The Galton–Watson tree conditioned on its height

Jochen Geiger and Götz Kersting*

Universität Frankfurt†

September 1998

Abstract

We give a probabilistic construction of the Galton–Watson tree conditioned on its height by decomposing it along the line of descent to the left–most particle at maximal height. This construction provides a representation of the final generation size as a sum of independent increments. Based on this representation we prove limit laws for the final generation size of a Galton–Watson tree conditioned on its height and total progeny, respectively.

1 Introduction

Let T denote the random family tree of a Galton–Watson branching process starting with a single founding ancestor, where each particle independently has probability p_k of producing k offspring. For a detailed definition and discussion of this process, we refer to [2], [1]. Regard T as a rooted planar tree with the distinguishable offspring of each vertex ordered from left to right. Let $\mu = \sum kp_k$ be the mean number of children per particle and denote by

*Research of both authors supported by the German Research Foundation DFG

†Fachbereich Mathematik, Postfach 11 19 32, D–60054 Frankfurt am Main, Germany

$Z_n(T)$ the nth generation size (= number of vertices with distance n from the root). We write $H(T)$ for the height of the tree (= maximal distance of a vertex from the root) and $q = P(Z_n(T) \to 0)$ for the extinction probability of the branching process. We assume $0 < p_0 < 1$ throughout, so that $P(H(T) = n) > 0$ for any n and $q = 1 \iff \mu \leq 1$.

In this paper we study the conditional structure of the family tree T given that $H(T) = n$. Clearly, particles no longer evolve independently of each other nor homogeneously in time. However, the Galton–Watson tree conditioned on its height satisfies some striking consistency and independence conditions. These properties allow to construct the tree inductively by attaching independent subtrees at the bottom of the line of descent to the left–most particle at maximal height. This provides a decomposition of the final generation size of a conditioned Galton–Watson tree into a sum of independent increments. The construction, which is the analog of the backward construction of the Galton–Watson tree conditioned on non–extinction in [4], is explained in section 2. In section 3 we prove a strong limit law for the final generation size for arbitrary offspring distribution. In the critical case with finite variance the limiting distribution of the final generation size is known to be the same, if instead the total population size of the tree were conditioned to be large [7]. We give a simple probabilistic proof of this fact. The concept of spinal decompositions as an approach to proving limit theorems for Galton–Watson processes goes back to Lyons, Pemantle, and Peres [10]. Similar and related constructions for the genealogical tree arising from a branching particle system have occured in [5], [3].

2 Growing a conditioned Galton–Watson tree from the top

Our starting point is the following simple but crucial observation: The law of the subtree founded by the left–most child of the root, who has a descendant

in the final generation of a Galton–Watson tree conditioned on height $n+1$, is the same as the law of a Galton–Watson tree conditioned on height n. This conditional law is independent of the first generation size and the rank of the child. The subtrees founded by the siblings to the left and right of the distinguished first generation particle are independent Galton–Watson trees conditioned on height strictly less than n and $n+1$, respectively.
More precisely, let $H(T)$ denote the height of the Galton–Watson tree T, i.e. for $n \geq 0$,

$$H(T) = n \;:\Longleftrightarrow\; Z_n(T) > 0,\; Z_{n+1}(T) = 0,$$

and let $T^{(i)}$, $1 \leq i \leq Z_1(T)$, be the subtrees founded by the first generation particles of T. For a finite tree T with $Z_1(T) > 0$ write $R(T)$ for the rank of the left–most such subtree with maximal height,

$$R(T) = \min\{1 \leq i \leq Z_1(T) |\, H(T^{(i)}) = H(T) - 1\}.$$

We sum up the properties stated above in the following lemma (we abbreviate $R = R(T)$, $H = H(T), \ldots$). The proof is immediate from the Galton–Watson conditions of independence and time–homogeneity.

