $p_\ell = 0$ for all $\ell > k$, then $g_k(z) = f(z)$ and we have equality also in the second inequality. In all other cases, the inequalities are strict.

Finally, we have

$$\lim_{k \rightarrow \infty} \frac{g_k(b_k)}{b_k} = \frac{f(a)}{a}. \tag{20}$$

Proof. First note that

$$g_{k+1}(z) - g_k(z) = \frac{z(z - f(z))(h_{k+1}(z) - h_k(z))}{(z - f(z) + h_k(z))(z - f(z) + h_k(z))}. \tag{21}$$

For the moment, we suppose that there exist q_1 and q_2 such that $f(q_i) = q_i, i = 1, 2$, and that $0 < q_1 < q_2 \leq \rho$. Since

$$\varphi_k(q_2) = 1 - \frac{h_k(q_2)}{q_2} < 1,$$

we have $q_2 \leq \sigma_k$ and, thus, $g_k(z)$ is well defined on $[q_1, q_2]$. However, for $z \in [q_1, q_2]$ we have $f(z) \leq z$ and, hence, $g_{k+1}(z) \geq g_k(z)$ holds on $[q_1, q_2]$, by (21). By the convexity of $f(z)/z$ and $g_k(z)/z$, and by $f(q_i)/q_i = g_k(q_i)/q_i = 1, i = 1, 2$, we see that a and b_k belong to $[q_1, q_2]$.

Now, since we have $h_k(z) \uparrow f(z)$, we also have that $g_k(z) \uparrow f(z)$ on $[q_1, q_2]$ as $k \rightarrow \infty$. By Dini’s lemma (see, e.g. [15]), this convergence is uniform on the compact interval $[q_1, q_2]$. That is, for any $\varepsilon > 0, g_k(z)/z > (f(z)/z) - \varepsilon$ holds for all $z \in [q_1, q_2]$ and for all sufficiently large k . Taking $z = b_k$ and $z = a$, we obtain

$$\frac{f(a)}{a} \geq \frac{g_k(a)}{a} \geq \frac{g_k(b_k)}{b_k} > \frac{f(b_k)}{b_k} - \varepsilon \geq \frac{f(a)}{a} - \varepsilon$$

for large k . This shows that $\lim_{k \rightarrow \infty} g_k(b_k)/b_k = f(a)/a$.

Now suppose that $f(z)$ is supercritical and let q be the extinction probability: $f(q) = q$. Then we can take $q_1 = q, q_2 = 1$ and (19) and (20) are proved in this case. When $f(z)$ is critical, or subcritical with $\rho = 1$, we have

$$a = f(a) = b_k = g_k(b_k) = 1$$

and there is nothing to be proved. Next, consider the case where $f(z)$ is subcritical (i.e. $f'(1) < 1$), $\rho > 1$, but $f(z) < z$ for $1 < z < \rho$. In this case we have, in particular, $\rho < \infty$ and $f(\rho) < \infty$, and $g_k(z) \uparrow f(z)$ holds uniformly on $[1, \rho]$. Thus, the above argument can be applied to prove that $g_k(b_k)/b_k \uparrow f(a)/a$.

Finally, suppose that $f(z)$ is supercritical, or subcritical with $\rho > 1$, and that $p_{k+1} > 0$. Then $h_{k+1}(z) > h_k(z)$ for $z > 0$ and, hence, we have $g_{k+1}(z) > g_k(z)$ on (q_1, q_2) or on $(1, \rho)$. Thus,

$$\frac{g_{k+1}(b_{k+1})}{b_{k+1}} > \frac{g_k(b_{k+1})}{b_{k+1}} \geq \frac{g_k(b_k)}{b_k}$$

and the (first) inequality is strict.

4. The joint distribution of Z and \mathcal{Y}_k

4.1. Generality

For complex numbers $u, v_j (j = 0, 1, \dots)$ satisfying $|u|, |v_j| \leq 1$, let

$$w := G(u; v_0, v_1, \dots) := E \left[Z < \infty; u^Z \prod_{j=0}^{\infty} v_j^{Y_j} \right]$$

be the joint probability generating function of Z, Y_0, Y_1, \dots . Here, $\prod_{j=0}^\infty v_j^{Y_j}$ is actually a finite product since $\sum_{j=0}^\infty Y_j = Z < \infty$. Due to the relations (1) and

$$Y_k(\omega) = \mathbf{1}_{\{v_\phi(\omega)=k\}} + \sum_{j=1}^{v_\phi(\omega)} Y_k(T_j(\omega)),$$

we can proceed, using the formula of Neveu (page 202 of [23]), to write

$$\begin{aligned} w &= \mathbb{E} \left[Z < \infty; u^{1+\sum_{j=1}^{v_\phi} Z \circ T_j} \prod_{k=0}^\infty v_k^{\mathbf{1}_{\{v_\phi=k\}} + \sum_{j=1}^{v_\phi} Y_k \circ T_j} \right] \\ &= u \mathbb{E} \left[\left(\prod_{j=1}^{v_\phi} \mathbf{1}_{\{Z < \infty\}} \circ T_j \right) \left(\prod_{k=0}^\infty v_k^{\mathbf{1}_{\{v_\phi=k\}}} \right) \left\{ \prod_{j=1}^{v_\phi} \left(u^Z \prod_{k=0}^\infty v_k^{Y_k} \right) \circ T_j \right\} \right] \\ &= u \mathbb{E} \left[\prod_{k=0}^\infty v_k^{\mathbf{1}_{\{v_\phi=k\}}} \mathbb{E} \left[\prod_{j=1}^{v_\phi} \left(\mathbf{1}_{\{Z < \infty\}} u^Z \prod_{k=0}^\infty v_k^{Y_k} \right) \circ T_j \mid \mathcal{F}_1 \right] \right] \\ &= u \mathbb{E} \left[\prod_{k=0}^\infty v_k^{\mathbf{1}_{\{v_\phi=k\}}} \prod_{j=1}^{v_\phi} \mathbb{E} \left[Z < \infty; u^Z \prod_{k=0}^\infty v_k^{Y_k} \right] \right] \\ &= u \mathbb{E} \left[\prod_{k=0}^\infty v_k^{\mathbf{1}_{\{v_\phi=k\}}} w^{v_\phi} \right] \\ &= u \sum_{k=0}^\infty p_k v_k w^k. \end{aligned}$$

Thus, if we define

$$\Phi(w) = \Phi(w; v_0, v_1, \dots) := \sum_{k=0}^\infty p_k v_k w^k,$$

then w is a solution of the equation $w = u\Phi(w)$. If we assume that $v_0 \neq 0$ then $\Phi(0) = p_0 v_0 \neq 0$, and we can again apply Lagrange inversion to obtain the formula

$$\mathbb{E} \left[Z = n; \prod_{k=0}^\infty v_k^{Y_k} \right] = \frac{1}{2\pi i n} \oint \left\{ \frac{\Phi(w)}{w} \right\}^n dw \tag{22}$$

for the coefficients of the power series

$$w = \sum_{n=1}^\infty \mathbb{E} \left[Z = n; \prod_{k=0}^\infty v_k^{Y_k} \right] u^n.$$

