On the Height Profile of a Conditioned Galton-Watson Tree *

Götz Kersting
University of Frankfurt am Main†

March 1998

Abstract

Recently Drmota and Gittenberger (1997) proved a conjecture due to Aldous (1991) on the height profile of a Galton-Watson tree with an offspring distribution of finite variance, conditioned on a total size of n individuals. The conjecture states that in distribution its shape, more precisely its scaled height profile coincides asymptotically with the local time process of a Brownian excursion of duration 1. We give a proof of the result, which extends to the case of an infinite variance offspring distribution. This requires a different strategy, since in the infinite variance case there is no longer a relationship to the local time of Brownian resp. Lévy excursions.

AMS classification numbers: Primary 60J80, Secondary 60F17
Key words: Galton-Watson trees, Lévy-excursions, functional limit theorems

1 Introduction and main result

In this paper we analyse the shape of a Galton-Watson tree, conditioned to have a total number of n individuals. More precisely we study its asymptotic height profile, as $n \to \infty$.

By a tree t we mean a rooted, ordered, finite tree. We consider the vertices to represent individuals, such that t can be regarded as the family tree of the progeny of some founding ancestor (the root). It is assumed that among siblings there is an order (of birth), which allows to imbed such trees into the plane. The total number of vertices of t, its *size* is denoted by $s(t)$. We shall be interested in the way, in which the size of the different generations vary along the tree. An individual i belongs to the k'th generation, if the path from i to the root contains exactly k edges. Let z_k denote the number of individuals in generation $k = 0, 1, \ldots$. The sequence

*partially supported by the German Research Foundation DFG
†Fachbereich Mathematik, Postfach 11 19 32, D-60054 Frankfurt/Main

$1 = z_0, z_1, z_2, \ldots$ is called the *height profile* of t (possibly after a suitable renormalisation).

For a Galton-Watson tree T the number of children of the different individuals are assumed to be independent and identically distributed random variables. $p_x, x = 0, 1, \ldots$ denotes the probability, that an individual possesses x children, and Z_k is the number of individuals in generation k. We shall consider T, conditioned on a size $s(T) = n$. Such a conditioned Galton-Watson tree will be abbreviated as $CGW(n)$-tree. Since it contains n individuals, it is natural to consider its height profile

$$H_u^n = a_n^{-1} Z_{[nu/a_n]},\ u \geq 0,$$

rescaled with some sequence (a_n) of positive numbers. The aim is to choose (a_n) such, that H^n converges in distribution. Then a_n gives the magnitude of the breadth of the $CGW(n)$-tree, whereas n/a_n indicates the order of $h(T) = \max\{k \mid Z_k > 0\}$, the height of the $CGW(n)$-tree. - We assume:

Assumption A

1.) The offspring distribution $(p_x)_x$ has mean 1:

$$\sum_{x=0}^{\infty} x p_x = 1.$$

The greatest common divisor of all x with $p_x > 0$ is 1.
2.) There are positive numbers a_n such that $a_n^{-1}(\xi_1 + \ldots + \xi_n - n)$ converges in distribution to a non-degenerate limit law ν, as $n \to \infty$. Here $\xi_1, \xi_2, \ldots$ denote independent random variables with distribution $(p_x)_x$.

The case of an offspring distribution with a finite mean can be treated in much the same way and is not a genuine generalisation (see Kennedy [25]). Similar the assumption on the g.c.d. may be removed. Criticality and a g.c.d. equal to 1 are assumed just for convenience.

The second assumption has been completely analysed. It has several implications, which are discussed in chapter XVII in Feller [16] and in the monograph of Gnedenko and Kolmogorov [18]. In particular (p_x) is in the domain of attraction of a stable law. Its index α belongs to $(1, 2]$, since we deal with an offspring distribution of finite mean. This implies $a_n = o(n)$. The limit law $\nu = \nu_\alpha$ is determined by α up to a scaling constant. For $\alpha = 2$ this is simply the normal law, and for $1 < \alpha < 2$ it is a one-sided stable law, because the negative tail of the offspring distribution is zero.