Lemma 2.1 *The subtrees $T^{(i)}$, $1 \leq i \leq Z_1$, are conditionally independent given $\{R = j, Z_1 = k, H = n+1\}$, $1 \leq j \leq k < \infty$, $n \geq 0$, with*

$$\mathcal{L}(T^{(i)} \,|\, R = j,\, Z_1 = k,\, H = n+1) = \begin{cases} \mathcal{L}(T |\, Z_n = 0), & 1 \leq i \leq j-1; \\ \mathcal{L}(T |\, H = n), & i = j; \\ \mathcal{L}(T |\, Z_{n+1} = 0), & j+1 \leq i \leq k. \end{cases}$$

The conditional joint distribution of R and Z_1 is

$$P(R = j, Z_1 = k \,|H = n+1) = c_n\, p_k\, P(Z_n = 0)^{j-1} P(Z_{n+1} = 0)^{k-j}, \quad (1)$$

where $c_n = P(H = n)/P(H = n+1)$, $n \geq 0$.

The properties stated in lemma 2.1 suggest to grow the conditioned Galton–Watson tree from the top backwards. We now construct an increasing sequence of trees $(\bar{T}_n)_{n\geq 0}$ such that $\bar{T}_n$ has law $\mathcal{L}(T|H = n)$. The tree $\bar{T}_{n+1}$ is obtained from $\bar{T}_n$ by attaching independent subtrees at the bottom of the line of descent of the left–most particle at maximal height.

Construction. Let (V_{n+1}, W_{n+1}), $n \geq 0$, be a sequence of independent random variables with distribution (1),

$$P(V_{n+1} = j, W_{n+1} = k) = c_n \, p_k \, P(Z_n = 0)^{j-1} P(Z_{n+1} = 0)^{k-j},$$

and let $\bar{T}_0$ be a Galton–Watson tree of height 0 (i.e. $\bar{T}_0$ consists of its root only). Inductively construct $\bar{T}_{n+1}$, $n \geq 0$, by the following procedure:

- Let the first generation size of $\bar{T}_{n+1}$ be W_{n+1}.
- Let $\bar{T}_n$ be the subtree founded by the V_{n+1}th first generation particle of $\bar{T}_{n+1}$.
- Attach independent Galton–Watson trees conditioned on height strictly less than n to the $V_{n+1} - 1$ siblings to the left of the distinguished first generation particle.
- Attach independent Galton–Watson trees conditioned on height strictly less than $n + 1$ to the $W_{n+1} - V_{n+1}$ siblings to the right of the distinguished first generation particle.

The tree $\bar{T}_n$ has the same probabilistic structure as a Galton–Watson tree conditioned on height n.

Proposition 2.2 *Suppose* $0 < p_0 < 1$. *Then*

$$\mathcal{L}(\bar{T}_n) = \mathcal{L}(T|H = n) \quad \textit{for any } n \geq 0.$$

Proof. Per induction on n using the properties stated in Lemma 2.1. □

Remark. For any offspring distribution $(p_k)_{k\geq 0}$, (V_n, W_n) has a weak limit (V_∞, W_∞), say, with distribution

$$P(V_\infty = j, W_\infty = k) \;=\; c_\infty \, p_k \, q^{k-1}, \;\; 1 \leq j \leq k, \tag{2}$$

where

$$c_\infty := \lim_{n\to\infty} c_n \;=\; \lim_{n\to\infty} \frac{P(H=n)}{P(H=n+1)} \;=\; \Big(\sum_{k=1}^{\infty} k q^{k-1} p_k\Big)^{-1} \geq 1. \tag{3}$$

Note that V_∞ is conditionally uniform on $\{1, \cdots, k\}$ given $W_\infty = k$, $k \geq 1$. This displays the fact that, as $n \to \infty$, all particles in the final generation of a Galton–Watson tree conditioned on height n are descendants of the same first generation particle. In subcritical and critical cases $\mu \leq 1$ the extinction probability q equals 1, i.e. W_∞ has the size–biased distribution $(\mu^{-1} k p_k)_{k\geq 1}$.