Now, if $Z(\omega) = n$ we then have $Y_k(\omega) = 0$ for $k \geq n$. Moreover, $\sum_{k=0}^{n-1} Y_k(\omega) = 1 + \sum_{k=1}^{n-1} k Y_k(\omega) = n$. Hence, for $n_{k_j} \geq 0, j = 1, \dots, r$, with $r \geq 1$ and $0 \leq k_1 < \dots < k_r$, the probability $\mathbb{P}(Z = n, Y_{k_j} = n_{k_j}, j = 1, \dots, r)$ is nonzero only if

$$n_{k_j} = 0 \quad \text{for } k_j \geq n, \quad \mu := \sum_{j=1}^r n_{k_j} \leq n, \quad \text{and} \quad \nu := \sum_{j=1}^r k_j n_{k_j} < n \tag{23}$$

and, in such a case, we obtain from (22) that

$$\begin{aligned}
 &P(Z = n, Y_{k_j} = n_{k_j}, j = 1, \dots, r) \\
 &= \left(\prod_{k \geq 0} \frac{1}{n_k!} \right) \frac{\partial^\mu}{\partial v_{k_1}^{n_{k_1}} \partial v_{k_2}^{n_{k_2}} \dots \partial v_{k_r}^{n_{k_r}}} E \left[Z = n; \prod_{k \geq 0} v_k^{Y_k} \right] \\
 &= \frac{1}{2\pi i n} \left(\prod_{k \geq 0} \frac{1}{n_k!} \right) \frac{\partial^\mu}{\partial v_{k_1}^{n_{k_1}} \partial v_{k_2}^{n_{k_2}} \dots \partial v_{k_r}^{n_{k_r}}} \oint \left\{ \frac{\Phi(z; v_0, v_1, \dots)}{z} \right\}^n dz \\
 &= \frac{1}{2\pi i n} \left(\prod_{k \geq 0} \frac{1}{n_k!} \right) \oint \frac{1}{z^n} \frac{n!}{(n - \mu)!} \left(\prod_{j=1}^r p_{k_j}^{n_{k_j}} \right) z^\nu \tilde{f}(z)^{n-\mu} dz \\
 &= \frac{(n - 1)!}{(n - \mu)!} \left(\prod_{k \geq 0} \frac{1}{n_k!} \right) \left(\prod_{j=1}^r p_{k_j}^{n_{k_j}} \right) \frac{1}{2\pi i} \oint \frac{\tilde{f}(z)^{n-\mu}}{z^{n-\nu}} dz,
 \end{aligned}$$

where we have set

$$\tilde{f}(z) = \sum_{\{k: k \neq k_j, j=1, \dots, r\}} p_k z^k,$$

and the partial derivatives are evaluated as $v_{k_1}, \dots, v_{k_r} \rightarrow 0$ and with $v_k = 1$ for $k \neq k_j$, $j = 1, \dots, r$.

Thus, we have obtained the following result.

Proposition 8. For (n_{k_j}) satisfying (23), we have

$$\begin{aligned}
 &P(Z = n, Y_{k_j} = n_{k_j}, j = 1, \dots, r) \\
 &= \frac{(n - 1)!}{(n - \mu)!} \left(\prod_{k \geq 0} \frac{1}{n_k!} \right) \left(\prod_{j=1}^r p_{k_j}^{n_{k_j}} \right) \frac{1}{2\pi i} \oint \frac{\tilde{f}(z)^{n-\mu}}{z^{n-\nu}} dz \\
 &= \frac{(n - 1)!}{(n - \mu)! (n - \nu - 1)!} \left(\prod_{k \geq 0} \frac{1}{n_k!} \right) \left(\prod_{j=1}^r p_{k_j}^{n_{k_j}} \right) \frac{1}{2\pi i} \frac{d^{n-\nu-1}}{dz^{n-\nu-1}} \tilde{f}(z)^{n-\nu} \Big|_{z=0}. \tag{24}
 \end{aligned}$$

In particular, if $(n_k)_{k \geq 0}$ are such that

$$n_k = 0 \text{ for } k \geq n, \quad \sum_{k=0}^{n-1} n_k = 1 + \sum_{k=1}^{n-1} k n_k = n, \tag{25}$$

then

$$P(Z = n, Y_k = n_k, k = 0, 1, \dots) = (n - 1)! \prod_{k=0}^{n-1} \frac{p_k^{n_k}}{n_k!}. \tag{26}$$

Next, let us consider the joint factorial moments of the Y_k , restricted on the event $\{Z = n\}$, which we will define as

$$M(n; r_0, r_1, \dots) := E \left[Z = n; \prod_{k \geq 0} Y_k^{[r_k]} \right],$$

where, as usual, we define

$$n^{[r]} = \begin{cases} n(n-1)\cdots(n-r+1) & \text{for } r \leq n, \\ 0 & \text{for } r > n, \end{cases}$$

for nonnegative integers n and r , so that, in particular, $n^{[0]} = 1$ and $n^{[1]} = n$ for any $n \geq 0$, and $r_k > 0$ only for a finite number of k s. More specifically, we will assume that $r_k = 0$ for $k \geq n$, in view of (23), and set $r = \sum_{k=0}^{n-1} r_k$. Now it is easy to compute, using (22), that

$$\begin{aligned} M(n; r_0, r_1, \dots) &= \frac{\partial^r}{\partial v_0^{r_0} \partial v_1^{r_1} \cdots \partial v_{n-1}^{r_{n-1}}} \mathbb{E} \left[Z = n; \prod_{k \geq 0} v_k^{Y_k} \right] \Big|_{v_0=v_1=\dots=1} \\ &= \frac{1}{2\pi i n} \oint \frac{1}{z^n} \frac{\partial^r}{\partial v_0^{r_0} \partial v_1^{r_1} \cdots \partial v_{n-1}^{r_{n-1}}} \Phi(z; v_0, v_1, \dots)^n \Big|_{v_0=v_1=\dots=1} dz \\ &= \frac{1}{2\pi i n} \oint \frac{1}{z^n} n^{[r]} f(z)^{n-r} \prod_{k=0}^{n-1} p_k^{r_k} z^{kr_k} dz \\ &= \frac{n^{[r]}}{2\pi i n} \left(\prod_{k=0}^{n-1} p_k^{r_k} \right) \oint \frac{f(z)^{n-r}}{z^{n-s}} dz, \end{aligned} \tag{27}$$

where we have set $s = \sum_{k=0}^{n-1} kr_k$. This formula may be used in obtaining means and covariances of the Y_k under the conditional distribution $P(\cdot \mid Z = n)$. Moreover, when Condition A holds, we can apply the analysis in Section 2.2 to the above formula, to obtain an asymptotic expansion of $M(n; r_0, r_1, \dots)$ as $n \rightarrow \infty$. For example, we obtain an asymptotic expression as $n \rightarrow \infty$ for the expectation $E_n(Y_k)$, variance $\text{var}_n(Y_k)$, and covariance $\text{cov}_n(Y_k, Y_{k'})$ ($k \neq k'$) of the Y_k under the conditional distribution $P(\cdot \mid Z = n)$, as follows:

$$E_n(Y_k) = \rho_k n + \mathcal{O}(1), \tag{28}$$

$$\text{var}_n(Y_k) = \frac{\rho_k}{\kappa_2} [(1 - \rho_k)\kappa_2 - (k - 1)^2 \rho_k] n + \mathcal{O}(1), \tag{29}$$

$$\text{cov}_n(Y_k, Y_{k'}) = -\frac{\rho_k \rho_{k'}}{\kappa_2} [\kappa_2 + (k - 1)(k' - 1)] n + \mathcal{O}(1), \tag{30}$$

where, as before, $\rho_k := p_k a^k / f(a)$ and κ_2 is the variance of the probability distribution $\Pi_a := \{\rho_k\}_{k \geq 0}$.