We shall prove, that under these assumptions the height profile H^n converges in distribution, where the a_n are just the numbers, given in the assumption. The limiting process $H = (H_u)_u$ turns out to be a functional of certain normalized excursions: $H = \psi(Y)$. By a *normalized excursion* we understand a process $Y = (Y_s)_{0 \leq s \leq 1}$, such that $Y_0 = Y_1 = 0$ and

$\inf_{\delta \le s \le 1-\delta} Y_s > 0$ for all $\delta > 0$. H will be obtained from Y as follows: The corresponding *cumulative height profile*

$$C_u = \int_0^u H_v \, dv, \; u \ge 0,$$

is given by

$$C_u = \sup\{s \le 1 \; : \; \int_0^s \frac{dt}{Y_t} \le u\}.$$

H, having a.s. paths continuous from the right, is completely determined by the cumulative height process C. If $\int_{0+} dt/Y_t < \infty$, then $h > 0$ and $H_u > 0$ for all $u \in (0, h)$, also $C_u \to 1$, as $u \to h$. Then we say, that H is *non-degenerate*. The case $\int_{0+} dt/Y_t = \infty$, implying $C \equiv 0$ and $H \equiv 0$, is in the sequel of no significance. - The same transformation already appears in the paper [29] by Lamperti, who used it to transform Lévy-processes into continuous branching processes. Our theorem is in a way a conditioned version of Lamperti's result.

The normalized excursions are obtained as follows. There is a unique Lévy-process $X = (X_s)_{s \ge 0}$ (a process with independent and stationary increments), such that $X_0 = 0$ and the law ν is the distribution of X_1, and to each of these processes there belongs a normalized excursion Y. This is explained in detail in Bertoin's monograph [5] (in particular chapter VIII.4).

Theorem 1 *Let H^n be the scaled height profile of a CGW(n)-tree, satisfying assumption A. Then, as $n \to \infty$, H^n converges in distribution to the process $H = \psi(Y)$, derived from the corresponding normalized Lévy-excursion Y. H is a.s. non-degenerate.*

In general the excursions contain jumps. These jumps also show up in H. Thus we regard H^n and H as random elements in the space of cádlág-functions, endowed with the usual Skorohod J_1 topology (compare [14]).

A neat case is that of an offspring distribution with finite variance:

$$\sigma^2 = \sum_{x=0}^{\infty} x^2 p_x \; - 1 < \infty.$$

Then, choosing $a_n = \sigma n^{1/2}$, Y is a normalized Brownian excursion. In this situation the limiting process allows another appealing description: $(H_u)_u \stackrel{d}{=} (\frac{1}{2} L_{u/2})_u$, where L denotes the local time process of a normalized Brownian excursion. The reason is: Up to a factor $1/2$ L is related to a normalized Brownian excursion in the same way, as H is derived above from Y, which follows from a result of Jeulin [19] (compare Biane [6], Théorème 3). This version of Theorem 1 has been observed by Aldous in special cases (as the geometric offspring distributions) and conjectured in the general finite variance case, compare [3]. A first proof of the conjecture was given by Drmota and Gittenberger [11]. They mastered the formidable task to obtain convergence of the finite-dimensional distributions as well as tightness, using generating functions and thereby generalizing work of Kennedy [25] on the one-dimensional distributions. Pitman [34] surrounded the difficulties by imbedding the problem into the context of convergence of strong solutions of stochastic differential equations. The limiting distribution of the maximum of the height profile, the 'width' of the family tree, has been found in the finite variance case already by Takács [39].