The particles in the final generation n of $\bar{T}_n$ are the particles in generation $n+1$ of $\bar{T}_{n+1}$ descended from the distinguished first generation particle. Any other particle in the final generation $n+1$ of $\bar{T}_{n+1}$ is a descendant of one of the $W_{n+1} - V_{n+1}$ siblings to the right of the distinguished first generation particle. This provides the following representation of the final generation size of a Galton–Watson tree conditioned on its height:

Let $Z_{n,i}$ and X_{n+1}, $n \geq 0$, $i \geq 1$, be independent random variables where

$$\mathcal{L}(Z_{n,i}) \;=\; \mathcal{L}(Z_n | Z_{n+1} = 0), \tag{4}$$

and X_{n+1} has distribution $\mathcal{L}(W_{n+1} - V_{n+1})$,

$$P(X_{n+1} = k) = c_n P(Z_{n+1} = 0)^k \sum_{j=k+1}^{\infty} p_j P(Z_n = 0)^{j-(k+1)}, \;\; k \geq 0. \tag{5}$$

Define $\bar{Z}_0 := 1$ and

$$\bar{Z}_{n+1} := \bar{Z}_n + \sum_{i=1}^{X_{n+1}} Z_{n,i}, \quad n \geq 0. \tag{6}$$

Corollary 2.3 *Suppose* $0 < p_0 < 1$. *Then*

$$\mathcal{L}(\bar{Z}_n) = \mathcal{L}(Z_H | H = n) \quad \textit{for any } n \geq 0.$$

Proof. By construction of $(\bar{T}_n)_{n \geq 0}$ and Proposition 2.2. □

Remark. A representation similar to (6) holds jointly for the last $k \leq n$ generations of a Galton–Watson tree conditioned on height n. However, note that also siblings to the left of particles in the distinguished line of descent can have descendants in generations $n - j$, $j \geq 1$. Alternatively, one can construct the tree along the backbone of the left–most particle in generation $n - k + 1$.

3 Limit laws for the final generation size

Our first result states that, as $n \to \infty$, the final generation size of a Galton–Watson tree conditioned on height n has a proper limit with finite mean. The statement holds for arbitrary offspring distribution. We remark that strong convergence of the joint distribution of the final k generation sizes can be derived in much the same way (compare the remark following Corollary 2.3).

Proposition 3.1 *Suppose* $0 < p_0 < 1$. *Then*

$$\bar{Z}_n \overset{a.s.}{\longrightarrow} \bar{Z}_\infty \quad \textit{as} \quad n \to \infty, \tag{7}$$

$$E\bar{Z}_\infty = \lim_{n \to \infty} E\bar{Z}_n < \infty. \tag{8}$$

Proof. Since $(\bar{Z}_n)_{n \geq 0}$ has non–negative increments, it has an almost sure limit $\bar{Z}_\infty$, say, and $\lim_{n \to \infty} E\bar{Z}_n = E\bar{Z}_\infty \leq \infty$. By (6), the independence properties of the X_{n+1} and $Z_{n,i}$, and Corollary 2.3

$$\begin{aligned} E\bar{Z}_{n+1} - E\bar{Z}_n &= EX_{n+1}\, E(Z_n | Z_{n+1} = 0) \\ &= EX_{n+1}\, E(Z_n | Z_n > 0, Z_{n+1} = 0)\, P(Z_n > 0 | Z_{n+1} = 0) \\ &= EX_{n+1}\, E\bar{Z}_n\, P(Z_n > 0 | Z_{n+1} = 0). \end{aligned} \tag{9}$$

For any increasing sequence $(a_n)_{n\geq 0}$ with $a_0 > 0$,

$$\log a_{n+1} - \log a_n \;\leq\; a_n^{-1}(a_{n+1} - a_n) \;\leq\; a_0^{-1}(a_{n+1} - a_n)$$

and, consequently,

$$\lim_{n\to\infty} a_n < \infty \iff \sum_{n=0}^{\infty} a_n^{-1}(a_{n+1} - a_n) < \infty.$$

Take $a_n = E\bar{Z}_n$, then (9) implies

$$E\bar{Z}_\infty < \infty \iff \sum_{n=0}^{\infty} EX_{n+1}\, P(Z_n > 0 | Z_{n+1} = 0) < \infty. \tag{10}$$

(Note that the quantity on the right–hand side of (10) is the expected number of 'contributing' siblings of the particles in the distinguished line of descent.) We begin with an upper bound of the first factor of the summands on the right–hand side of (10). We write $f \asymp g$ if there exist positive constants c_1 and c_2 such that $c_1 f(n) \leq g(n) \leq c_2 f(n),\ n \geq 0$.