Let us now discuss several examples. See [1] for related topics.

4.2. Example 1: geometric offspring distribution

Suppose that $\Pi = \{p_k\}_{k=0}^\infty$ is the geometric distribution with parameter p , i.e. $p_k = (1 - p)p^k$, $k = 0, 1, 2, \dots$, for some $0 < p < 1$. In this case, $f(z) = (1 - p)/(1 - pz)$ and we find that

$$\begin{aligned} P(Z = n) &= \frac{1}{2\pi i n} \oint \frac{1}{z^n} \left(\frac{1 - p}{1 - pz} \right)^n dz \\ &= \frac{1}{n!} \frac{d^{n-1}}{dz^{n-1}} \left(\frac{1 - p}{1 - pz} \right)^n \Big|_{z=0} \\ &= (1 - p)^n p^{n-1} \frac{1}{n} \binom{2n - 2}{n - 1}. \end{aligned} \tag{31}$$

For a finite tree $\omega \in \Omega$ with $Z(\omega) = n$, we thus obtain

$$P(\{\omega\} \mid Z = n) = \frac{1}{P(Z = n)} \prod_{u \in \omega} p_{v_u(\omega)} = \left[\frac{1}{n} \binom{2n-2}{n-1} \right]^{-1}.$$

The right-hand side being independent of the detailed data $\{v_u(\omega)\}_{u \in \omega}$ on ω , this shows that, conditioned on the event $\{Z = n\}$, all trees of size n are equiprobable, which in turn implies that the total number of trees of size n is $(1/n) \binom{2n-2}{n-1}$. This is well known in combinatorics.

For each integer $n \geq 1$ and for each sequence (n_k) of nonnegative integers satisfying (25), we have, by (26) and (31),

$$\begin{aligned} P(Y_k = n_k, k \geq 0 \mid Z = n) &= \frac{(n-1)! (\prod_{k \geq 0} (1-p)^{kn_k} / n_k!)}{(1-p)^n p^{n-1}} \left[\frac{1}{n} \binom{2n-2}{n-1} \right]^{-1} \\ &= \frac{1}{n} \binom{n}{n_0 n_1 \dots} \left[\frac{1}{n} \binom{2n-2}{n-1} \right]^{-1}, \end{aligned}$$

where

$$\binom{n}{n_0 n_1 \dots} := \frac{n!}{n_0! n_1! \dots}, \quad \text{with } \sum_{j \geq 0} n_j = n,$$

is the multinomial coefficient. By the equiprobability of the trees under the condition $\{Z = n\}$, this shows that the total number of trees ω such that $Z(\omega) = n$ and $Y_k(\omega) = n_k, k \geq 0$, is given by

$$\frac{1}{n} \binom{n}{n_0 n_1 \dots}, \tag{32}$$

where (n_k) satisfies (25). This result is also well known and is attributed to Tutte [28], but it had already been stated by Otter (page 213 of [24]), although with a somewhat incomplete proof.

Let us now take $r = 1, k_1 = 0$, and $j := n_{k_1} = \mu < n$ in (24). Then we have $\nu = 0$ and

$$\tilde{f}(z) = (1-p) \sum_{k \geq 1} (pz)^k = pz f(z) = (1-p) \frac{pz}{1-pz},$$

so that

$$\begin{aligned} P(Z = n, Y_0 = j) &= \frac{(n-1)!}{(n-j)! j!} (1-p)^j \frac{1}{2\pi i} \oint \frac{(pz)^{n-j}}{z^n} f(z)^{n-j} dz \\ &= \frac{(n-1)!}{(n-j)! j!} p^{n-j} (1-p)^n \frac{1}{(j-1)!} \frac{d^{j-1}}{dz^{j-1}} (1-pz)^{-(n-j)} \Big|_{z=0} \\ &= p^{n-1} (1-p)^n \frac{1}{n} \binom{n}{j} \binom{n-2}{j-1}. \end{aligned}$$

Again by (31) and the equiprobability of trees, we see that the number $p_{n,j}$ of trees of size n with exactly j vertices having no children (namely ‘leaves’), is given by

$$p_{n,j} = \frac{1}{n} \binom{n}{j} \binom{n-2}{j-1}.$$

This formula is also known in combinatorics (see, e.g. [8]). Note that ‘trees’ in the sense we mean are often called ‘planted plane trees’ in combinatorics. Thus, $p_{n,j}$ is the number of planted plane trees with j leaves.

4.3. Example 2: binomial offspring distribution

Next, consider the case where Π is the binomial distribution $B(d, p)$, so that $f(z) = (1 - p + pz)^d$. Applying (2), we can easily compute that

$$P(Z = n) = \frac{1}{n} \binom{dn}{n-1} (1-p)^{dn-n+1} p^{n-1}.$$

On the other hand, for a tree ω of size n ,

$$\begin{aligned} P(\{\omega\}) &= \prod_{u \in \omega} p_{v_u(\omega)} = \prod_{u \in \omega} \binom{d}{v_u(\omega)} p^{v_u(\omega)} (1-p)^{d-v_u(\omega)} \\ &= p^{n-1} (1-p)^{dn-n+1} \prod_{u \in \omega} \binom{d}{v_u(\omega)} \end{aligned}$$

and, hence,

$$P(\{\omega\} \mid Z = n) = \left[\frac{1}{n} \binom{dn}{n-1} \right]^{-1} \prod_{u \in \omega} \binom{d}{v_u(\omega)}. \tag{33}$$

These results have the following interpretation. Consider the bond percolation on the rooted d -ary tree. Here we mean by the ‘rooted d -ary tree’ an element $T \in \Omega$ such that $v_u(T) = d$ for all $u \in T$, and we suppose that each bond, or edge, of the graph T is ‘open’ with a certain probability $0 < p < 1$ and ‘closed’ with probability $1 - p$, independently of the other edges. We then call a percolation cluster on T a subset γ of T consisting of those vertices $u \in T$ that are connected to the root ϕ through open bonds. A percolation cluster γ thus obtained is a ‘tree’ distinct from our Galton–Watson tree defined in Section 1, in the sense that each vertex $u \in \gamma$ can have k children, among d possible ones, in $\binom{d}{k}$ different ways. (When $d = 2$, the percolation clusters thus obtained are often called ‘uniform binary trees’ – see, e.g. [22].) Each realization γ of the percolation cluster is an abstract tree in the sense of Definition 1 and, hence, is isomorphic to a tree $\omega \in \Omega$. We denote this correspondence by $\omega = \Psi(\gamma)$. When ω is finite, there are exactly $\prod_{u \in \omega} \binom{d}{v_u(\omega)}$ percolation clusters corresponding to the same ω .