The relations to stochastic analysis can be further developed in the case $\sigma^2 < \infty$. Note that the defining equation for C can be written as

$$u = \int_0^{C_u} \frac{dt}{Y_t} \text{ for } 0 < u \leq h, \text{ and } C_u = 1 \text{ for } u > h,$$

where

$$h = \int_0^1 \frac{dt}{Y_t}.$$

h is the asymptotic height of the rescaled tree, as will become clear in the next section. By differentiation with respect to u we get

$$H_u = \frac{dC_u}{du} = Y(C_u).$$

Now a Brownian excursion Y solves the stochastic equation $dY = dW + (1 - \frac{Y^2}{1-t})\frac{dt}{Y}$, with a standart Brownian motion W (compare [36], chapter IV, section (40.4)). Viewing C as a time-change, this leads to the stochastic equation

$$dH = \sqrt{H} dB + \left(1 - \frac{H^2}{1-C}\right) dt,$$

with a standart Brownian motion B. Pitman [34] also obtains this equation and discusses it in detail, therefore there is no need to trace this aspect further. The idea of using C as a time-change goes back to Lamperti [29].

The mentioned proofs all focus on the relationship to Brownian local time. In contrast we shall not rely on local times in this paper. (A short description of our proof in the finite variance case already appeared in the technical report [26].) The reason is that in the case $\alpha < 2$ the connection to local times breaks down. Then Y and consequently H exhibits a.s. jumps. Since local time processes of Lévy-processes (if existent) have a.s. continuous paths (compare f.e. [5], chapter V.1), they are no longer suited.

This can be explained on a heuristic level, too. Consider the following construction, going back to Harris [23] and used by different people. Traverse the individuals of T in the following manner: From individual i pass over to its oldest child, which has not yet been visited, resp. return to its predecessor, if all children of i have already been visited. This gives a traversal through T with the root as starting and end point. Each edge is passed twice, once forward and once backward. Next consider the associated random path, which increases one unit, if we change over to a child, and decreases one unit, if we go back to a predecessor. In the case of a geometric offspring distribution we get a true random walk excursion, conditioned to return to zero after $2n$ steps for a $CGW(n)$-tree. This is due to the lack of memory of the geometric distribution. The number of upcrossings from level $k-1$ to level k is equal to the size Z_k of the k'th generation, which makes the relation to local times obvious in this situation. In general the random path exhibits complicated dependence properties. In the case $\sigma^2 < \infty$ they are of a local nature and vanish in the limit $n \to \infty$, as was shown by Aldous, such that the asymptotic height profile can still be described by Brownian local time. For $\alpha < 2$ however, the dependence structure survives in the limit.

The combinatorics of $CGW(n)$-trees have been widely studied by means of generating functions (see f.e. [11, 17, 25, 31]). Probabilistic methods have been introduced in Kolchin [28] and in particular by Aldous [2, 3]. Our proof of theorem 1 is based on two probabilistic constructions, which are valid for the infinite variance case, too. The first one will be described in section 2, it establishes a connection between Galton-Watson trees and suitable random walk excursions. Though the relationship has been known for quite a while (compare [10]), the scope of this approach has been enlarged considerably only recently (see [4, 8, 20]). In our context it allows to reduce convergence of H^n to convergence of excursions. The required continuity theorem will be developed in section 3. The second probabilistic construction, which will be presented in section 4, is size-biasing of Galton-Watson trees. We use it to describe the bottom of the trees, which is of some interest of its own and will help to check a main condition of the continuity theorem. This concept goes back to Geiger [21], who developed a construction due to Lyons, Pemantle and Peres [30]. Section 5 addresses the question of convergence of excursions.

Thus the height profile will be considered as a functional of a random walk excursion S. Other quantities of the tree can be viewed as well as functionals of S. In this manner we can also treat the height profile of random forests, as discussed by Drmota and Gittenberger [12] and Pitman [34] in the finite variance case. We contend ourselves by stating a version of the theorem, which is valid for infinite variances, too. A *conditioned random forest* consists of l Galton-Watson trees, conditioned to contain altogether n individuals. Let now Z_k be the total number of all individuals in generation k in one of the l trees. Suppose that $l \sim \gamma a_n$, as $n \to \infty$, with $\gamma \geq 0$. Then the height profile $(H_u^n)_u = (a_n^{-1} Z_{[nu/a_n]})_u$ converges in distribution. The limiting process can be described as follows. Let $X_\gamma = (X_{\gamma,s})_{0 \leq s \leq 1}$ be a Lévy-process as above, now conditioned to hit $-\gamma$ at the moment $s = 1$ for the first time. It is built up from the excursions

$$Y_\eta = (Y_{\eta,s})_{s \leq L_\eta} = (X_s + \eta)_{T_\eta \leq s \leq T_{\eta+}}, \ 0 \leq \eta \leq \gamma,$$

with $T_\eta = \inf\{s : X_s = -\eta\}$, $T_{\eta+} = \inf\{s : X_s < -\eta\}$ and $L_\eta = T_{\eta+} - T_\eta$.