$$\begin{aligned} EX_{n+1} &\overset{(3),(5)}{\asymp} \sum_{k=1}^{\infty} k\, P(Z_{n+1} = 0)^k \sum_{j=k+1}^{\infty} p_j P(Z_n = 0)^{j-(k+1)} \\ &\leq \sum_{j=2}^{\infty}\sum_{k=1}^{j-1} k\, p_j P(Z_{n+1} = 0)^{j-1} \\ &\leq \sum_{j=1}^{\infty} j^2 p_j P(Z_{n+1} = 0)^{j-1}. \end{aligned} \tag{11}$$

For the other term in the sum on the right–hand side of (10) note that

$$\begin{aligned} P(Z_n > 0 | Z_{n+1} = 0) &= \frac{P(H = n)}{P(Z_{n+1} = 0)} \\ &\asymp \quad P(H = n) \overset{(3)}{\asymp} P(H = n+1). \end{aligned} \tag{12}$$

In view of (11), (12), and the binomial theorem, we conclude (here c denotes some positive constant)

$$\sum_{n=0}^{\infty} EX_{n+1}\, P(Z_n > 0 | Z_{n+1} = 0)$$

$$
\begin{aligned}
&\le \; c \sum_{n=0}^{\infty} \sum_{j=1}^{\infty} j^2\, p_j P(Z_{n+1}=0)^{j-1} P(H=n+1) \\
&= \; c \sum_{j=1}^{\infty} j p_j \sum_{n=0}^{\infty} j\, P(H \le n)^{j-1} P(H=n+1) \\
&\le \; c \sum_{j=1}^{\infty} j p_j \sum_{n=0}^{\infty} P(H \le n+1)^j - P(H \le n)^j \\
&\le \; c \sum_{j=1}^{\infty} j p_j P(H < \infty)^j \; < \; \infty, \qquad (13)
\end{aligned}
$$

since either $\mu = \sum_{j=1}^{\infty} j p_j \le 1$ or $q = P(H < \infty) < 1$ (in fact, the inequality $\sum_{j \ge 1} j p_j q^j \le q$ holds). The claim of Proposition 3.1 follows from (10) and (13). □

From Proposition 3.1 we derive the following limit law for the final generation size of a Galton–Watson tree with arbitrary offspring distribution conditioned on height n. In the critical, finite variance case the result is due to Kesten, Ney, and Spitzer [6]. Seneta [11] removed the second moment assumption. In fact, no conditions on the offspring distributions are required.

Theorem 3.2 *Let $\nu_n = \mathcal{L}(Z_H | H = n)$, $n \ge 0$, and suppose $0 < p_0 < 1$. The sequence $(\nu_n)_{n \ge 0}$ converges in a strong sense,*

$$\sum_{n=0}^{\infty} ||\nu_{n+1} - \nu_n|| < \infty, \qquad (14)$$

where $||\cdot||$ denotes total variation norm. If ν_∞ denotes the limiting distribution of $(\nu_n)_{n \ge 0}$, then

$$\nu_\infty(k) = p_0^k \pi_k, \quad k \ge 1, \qquad (15)$$

where $\pi_k = \lim_{n \to \infty} P(Z_n = k)/P(H = n)$ is a solution of

$$\pi_k \; = \; c_\infty \sum_{j=1}^{\infty} \pi_j\, P(Z_1 = k | Z_0 = j). \qquad (16)$$

Proof. By Corollary 2.3 we have $\nu_n = \mathcal{L}(\bar{Z}_n)$. Consequently,

$$||\nu_{n+1} - \nu_n|| \leq 2P(\bar{Z}_{n+1} \neq \bar{Z}_n).$$

Since $(\bar{Z}_n)_{n\geq 0}$ has independent non–negative, integer–valued increments, $\sum_{n=0}^{\infty}(\bar{Z}_{n+1} \neq \bar{Z}_n) < \infty$ iff $P(\bar{Z}_\infty < \infty) = 1$ (by the Borel–Cantelli lemma). Clearly, (8) implies $\bar{Z}_\infty < \infty$ a.s., which establishes (14). For the representation of ν_∞ in (15) note that