Now let \mathcal{C}_n be the totality of percolation clusters of size n . From what we have seen above,

$$\#\mathcal{C}_n = \sum_{\{\omega \in \Omega : Z(\omega) = n\}} \#\Psi^{-1}(\{\omega\}) = \sum_{\{\omega \in \Omega : Z(\omega) = n\}} \prod_{u \in \omega} \binom{d}{v_u(\omega)}.$$

On the other hand, we have

$$1 = \sum_{\{\omega \in \Omega : Z(\omega) = n\}} P(\{\omega\} \mid Z = n) = \left[\frac{1}{n} \binom{dn}{n-1} \right]^{-1} \sum_{\{\omega \in \Omega : Z(\omega) = n\}} \prod_{u \in \omega} \binom{d}{v_u(\omega)}$$

and, thus,

$$\#\mathcal{C}_n = \frac{1}{n} \binom{dn}{n-1}.$$

If we let $E_u(\gamma)$ be the number of outward edges issued from the vertex u of a percolation cluster $\gamma \in \mathcal{C}_n$, then it is clear that

$$X_k(\gamma) := \#\{u \in \gamma : E_u(\gamma) = k\} = Y_k(\Psi(\gamma)).$$

Thus, if we are given $(n_k)_{k \geq 0}$, which is subject to the condition (25), then we have

$$\begin{aligned} \#\{\gamma \in \mathcal{C}_n : X_k(\gamma) = n_k, k \geq 0\} &= \sum_{\{\omega \in \Omega : Z(\omega) = n, Y_k(\omega) = n_k, k \geq 0\}} \#\Psi^{-1}(\{\omega\}) \\ &= \#\{\omega \in \Omega : Z(\omega) = n, Y_k(\omega) = n_k, k \geq 0\} \times \prod_{v=0}^d \binom{d}{v}^{n_v} \\ &= \frac{1}{n} \binom{n}{n_0 n_1 \dots} \prod_{v=0}^d \binom{d}{v}^{n_v}. \end{aligned}$$

This shows that if Q_n is the uniform distribution on \mathcal{C}_n , then, from (32) and (33),

$$\begin{aligned} Q_n(X_k = n_k, k \geq 0) &= \frac{1}{n} \binom{n}{n_0 n_1 \dots} \prod_{v=0}^d \binom{d}{v}^{n_v} \left[\frac{1}{n} \binom{nd}{n-1} \right]^{-1} \\ &= P(Y_k = n_k, k \geq 0 \mid Z = n). \end{aligned}$$

Next, consider the special case $d = 2$ and let $r = 1, k_1 = 0$, and $j := n_{k_1} = \mu < n$ in (24). Then, $v = 0$ and $\tilde{f}(z) = pz(pz + 2q)$ with $q = 1 - p$, and we have

$$\begin{aligned} P(Z = n, Y_0 = j) &= \frac{(n-1)!}{(n-j)!} \frac{1}{j!} q^{2j} \frac{1}{2\pi i} \oint \frac{1}{z^n} (pz)^{n-j} (pz + 2q)^{n-j} dz \\ &= \frac{1}{n} \binom{n}{j} 2^{n+1-2j} q^{n+1} p^{n-1} \binom{n-j}{j-1} \end{aligned}$$

and, hence,

$$Q_n(X_0 = j) = P(Y_0 = j \mid Z = n) = \frac{1}{b_n} 2^{n+1-2j} \frac{1}{n} \binom{n}{j} \binom{n-j}{j-1}, \tag{34}$$

where $b_n = \binom{2n}{n-1}/n$ is the n th Catalan number. Note that, from (25), we have

$$Y_0 + Y_1 + Y_2 = n, \quad Y_1 + 2Y_2 = n - 1,$$

so that, in particular, $n - j \geq j - 1$ when $Y_0 = j$. Moreover, these relations imply that

$$Q_n(X_2 = j) = Q_n(X_0 = j + 1) \tag{35}$$

and

$$Q_n(X_1 = j) = \begin{cases} Q_n(X_0 = (n + 1 - j)/2) & \text{if } n - j \text{ is odd,} \\ 0 & \text{if } n - j \text{ is even.} \end{cases} \tag{36}$$

Explicit formulae for $Q_n(X_i = j), i = 0, 1, 2$, were first given by Mahmoud in [22], but his expression consisted of very complicated alternating sums. Later, Prodinger [25] computed these sums, to obtain (34), (35), and (36). (See also [13], [20], and [26].)

We can also apply (27) to compute the factorial moments $E_{Q_n}[X_i^{[r]}], i = 0, 1, 2$, reproducing Theorem 3 of [25]. Thus, for example,

$$\begin{aligned} E_{Q_n}[X_0^{[r]}] &= \frac{E[Z = n; Y_0^{[r]}]}{P(Z = n)} \\ &= \frac{1}{P(Z = n)} \frac{n^{[r]}}{2\pi i n} q^{2r} \oint \frac{f(z)^{n-r}}{z^n} dz \\ &= \frac{1}{P(Z = n)} n^{[r]} q^{n+1} p^{n-1} \binom{2n - 2r}{n - 1} \\ &= \frac{1}{b_n} n^{[r]} \binom{2n - 2r}{n - 1}. \end{aligned}$$

4.4. Example 3: power series offspring distribution

As is seen in the above examples, the conditional distribution $P(\{\omega\} | Z = n)$ does not depend on the parameter p . This is generally true when the offspring distribution is a power series distribution. More precisely, let $B(\theta) = \sum_{k \geq 0} b_k \theta^k$ be a power series with nonnegative coefficients and with positive radius of convergence R . Assuming that $b_0 > 0$ and $b_k > 0$ for some $k \geq 2$, we let

$$p_k(\theta) = \frac{b_k \theta^k}{B(\theta)}, \quad k = 0, 1, 2, \dots,$$

for $0 < \theta < R$. Then the probability distribution $\Pi(\theta) = \{p_k(\theta)\}_{k=0}^\infty$ satisfies the basic assumptions $p_0(\theta) > 0$ and $p_0(\theta) + p_1(\theta) < 1$, and we have the following proposition.

Proposition 9. *Let P_θ be the probability measure for the Galton–Watson tree with offspring distribution $\Pi(\theta)$. Then the total progeny Z provides sufficient statistics for the statistical model $(\Omega, \mathcal{F}, \{P_\theta(\cdot | Z < \infty)\}_{\theta \in (0, R)})$ in the sense that, for any $n \geq 1$, the conditional distribution $P_\theta(\cdot | Z = n)$ does not depend on θ .*

Proof. The generating function of $\Pi(\theta)$ is given by $B(\theta z)/B(\theta)$. Hence, by (2),

$$P(Z = n) = \frac{1}{2\pi i n} \oint \frac{1}{z^n} \left(\frac{B(\theta z)}{B(\theta)} \right)^n dz = \frac{1}{2\pi i n} \frac{\theta^{n-1}}{B(\theta)^n} \oint \frac{B(w)^n}{w^n} dw.$$

On the other hand, for a tree ω with $Z(\omega) = n$,

$$P(\{\omega\}) = \prod_{u \in \omega} p_{v_u(\omega)}(\theta) = \frac{\theta^{n-1}}{B(\theta)^n} \prod_{u \in \omega} b_{v_u(\omega)},$$

so that by dividing, the factors involving θ cancel out of $P(\{\omega\} | Z = n)$.

5. The weak law of large numbers and the central limit theorem for the distribution of $\{Y_k\}$, conditioned on Z

In this section, we prove two conditional limit theorems for $\{Y_k\}$, namely the assertions (i) and (ii) of Theorem 3 below. Although assertion (i), the weak law of large numbers, is actually due to Otter (Theorem 6 in [24]), we will give an alternative proof, which will be completed almost simultaneously with the proof of the central limit theorem of assertion (ii).