Then the limiting process is given by

$$\sum_{0 \le \eta \le \gamma} H_{\eta,u}, \ u \ge 0,$$

where $H_{\eta,u}$ is derived similarly as above from

$$C_{\eta,u} = \sup\{s \le L_\eta \ : \ \int_0^s \frac{dt}{Y_{\eta,t}} \le u\}.$$

The sum is a.s. finite for every $u > 0$, which reflects the fact, that also in the limit only finitely many trees contribute to the height profile. This result can be proved in much the same manner (and with only little additional effort), as we shall obtain Theorem 1 below. For $\gamma = 0$ we are back in the situation of Theorem 1.

2 Trees and Random Walk Excursions

It has been known for some time, that Galton-Watson trees can be imbedded into random walks (compare [8] and the references therein). This is implicit in Dwass' important paper [10]. Here we give a combinatorical treatment, which allows generalization. Let T be a tree of size n. Suppose that we label the individuals in T with the numbers $1, 2, \ldots, n$ such that the root gets label 1, the individuals in the first generation the labels $2, \ldots, Z_0 + Z_1$ (say from left to right), the individuals in the second generation the labels $Z_0 + Z_1 + 1, \ldots, Z_0 + Z_1 + Z_2$ and so forth. Using these labels we define a random path $S = (S(0), S(1), \ldots, S(n))$ recursively by

$$S(0) = 1, \ S(i) = S(i-1) + \xi_i - 1, \ i = 1, \ldots n,$$

where ξ_i denotes the offspring number of the individual with label i. We can imagine that the path arises as follows: To its i'th increment individual i contributes the downward step -1, whereas each of its children contributes one step $+1$ upwards. Then each individual is responsible for one upward and one downward step, except the root, which has no predecessor and thus contributes only a step downwards. Since $S(0) = 1$,

$$S(n) = 0.$$

Furthermore, individuals always have smaller labels than their offspring, therefore the upward step of an individual appears before its downward step. Clearly this implies

$$S(i) > 0 \text{ for all } i < n,$$

i.e. S is an *excursion* of length n. Conversely, given such an excursion S of length n, with $S(i) \ge S(i-1) - 1$, we can construct a tree, fitting to the excursion S. Namely, from S we read off $\xi_i = S(i) - S(i-1) + 1$, and the given labelling rule allows us to grow the tree from its root. Thus there is a one-to-one correspondence between trees of size n and excursion of length n.

As to the probabilistic aspect of the construction note, that for a Galton-Watson tree the ξ_i are independent random variables, such that S becomes an ordinary random walk excursion. Likewise S is a random walk excursion of duration n, if T is a $CGW(n)$-tree.

Remarks 1.) These considerations remain valid for other ways of labelling. One possibility is to label according to depth-first search, which has been exploited in [4, 20]. In general the following properties are required:
i) 1 is the label of the root.
ii) The individuals with labels $1, \ldots, i$ form a subtree for any $i < n$. In other words: Any individual has a smaller label than any of its children.
iii) Given the subtree with labels $1, \ldots, i$ and the numbers $\xi_1, \ldots, \xi_i$ there is a rule, which specifies, which child of the individuals $1, \ldots, i$ gets the label $i+1$.
2.) A random forest of l trees and n individuals can be described by a random walk path S with $S(0) = l$, $S(i) > 0$ for $i < n$ and $S(n) = 0$. If we label the trees one after the other, then the r'th tree is represented by the part of the random walk between the hitting times of $l + 1 - r$ and $l - r$.