$$P(Z_H = k|H = n) \;=\; \frac{P(Z_n = k)\, p_0^k}{P(H = n)} \;\to\; \nu_\infty(k) \quad \text{as} \quad n \to \infty,$$

which implies existence of $\pi_k := \lim_{n\to\infty} P(Z_n = k)/P(H = n)$. Finally, for any $k \geq 1$,

$$\begin{aligned} \frac{P(Z_{n+1} = k)}{P(H = n+1)} \;&=\; c_n \sum_{j=1}^{\infty} \frac{P(Z_n = j)}{P(H = n)} P(Z_{n+1} = k|Z_n = j) \\ &=\; c_n \sum_{j=1}^{\infty} P(Z_n = j|H = n)\, p_0^{-j} P(Z_1 = k|Z_0 = j). \quad (17) \end{aligned}$$

To justify that summation and limiting procedures may be interchanged on the right–hand side of (17), note that the non–negativity of the increments of $(\bar{Z}_n)_{n\geq 0}$ implies

$$P(\bar{Z}_\infty = j) \;\geq\; P(\bar{Z}_n = j)P(\bar{Z}_\infty = \bar{Z}_n) \;\geq\; P(\bar{Z}_n = j)P(\bar{Z}_\infty = \bar{Z}_1), \quad (18)$$

where the second factor on the right–hand side of (18) is positive by (14) and the fact, that $P(\bar{Z}_{m+1} = \bar{Z}_m) > 0$ for any $m \geq 1$ due to the assumption $p_0 > 0$. Hence, by (17), Fatou's lemma and the dominated convergence theorem,

$$\begin{aligned} \pi_k \;&=\; c_\infty \lim_{n\to\infty} \sum_{j=1}^{\infty} P(\bar{Z}_n = j)\, p_0^{-j} P(Z_1 = k|Z_0 = j) \\ &=\; c_\infty \sum_{j=1}^{\infty} \nu_\infty(j)\, p_0^{-j} P(Z_1 = k|Z_0 = j), \end{aligned}$$

which is equation (16). □

In the critical case $\mu = 1$ with finite variance $\sigma^2 = \sum k(k-1)p_k < \infty$, the limiting distribution of the final generation size is the same, if the Galton–Watson tree T is conditioned on its total population size instead,

$$\lim_{n\to\infty,\, n-1\in h\mathbb{N}} P(Z_H = k | N = n) \;=\; \nu_\infty(k),\;\; k \geq 1, \tag{19}$$

where $N(T) = \sum_{k=0}^{\infty} Z_k(T)$ and h denotes the span of the offspring distribution $(p_k)_{k\geq 0}$ (see Theorem 2 in [7]). The criticality assumption is crucial for (19) to hold: Suppose that for some $\theta > 0$,

$$p'_k = c\theta^k p_k,\;\; k \geq 0,\;\text{ and }\; \textstyle\sum_{k\geq 0} p'_k = 1.$$

Then the distribution of T conditioned on its total population size being n is the same for the two offspring distributions $(p_k)_{k\geq 0}$ and $(p'_k)_{k\geq 0}$, whereas the distributions ν_∞ and ν'_∞ differ (e.g., if $p_2 = 1 - p_0 = \varepsilon$, then $\nu_\infty(2) \to 1$ as $\varepsilon \to 0$). In [7] the limit law (19) is derived from a joint local limit theorem for H, Z_H, and N. We give here a simple probabilistic proof of a somewhat weaker result.

Theorem 3.3 *Suppose* $\mu = \sum_{k=1}^{\infty} kp_k = 1$, $\sigma^2 = \sum_{k=1}^{\infty} k(k-1)p_k < \infty$, *and* $p_0 > 0$. *Then*

$$\lim_{n\to\infty} P(Z_H = k | N \geq n) \;=\; \nu_\infty(k), \quad k \geq 1, \tag{20}$$

where ν_∞ *is the limit of* $\mathcal{L}(Z_H | H = n)$ *from Theorem 3.2.*

Proof. We will show that, as $m \to \infty$, Z_H is conditionally independent of N given $H = m$. To this end recall the construction of $\bar{T}_m$ along the line of descent to the left–most particle at maximal height m and let v_j, $0 \leq j \leq m$, be the distinguished particle's ancestor in generation $m - j$. Clearly, v_j is a child of v_{j+1}; v_m is the founding ancestor, and v_0 is the distinguished particle itself. Decompose the total population size of $\bar{T}_m$ by writing