Theorem 3. *Suppose that Condition A holds. Then the following limit theorems hold, where we let $n \rightarrow \infty$ keeping $n = 1 \pmod{d(\Pi)}$.*

(i) *The weak law of large numbers. For any $k \geq 0$,*

$$P(Y_k/n \in \cdot \mid Z = n) \longrightarrow \delta_{\rho_k}(\cdot),$$

where $\delta_b(\cdot)$ is the point mass concentrated on b . Here, $\{\rho_k\}_{k \geq 0}$ is the probability distribution on \mathbb{Z}_+ defined by $\rho_k = p_k a^k / a^k$.

(ii) *The central limit theorem. For any finite subset A of \mathbb{Z}_+ ,*

$$P(\{n^{1/2}(Y_k/n - \rho_k)\}_{k \in A} \in \cdot \mid Z = n) \longrightarrow N_A(0, V_A),$$

where $N_A(0, V_A)$ is the $|A|$ -dimensional Gaussian distribution with mean 0, and with the covariance matrix V_A determined from the quadratic form

$$V(\{t_k\}) = \sum_{k \in A} t_k^2 \rho_k - \left(\sum_{k \in A} t_k \rho_k \right)^2 - \left[\sum_{k \geq 2} k(k-1) \rho_k \right]^{-1} \left(\sum_{k \in A} k t_k \rho_k - \sum_{k \in A} t_k \rho_k \right)^2.$$

Here, V_A is strictly positive definite if and only if $A \subset S$ and $|S \setminus A| \geq 2$, where we let $S := \{k \in \mathbb{Z}_+ : p_k > 0\}$.

Remark 1. The special case $A = \{0\}$ of the central limit theorem is known to combinatorialists – see, e.g. Theorem 12 of [8].

On the other hand, under the condition that $f(z)$ be entire, Dembo *et al.* [7] proved a general large deviation principle that reduces to the large deviation principle corresponding to the law of large numbers in assertion (i). (To see this, set $\mathcal{X} = \text{singleton}$ in Theorem 2.2 of [7].)

For other types of conditional limit theorem considered on the event $\{Z = n\}$, see, e.g. [18] and [19].

Proof. Take a real sequence $\{t_k\}_{k \geq 0}$, where $t_k \neq 0$ for only finitely many k s, and let $v_k = e^{\zeta t_k}$, $\zeta \in \mathbb{C}$, in (22). Then, we obtain

$$E \left[Z = n; \exp \left(\zeta \sum_{k \geq 0} t_k Y_k \right) \right] = \frac{1}{2\pi i n} \oint \exp(nh(z, \zeta)) dz, \tag{37}$$

where we let

$$h(z, \zeta) := \log \left[\frac{\Phi(z, \zeta)}{z} \right] \quad \text{and} \quad \Phi(z, \zeta) := \sum_{k \geq 0} p_k e^{\zeta t_k} z^k.$$

If we further define

$$\begin{aligned} M(z, \zeta) &:= \sum_{k \geq 1} k p_k e^{\zeta t_k} z^k - \sum_{k \geq 0} p_k e^{\zeta t_k} z^k \\ &= z \frac{\partial}{\partial z} \Phi(z, \zeta) - \Phi(z, \zeta), \end{aligned}$$

then we get

$$\frac{\partial}{\partial z} h(z, \zeta) = \frac{M(z, \zeta)}{z \Phi(z, \zeta)}.$$

In order to apply the saddle point approximation to the integral in (37), we need to know the zero of the function $\partial h/\partial z$ or, equivalently, we need to solve the equation $M(z, \zeta) = 0$ with respect to z . Since $\Phi(z, 0) = f(z)$, we have $M(a, 0) = af'(a) - f(a) = 0$. On the other hand, we have $(\partial M/\partial z)(a, 0) = af''(a) > 0$. Hence, by the implicit function theorem, there is a unique analytic function $a(\zeta)$, defined in a neighbourhood of the origin, such that $a(0) = a$ and $M(a(\zeta), \zeta) = 0$.

Now, if we take the contour $|z| = |a(\zeta)|$ in the integral in (22), and if we assume that $n = 1 \pmod{d(\Pi)}$, then, as in (6), we obtain

$$\mathbb{E}\left[Z = n; \prod_{k \geq 0} \exp(\zeta t_k Y_k)\right] = \frac{a(\zeta)d}{2\pi n} \int_{-\pi/d}^{\pi/d} \exp(n\psi(\theta, \zeta)) e^{i\theta} d\theta, \tag{38}$$

where we have set

$$\psi(\theta, \zeta) := h(a(\zeta) e^{i\theta}, \zeta) = \log\left[\frac{\Phi(a(\zeta) e^{i\theta}, \zeta)}{a(\zeta) e^{i\theta}}\right].$$

For this function $\psi(\theta, \zeta)$, the following assertions can be easily verified.

Assertion 1. As $\zeta \rightarrow 0$ in \mathbb{C} , $\psi(\theta, \zeta)$ and all its derivatives with respect to θ converge uniformly to

$$\psi(\theta) := h(a e^{i\theta}, 0) = \log\left[\frac{f(a e^{i\theta})}{a e^{i\theta}}\right]$$

and the corresponding derivatives.

Assertion 2. In particular,

$$\psi(0, \zeta) = \log\left[\frac{\Phi(a(\zeta), \zeta)}{a(\zeta)}\right] \rightarrow \log\left[\frac{f(a)}{a}\right]$$

as $\zeta \rightarrow 0$.

Assertion 3. $\left.\frac{\partial}{\partial \theta} \psi(\theta, \zeta)\right|_{\theta=0} = ia(\zeta) \frac{\partial h}{\partial \zeta}(a(\zeta), \zeta) = 0$.

Assertion 4. We have

$$-K(\zeta) := \left.\frac{\partial^2}{\partial \theta^2} \psi(\theta, \zeta)\right|_{\theta=0} = -a(\zeta)^2 \frac{\partial^2 h}{\partial z^2}(a(\zeta), \zeta)$$

and, as $\zeta \rightarrow 0$,

$$-K(\zeta) \rightarrow -\frac{a^2 f''(a)}{f(a)} = -K < 0.$$

Assertion 5. By Assertion 1, we can write

$$\psi(\theta, \zeta) = \psi(0, \zeta) - \frac{1}{2} K(\zeta) \theta^2 + \theta^3 \beta(\theta, \zeta),$$

where $\beta(\theta, \zeta)$ is analytic and uniformly bounded in (θ, ζ) .

Assertion 6. By Assertions 3, 4, and 5, we can choose positive constants γ, δ , and η such that

$$\operatorname{Re} \psi(\theta, \zeta) \leq \operatorname{Re} \psi(0, \zeta) - \eta \theta^2$$

holds for $|\theta| \leq \delta$ and $|\zeta| \leq \gamma$.

Assertion 7. From Assertions 1, 2, and 6, we infer that, if $|\zeta|$ is sufficiently small, then

$$|\Phi(a(\zeta) e^{i\theta}, \zeta)/a(\zeta)| = \exp(\operatorname{Re} \psi(\theta, \zeta))$$

takes its maximum value at $\theta = 0$ and that, if $0 < |\theta| \leq \pi/d$,

$$\left| \frac{\Phi(a(\zeta) e^{i\theta}, \zeta)}{a(\zeta)} \right| < \left| \frac{\Phi(a(\zeta), \zeta)}{a(\zeta)} \right|.$$

Assertion 8. Let Δ_δ be defined by (7). From Assertions 1 and 6, if we take an ε such that $0 < \varepsilon < \frac{1}{3}\{(f(a)/a) - \Delta_\delta\}$ then, for ζ sufficiently near to 0, we have

$$\Delta_{\zeta, \delta} := \sup_{\delta \leq |\theta| \leq \pi/d} \left| \frac{\Phi(a(\zeta) e^{i\theta}, \zeta)}{a(\zeta)} \right| < \Delta_\delta + \frac{\varepsilon}{3} < \frac{f(a)}{a} - \frac{\varepsilon}{3} < \left| \frac{\Phi(a(\zeta), \zeta)}{a(\zeta)} \right|.$$

Now we can perform an analysis similar to that in Section 2.2. First, we rewrite (38) as

$$\begin{aligned} & \mathbb{E} \left[Z = n; \exp \left(\sum_{k \geq 0} \zeta t_k Y_k \right) \right] \\ &= \frac{a(\zeta)d}{2\pi n} \left(\frac{\Phi(a(\zeta), \zeta)}{a(\zeta)} \right)^n \int_{-\delta}^{\delta} \exp(-\frac{1}{2}nK(\zeta)\theta^2 + n\theta^3\beta(\theta, \zeta) + i\theta) d\theta + \mathcal{O} \left(\frac{1}{n} \Delta_{\zeta, \delta}^n \right). \end{aligned} \tag{39}$$

If we introduce

$$P_\zeta(\omega, \theta) := \exp(i\theta + \omega\beta(\theta, \zeta)) = \sum_{\ell, m \geq 0} c_{\ell m}(\zeta) \omega^\ell \theta^m$$

and

$$Q_\zeta(\omega, \theta) := c_{00}(\zeta) + c_{10}(\zeta)\omega + c_{01}(\zeta)\theta,$$

we then have

$$|P_\zeta(\omega, \theta) - Q_\zeta(\omega, \theta)| = \mathcal{O}(|\omega|^2) + \mathcal{O}(|\theta|^2),$$

where the estimate $\mathcal{O}(\cdot)$ is valid uniformly in ζ sufficiently near 0. Now replace δ , in the limit of integration in (39), by $\tau_n := n^{-1/3}$, and then replace $\exp(i\theta + n\theta^3\beta(\theta, \zeta))$ by $Q_\zeta(n\theta^3, \theta)$. The errors introduced in doing so are estimated as in Section 2.2, and we obtain

$$\begin{aligned} & \mathbb{E} \left[Z = n; \exp \left(\sum_{k \geq 0} \zeta t_k Y_k \right) \right] \\ &= \frac{a(\zeta)d}{2\pi n} \left(\frac{\Phi(a(\zeta), \zeta)}{a(\zeta)} \right)^n \left[\sum_{\ell+m \leq 1} c_{\ell m}(\zeta) n^\ell \int_{-\tau_n}^{\tau_n} \exp(-\frac{1}{2}nK(\zeta)\theta^2) \theta^{3\ell+m} d\theta + \mathcal{O}(n^{-3/2}) \right]. \end{aligned}$$

Finally, we extend the integration to $(-\infty, \infty)$, with a small error that is absorbed in the term $\mathcal{O}(n^{-3/2})$. We thus arrive at

$$\mathbb{E} \left[Z = n; \exp \left(\sum_{k \geq 0} \zeta t_k Y_k \right) \right] = \frac{a(\zeta)d}{n} \left(\frac{\Phi(a(\zeta), \zeta)}{a(\zeta)} \right)^n \sqrt{\frac{1}{2\pi nK(\zeta)}} (1 + \mathcal{O}(n^{-3/2})).$$

Note that the integrals for $\ell + m = 1$ vanish because of symmetry, and that the error estimate $\mathcal{O}(n^{-3/2})$ is uniform in ζ .

Combining this result with that of Theorem 1 or, rather, of Otter’s theorem, we get

$$E \left[\exp \left(\sum_{k \geq 0} \zeta t_k Y_k \right) \middle| Z = n \right] = \frac{a(\zeta)(\Phi(a(\zeta), \zeta)/a(\zeta))^n [K(\zeta)]^{-1/2} (1 + \mathcal{O}(n^{-3/2}))}{a(f(a)/a)^n [f(a)/a^2 f''(a)]^{1/2} (1 + \mathcal{O}(n^{-3/2}))}.$$

If we take a sequence $\zeta_n \rightarrow 0$ then $K(\zeta_n) \rightarrow a^2 f''(a)/f(a)$ and $a(\zeta_n) \rightarrow a$. Hence, we finally obtain

$$E \left[\exp \left(\sum_{k \geq 0} \zeta_n t_k Y_k \right) \middle| Z = n \right] \sim \left\{ \frac{a}{f(a)} \frac{\Phi(a(\zeta_n), \zeta_n)}{a(\zeta_n)} \right\}^n, \tag{40}$$

where $n \rightarrow \infty$ keeping $n = 1 \pmod{d(\Pi)}$.

Now let us take $\zeta_n = -1/n$, and suppose that $t_k \geq 0$. Noting that $f'(a) = f(a)/a$, it is easy to compute that

$$\log \Phi(a(-1/n), -1/n) = \log f(a) - \frac{1}{n} \left\{ \frac{1}{f(a)} \sum_{k \geq 0} p_k t_k a^k + \frac{a'(0)}{a} \right\} + \mathcal{O}(n^{-2})$$

and

$$\log a(-1/n) = \log a - \frac{1}{n} \frac{a'(0)}{a} + \mathcal{O}(n^{-2}).$$

Thus, we conclude from (40) that

$$E \left[\exp \left(- \sum_{k \geq 0} t_k \frac{Y_k}{n} \right) \middle| Z = n \right] \rightarrow \exp \left\{ - \sum_{k \geq 0} \frac{p_k a^k}{f(a)} t_k \right\},$$

completing the proof of part (i) of Theorem 3.

Let us remark that $\{\rho_k\}_{k \geq 0} := \{p_k a^k / f(a)\}_{k \geq 0}$ is a probability distribution on \mathbb{Z}_+ with mean 1:

$$\sum_{k \geq 1} k \rho_k = \frac{1}{f(a)} \sum_{k \geq 1} k p_k a^k = \frac{a f'(a)}{f(a)} = 1.$$

In order to prove part (ii), we expand $\log[\Phi(a(\zeta), \zeta)]$ and $\log a(\zeta)$ to second order for small ζ , to obtain

$$\begin{aligned} \log \left[\frac{a}{f(a)} \frac{\Phi(a(\zeta), \zeta)}{a(\zeta)} \right] &= \frac{1}{f(a)} \left(\frac{\partial \Phi}{\partial \zeta}(a, 0) \right) \zeta \\ &\quad + \left[\frac{1}{2 f(a)} \left(f''(a) a'(0)^2 + \frac{\partial^2 \Phi}{\partial \zeta^2}(a, 0) + 2 \frac{\partial^2 \Phi}{\partial \zeta \partial z}(a, 0) a'(0) \right) \right. \\ &\quad \left. - \frac{1}{2 f(a)^2} \left(f'(a) a'(0) + \frac{\partial \Phi}{\partial \zeta}(a, 0) \right)^2 + \frac{1}{2} \frac{a'(0)^2}{a^2} \right] \zeta^2 + \mathcal{O}(\zeta^3). \end{aligned}$$