$$N(\bar{T}_m) = 1 + \textstyle\sum_{j=1}^{m} Y_j,$$

where Y_j is the number of particles in $\bar{T}_m$, whose most recent ancestor in the distinguished line of descent is v_j. By construction of $\bar{T}_m$, the random variables Y_j, $1 \le j \le m$, are independent and the law of $Y_j - 1$ can be represented as a random sum of independent copies of the total population size of Galton–Watson trees conditioned on extinction at generation $j-1$ and j, respectively. In particular, the law of Y_j does not depend on m and

$$\begin{aligned} P(Y_j \ge n) &\ge P(X_j \ge 1)\, P(N \ge n | H < j) \\ &\overset{(2)}{\asymp} P(H < j | N \ge n)\, P(N \ge n), \quad n, j \ge 1, \end{aligned} \tag{21}$$

where $N = N(T)$, $H = H(T)$. To estimate the right–hand side of (21) we use the following asymptotics for the height and the size of a Galton–Watson tree T,

$$\lim_{n\to\infty,\, n-1\in h\mathbb{N}} n^{\frac{3}{2}}\, P(N = n) = \frac{h}{\sigma\sqrt{2\pi}}; \tag{22}$$

$$\mathcal{L}(n^{-\frac{1}{2}} H | N = n) \underset{n\to\infty,\, n-1\in h\mathbb{N}}{\longrightarrow} \frac{2W^*}{\sigma}, \tag{23}$$

where W^* is the maximum of standard Brownian excursion of duration 1 (see [8], Lemma 2.1.4 and Theorem 2.4.3). It is immediate from (22) that

$$\mathcal{L}(n^{-1} N | N \ge n) \overset{d}{\longrightarrow} U^{-2} \quad \text{as } n \to \infty, \tag{24}$$

where U is uniformly distributed on $(0, 1)$. Consequently,

$$\mathcal{L}(n^{-\frac{1}{2}} H | N \ge n) \overset{d}{\longrightarrow} \frac{2W^*}{\sigma U} \quad \text{as } n \to \infty, \tag{25}$$

with W^* and U independent. From (21), (22), and (25) we obtain the following asymptotic lower bound for the tail of $\mathcal{L}(Y_j)$.

$$\liminf_{j\to\infty} j P(Y_j \ge x j^2) \ge c x^{-\frac{1}{2}} P(2W^* \le x^{-\frac{1}{2}} \sigma U) > 0, \quad x > 0. \tag{26}$$

Estimating $\bar{N}_m = N(\bar{T}_m)$ through the maximum of the Y_j, $m/2 \le j \le m$, we obtain from (26),

$$\liminf_{m\to\infty} P(\bar{N}_m \ge x m^2) > 0, \quad x \ge 0. \tag{27}$$

Now recall that $\mathcal{L}(Z_H|H = m, N \geq n) = \mathcal{L}(\bar{Z}_m|\bar{N}_m \geq n)$ by Proposition 2.2. Hence, for any $\varepsilon > 0$,

$$\begin{aligned}
&|P(Z_H = k|N \geq n) - \nu_\infty(k)| \\
&\quad \leq \sum_{m=1}^{\infty} |P(Z_H = k|H = m, N \geq n) - \nu_\infty(k)|\; P(H = m|N \geq n) \\
&\quad \leq \sup_{m \geq \varepsilon\sqrt{n}} |P(\bar{Z}_m = k|\bar{N}_m \geq n) - \nu_\infty(k)| \;+\; P(H \leq \varepsilon\sqrt{n}|N \geq n). \qquad (28)
\end{aligned}$$