On the other hand, differentiating $M(a(\zeta), \zeta) = 0$, we get

$$a'(0) = - \frac{1}{a f''(a)} \left\{ a \frac{\partial^2 \Phi}{\partial \zeta \partial z}(a, 0) - \frac{\partial \Phi}{\partial \zeta}(a, 0) \right\},$$

which, when inserted in the above formula, gives

$$\begin{aligned} \log \left[\frac{a}{f(a)} \frac{\Phi(a(\zeta), \zeta)}{a(\zeta)} \right] &= \frac{1}{f(a)} \left(\frac{\partial \Phi}{\partial \zeta}(a, 0) \right) \zeta \\ &+ \left[-\frac{f''(a)}{2f(a)} a'(0)^2 + \frac{1}{2f(a)} \frac{\partial^2 \Phi}{\partial \zeta^2}(a, 0) \right. \\ &\quad \left. - \frac{1}{2f(a)^2} \left(\frac{\partial \Phi}{\partial \zeta}(a, 0) \right)^2 \right] \zeta^2 + \mathcal{O}(\zeta^3). \end{aligned}$$

Letting $t_k \in \mathbb{R}$ and $\zeta = i/n^{1/2}$ in (40), we finally obtain

$$\begin{aligned} &\lim_{n \rightarrow \infty} \mathbb{E} \left[\exp \left(i \sum_{k \geq 0} t_k n^{1/2} \left(\frac{Y_k}{n} - \frac{p_k a^k}{f(a)} \right) \right) \middle| Z = n \right] \\ &= \lim_{n \rightarrow \infty} \exp \left(-in^{1/2} \sum_{k \geq 0} t_k \frac{p_k a^k}{f(a)} \right) \left\{ \frac{a}{f(a)} \frac{\Phi(a(i/n^{1/2}), i/n^{1/2})}{a(i/n^{1/2})} \right\}^n \\ &= \lim_{n \rightarrow \infty} \exp \left(-n \frac{1}{f(a)} \left(\frac{\partial \Phi}{\partial \zeta}(a, 0) \right) \frac{i}{n^{1/2}} \right) \left\{ \frac{a}{f(a)} \frac{\Phi(a(i/n^{1/2}), i/n^{1/2})}{a(i/n^{1/2})} \right\}^n \\ &= \exp \left(\frac{f''(a)}{2f(a)} a'(0)^2 - \frac{1}{2f(a)} \frac{\partial^2 \Phi}{\partial \zeta^2}(a, 0) + \frac{1}{2f(a)^2} \left(\frac{\partial \Phi}{\partial \zeta}(a, 0) \right)^2 \right) \\ &= \exp[-\frac{1}{2} V(\{t_k\})], \end{aligned}$$

where we have set

$$\begin{aligned} V(\{t_k\}) &= \frac{1}{f(a)} \sum_{k \geq 0} p_k t_k^2 a^k - \frac{1}{f(a)^2} \left(\sum_{k \geq 0} p_k t_k a^k \right)^2 \\ &\quad - \frac{1}{a^2 f(a) f''(a)} \left(\sum_{k \geq 1} k p_k t_k a^k - \sum_{k \geq 0} p_k t_k a^k \right)^2. \end{aligned}$$

In terms of ρ_k , introduced above, this can be expressed as

$$V(\{t_k\}) = \sum_{k \geq 0} t_k^2 \rho_k - \left(\sum_{k \geq 0} t_k \rho_k \right)^2 - \left[\sum_{k \geq 2} k(k-1) \rho_k \right]^{-1} \left(\sum_{k \geq 1} k t_k \rho_k - \sum_{k \geq 0} t_k \rho_k \right)^2.$$

Now let X be a \mathbb{Z}_+ -valued random variable with distribution $\{\rho_k\}$ and let the function $T : \mathbb{Z}_+ \rightarrow \mathbb{R}$ be defined by $T(k) = t_k$. Then $\mathbb{E}[X] = 1$ and

$$\sum_{k \geq 2} k(k-1) \rho_k = \mathbb{E}[X(X-1)] = \mathbb{E}[(X-1)^2] > 0$$

because the distribution $\{p_k\}$ is not degenerate to a point, by our basic assumptions. Therefore,

$$\begin{aligned} &\left(\sum_{k \geq 2} k(k-1) \rho_k \right) V(\{t_k\}) \\ &= \mathbb{E}[(X-1)^2] \{ \mathbb{E}[T(X)^2] - \mathbb{E}[T(X)]^2 - \{ \mathbb{E}[XT(X)] - \mathbb{E}[T(X)] \}^2 \} \\ &= \mathbb{E}[(X-1)^2] \mathbb{E}[\{T(X) - \mathbb{E}[T(X)]\}^2] - \mathbb{E}[(X-1)\{T(X) - \mathbb{E}[T(X)]\}]^2, \end{aligned}$$

which is actually nonnegative, by the Schwarz inequality applied to the random variables $X - 1$ and $T(X) - E[T(X)]$.

Now, suppose that $A \subset S$, $A \neq \emptyset$, and $|S \setminus A| \geq 2$. Suppose further that, for some $\lambda \in \mathbb{R}$, the relation

$$T(X) - E[T(X)] = \lambda(X - 1)$$

almost surely holds, where we assume that $T(k) = t_k = 0$ for $k \notin A$. Then if $\ell \in S \setminus A$, we have $-E[T(X)] = \lambda(X - 1)$, but in the case $|S \setminus A| \geq 2$, this is possible only if $\lambda = 0$. Thus, the inequality must be strict in the above-mentioned Schwarz inequality, unless $t_k = 0$ for all $k \in A$. Thus, the limiting Gaussian distribution of $\{n^{1/2}(n^{-1}Y_k - \rho_k)\}_{k \in A}$ conditioned on $\{Z = n\}$ is nondegenerate.

Conversely, suppose that $A \setminus S \neq \emptyset$. Since $Y_k = 0$ almost surely for $k \notin S$, the distribution of $\{Y_k\}_{k \in A}$ is already degenerate. If, on the other hand, $A = S \setminus \{\ell\}$ for some $\ell \in S$ then, since the relations $\sum_{k \in S} Y_k = Z$ and $\sum_{k \in S} kY_k = Z - 1$ yield

$$\sum_{k \in S \setminus \{\ell\}} (\ell - k)Y_k = (\ell - 1)Z + 1,$$

the conditional distribution $P(\{n^{1/2}(n^{-1}Y_k - \rho_k)\}_{k \in A} \in \cdot \mid Z = n)$ is concentrated on a hyperplane of dimension less than or equal to $|A| - 2$. Hence, its limiting Gaussian distribution is degenerate. This is also the case when $A = S$. The proof of Theorem 3 is now complete.