We may assume $\nu_\infty(k) > 0$ since $\nu_\infty(k) = 0$ implies $P(\bar{Z}_m = k) = 0$ for any $m \geq 1$. Then (27) implies

$$\begin{aligned}
&\sup_{m \geq \varepsilon\sqrt{n}} |P(\bar{Z}_m = k|\bar{N}_m \geq n) - \nu_\infty(k)| \\
&\quad \asymp \sup_{m \geq \varepsilon\sqrt{n}} \left|\nu_\infty(k)^{-1} P(\bar{N}_m \geq n,\, \bar{Z}_m = k) - P(\bar{N}_m \geq n)\right| \\
&\quad \leq \sup_{m \geq \varepsilon\sqrt{n}} |P(\bar{N}_m \geq n|\bar{Z}_\infty = k) - P(\bar{N}_m \geq n)| \\
&\qquad\qquad + \nu_\infty(k)^{-1} \sup_{m \geq \varepsilon\sqrt{n}} P(\bar{Z}_m \neq \bar{Z}_\infty). \qquad (29)
\end{aligned}$$

The second term on the right–hand side of (29) goes to 0 as $n \to \infty$ by Proposition 3.1. We will show below that

$$|P(\bar{N}_m \geq n|\bar{Z}_\infty = k) - P(\bar{N}_m \geq n)| \;\leq\; P(\sigma > m), \qquad (30)$$

where σ is a finite coupling time depending on k but not on m or n. Hence, if we pass to the limit $n \to \infty$ in (28), then (29), (30) and the limit law (25) imply

$$\limsup_{n \to \infty} |P(Z_H = k|N \geq n) - \nu_\infty(k)| \leq P(2W^* \leq \varepsilon\sigma U) \qquad (31)$$

for any $\varepsilon > 0$. Then let $\varepsilon \to 0$ in (31) to establish (20). We finally explain the coupling that leads to (30): Let $(Y_j', Z_j')_{j \geq 1}$ have distribution $\mathcal{L}((Y_j, \bar{Z}_j)_{j \geq 1}|\bar{Z}_\infty = k)$ and let

$$\tau = \min\{j \geq 0|\bar{Z}_j = \bar{Z}_\infty\}, \quad \tau' = \min\{j \geq 0|Z_j' = Z_\infty'\}.$$

Observe that for any $j \geq 1$

$$\mathcal{L}((Y_i')_{i \geq j} | \, j > \tau') = \mathcal{L}((Y_i)_{i \geq j} | \, j > \tau), \tag{32}$$

since Y_i is conditionally independent of $\bar{Z}_\infty = k$ given $i > \tau$. Also, note that (30) is the coupling inequality for the coupling time

$$\sigma = \min\{j > \tau \vee \tau' | \textstyle\sum_{i=1}^{j} Y_i = \sum_{i=1}^{j} Y_i'\}.$$

In view of (32) we have to produce a coupling for sums of independent but not identically distributed random variables. However, note that the Y_j have a weak limit by (2) and the fact that $\{H \leq m\} \nearrow \{H < \infty\}$. It is well–known that in this case successfull couplings exist (a Mineka or Ornstein coupling will do, see e.g. [9], Chap. 2), i.e. $\sigma < \infty$ a.s. □

References

[1] Asmussen, S. and Hering, H. (1983) *Branching Processes,* Birkhäuser, Boston.

[2] Athreya, K. B. and Ney, P. (1972) *Branching Processes*, Springer, New York.

[3] Chauvin, B., Rouault, A., and Wakolbinger, A. (1991) Growing conditioned trees, *Stoch. Proc. Appl.* **39**, 117–130.

[4] Geiger, J. (1997) Elementary new proofs of classical limit theorems for Galton–Watson processes, to appear in *J. Appl. Prob.* **36**

[5] Kallenberg, O. (1977) Stability of critical cluster fields. *Math. Nachrichten* **77**, 7–43.

[6] Kesten, H., Ney, P., and Spitzer, F. (1966) The Galton–Watson process with mean one and finite variance. *Theor. Prob. Appl.* **11**, 513–540.

[7] 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 and Algorithms* **8**, 243–299.

[8] Kolchin, V.F. (1986) *Random Mappings*, Optimization Software, New York.

[9] Lindvall, T. (1992) *Lectures on the Coupling Method*, Wiley, New York.

[10] Lyons, R., Pemantle, R., and Peres, Y. (1995) Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes. *Ann. Prob.* **25**, 1125–1138.

[11] Seneta, E. (1967) The Galton–Watson process with mean one, *J. Appl. Prob.* **4**, 489–495.