As an application of Theorem 3, we consider the following question, treated by Mahmoud [22]. Let \mathcal{C}_n be the totality of uniform binary trees of size n . (See Section 4.2 for the definition.) Furthermore, let $X_i(\gamma)$, $i = 0, 1, 2$, be the number of vertices of the tree $\gamma \in \mathcal{C}_n$ having i children. As we saw in Section 4.2, we have

$$Q_n(\{\gamma \in \mathcal{C}_n : X_i(\gamma) = n_i, i = 0, 1, 2\}) = P(Y_i = n_i, i = 0, 1, 2 \mid Z = n), \tag{41}$$

where Q_n is the uniform distribution on \mathcal{C}_n and P is the distribution of the Galton–Watson tree whose offspring distribution is the binomial distribution $\text{Bin}(2, p)$ with $0 < p < 1$. The right-hand side being independent of p , we can take $p = \frac{1}{2}$, in which case we have $a = f(a) = 1$ and $\rho_i = p_i a^i / f(a) = \frac{1}{4}$ for $i = 0, 2$, while $\rho_1 = p_1 a^1 / f(a) = \frac{1}{2}$. Moreover, $f'(a) = f''(a) = \frac{1}{2}$. We can now apply part (ii) of Theorem 3 to this P and, in view of (41), we obtain

$$Q_n\left(\left(\frac{X_0 - n/4}{n^{1/2}}, \frac{X_1 - n/2}{n^{1/2}}, \frac{X_2 - n/4}{n^{1/2}}\right) \in \cdot\right) \rightarrow N(0, \frac{1}{16}C), \quad n \rightarrow \infty,$$

where

$$C = \begin{pmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{pmatrix},$$

which is precisely Theorem 2 of [22] because, as a special case of (28), we have

$$E_{Q_n}[X_k] = \rho_k n + \mathcal{O}(1)$$

as $n \rightarrow \infty$ with $(\rho_0, \rho_1, \rho_2) = (\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$. Note that $d(\Pi) = 1$ and $S = A = \{0, 1, 2\}$ in our case, so that the matrix C is naturally degenerate by virtue of the second part of Theorem 3(ii). We also remark that we can apply (29) and (30) to show that the covariance matrix of (X_0, X_1, X_2) under the distribution Q_n is asymptotically equal to $nC/16$, as in Equation (2.5) of [22].

Acknowledgements

The author is grateful to Professors M. Drmota, H. Kesten, and T. Kumagai for valuable comments and information, and to Professors K. Fukuyama and Y. Higuchi for their interest in this work and for giving him the opportunity to give a lecture on the subject at Kobe University. Last but not least, the author also thanks the anonymous referee for some additional references and valuable comments.

This work was partially supported by a Grant-in-Aid for Scientific Research, the Ministry of Education, Sports, Science, Culture and Technology, Japan.

References

- [1] ALDOUS, D. (1991). The continuum random tree. II. An overview. In *Stochastic Analysis* (London Math. Soc. Lecture Notes Ser. **167**), eds M. T. Barlow and N. H. Bingham, Cambridge University Press, pp. 23–70.
- [2] ATHREYA, K. B. AND NEY, P. E. (1972). *Branching Processes*. Springer, New York.
- [3] BOYD, A. V. (1971). Formal power series and the total progeny in a branching process. *J. Math. Anal. Appl.* **34**, 565–566.
- [4] CONSUL, P. C. AND SHENTON, L. R. (1972). Use of Lagrange expansion for generating discrete generalized probability distributions. *SIAM J. Appl. Math.* **23**, 239–248.
- [5] COPSON, E. T. (1965). *Asymptotic Expansions*. Cambridge University Press.
- [6] DE BRUIJN, N. G. (1981). *Asymptotic Methods in Analysis*. Dover, New York.
- [7] DEMBO, A., MÖRTERS, P. AND SHEFFIELD, S. (2003). A large-deviation theorem for tree-indexed Markov chains. Preprint ArXiv:math.PR/0306045, University of California, Davis.
- [8] DRMOTA, M. (2004). Combinatorics and asymptotics on trees. *Cubo Mat. Educ.* **6**, 105–136.
- [9] DWASS, M. (1969). The total progeny in a branching process and a related random walk. *J. Appl. Prob.* **6**, 682–686.
- [10] GOOD, I. J. (1949). The number of individuals in a cascade process. *Proc. Camb. Phil. Soc.* **45**, 360–363.
- [11] GOOD, I. J. (1960). Generalizations to several variables of Lagrange’s expansion, with applications to stochastic processes. *Proc. Camb. Phil. Soc.* **56**, 367–380.
- [12] GOOD, I. J. (1975). The Lagrange distributions and branching processes. *SIAM J. Appl. Math.* **28**, 270–275.
- [13] HAN, G. N. (1996). Démonstration combinatoire des formules de Prodinger concernant les arbres binaires. *Séminaire Lotharingien de Combinatoire* **38**. Available at <http://www.mat.univie.ac.at/~slc/>.
- [14] HARRIS, T. (1989). *The Theory of Branching Processes*. Dover, New York.
- [15] HIRSCH, F. AND LACOMBE, G. (1999). *Elements of Functional Analysis* (Graduate Texts Math. **192**). Springer, New York.
- [16] HOFBAUER, J. (1982). Lagrange inversion. *Séminaire Lotharingien de Combinatoire* **6**. Available at <http://www.mat.univie.ac.at/~slc/>.
- [17] JAGERS, P. (1975). *Branching Processes with Biological Applications*. John Wiley, London.
- [18] KENNEDY, D. P. (1975). The Galton–Watson process conditioned on the total progeny. *J. Appl. Prob.* **12**, 800–806.
- [19] KESTEN, H. AND PITTEL, B. (1996). A local limit theorem for the number of nodes, the height, and the number of final leaves in a critical branching process tree. *Random Structures Algorithms* **8**, 243–299.
- [20] KRATTENTHALER, C. (1996). Comment on ‘A note on the distribution of the three types of nodes in uniform binary trees’ by Helmut Prodinger. *Séminaire Lotharingien de Combinatoire* **38**. Available at <http://www.mat.univie.ac.at/~slc/>.
- [21] KURATA, K. AND MINAMI, N. (2004). The equivalence of two constructions of Galton–Watson processes. Preprint mp_arc 04-295, Department of Mathematics, University of Texas at Austin. Available at http://www.ma.utexas.edu/mp_arc-bin/mpa?yn=04-295.
- [22] MAHMOUD, H. M. (1995). The joint distribution of the three types of nodes in uniform binary trees. *Algorithmica* **13**, 313–323.
- [23] NEVEU, J. (1986). Arbres et processus de Galton–Watson. *Ann. Inst. H. Poincaré Prob. Statist.* **22**, 199–207.
- [24] OTTER, R. (1949). The multiplicative processes. *Ann. Math. Statist.* **20**, 206–224.
- [25] PRODINGER, H. (1996). A note on the distribution of the three types of nodes in uniform binary trees. *Séminaires Lotharingien de Combinatoire* **38**. Available at <http://www.mat.univie.ac.at/~slc/>.
- [26] ROTE, G. (1996). Binary trees having a given number of nodes with 0, 1, and 2 children. *Séminaire Lotharingien de Combinatoire* **38**. Available at <http://www.mat.univie.ac.at/~slc/>.
- [27] SEVAST’YANOV, B. A. (1971). *Branching Processes*. Nauka, Moscow (in Russian).
- [28] TUTTE, W. T. (1964). The number of planted plane trees with a given partition. *Amer. Math. Monthly* **71**, 272–277.