In our labelling $1, \ldots, Z_0 + Z_1 + \ldots + Z_{k-1}$ are just the members of generation 0 to $k-1$. They contribute a negative step to $S(Z_0 + Z_1 + \ldots + Z_{k-1})$. Their children, i.e the individuals in generation 1 to k, add a positive step. Therefore

$$Z_k = S(Z_0 + Z_1 \ldots + Z_{k-1}).$$

This is the announced random walk representation for the height profile, which has been used by several authors (see [8]). We transform it into a differential equation. Define

$$C_u^n = \int_0^u a_n^{-1}\, Z_{[nv/a_n]}\, dv,$$

in particular

$$C_{ka_n/n}^n = \frac{1}{n}(Z_0 + \ldots + Z_{k-1}).$$

Further let for $0 \le s \le 1$

$$S^n(s) = a_n^{-1} S([ns])$$

and

$$Y^n(s) = S^n(C_{ka_n/n}^n) \text{ for } s \in [C_{ka_n/n}^n, C_{(k+1)a_n/n}^n).$$

Then the above equation translates into the ordinary differential equation

$$\frac{d}{du} C_u^n = Y^n(C_u^n),$$

which by integration leads to

$$u = \int_0^{C_u^n} \frac{dt}{Y^n(t)} \text{ for } u \leq h_n \text{ and } C_u^n = 1 \text{ for } u > h_n,$$

with

$$h_n = \int_0^1 \frac{dt}{Y^n(t)}.$$

h_n obviously is the height of the rescaled tree. — It is now our plan to reduce the question of convergence of H^n and C^n to that of S^n and Y^n. The next section provides the required continuity statement.

3 A Continuity Theorem

Let D and D' be the spaces of all cádlág-functions $f : [0,1] \to \mathbb{R}$ resp. $g : [0,\infty) \to \mathbb{R}$. By endowing them with the common Skorohod-distance (compare [14]) we make them to complete metric spaces. Thus functions f_n converge to f in D, if there are increasing bijections $\alpha_n : [0,1] \to [0,1]$, such that, as $n \to \infty$,

$$\sup_t |\alpha_n(t) - t| \to 0, \qquad \sup_t |f_n(t) - f(\alpha_n(t))| \to 0.$$

Similarly $g_n \to g$ in D', if there are increasing bijections $\beta_n : [0,\infty) \to [0,\infty)$, such that for all $v < \infty$

$$\sup_{u \leq v} |\beta_n(u) - u| \to 0, \qquad \sup_{u \leq v} |g_n(u) - g(\beta_n(u))| \to 0.$$

If g is a continuous function, we may choose $\beta_n(u) = u$.

We begin by collecting some analytical facts. Let D_+ be the space of non-negative $f \in D$.

Lemma 2 *Let $f, f_n \in D_+$.*
i) For given s the mapping $f \mapsto \int_0^s dt/f(t)$ (with the possible value ∞) is lower-semicontinuous (and thus measurable).
ii) If $\inf_{a \leq t \leq b} f(t) > 0$ for some $0 \leq a < b \leq 1$, and if $f_n \to f$, then

$$\sup_{a \leq s \leq b} \left| \int_a^s \frac{dt}{f(t)} - \int_a^s \frac{dt}{f_n(t)} \right| \to 0.$$

Proof Let α_n as above. As is well-known, it may be assumed without loss of generality, that they are differentiable functions such that

$$\sup_t |\alpha_n^{'}(t) - 1| \to 0.$$

i) If $f_n \to f$, then for $\delta > 0$ by Fatou's Lemma

$$\int_0^{s-\delta} \frac{dt}{f(t)} \le \liminf_n \int_0^{s-\delta} \frac{dt}{f_n(\alpha_n^{-1}(t))} \le \liminf_n \int_0^s \frac{\alpha_n^{'}(t)\,dt}{f_n(t)}$$

and consequently

$$\int_0^s \frac{dt}{f(t)} \le \liminf_n \int_0^s \frac{dt}{f_n(t)}.$$

ii) Due to uniform convergence of the integrands

$$\sup_{a \le s \le b} \left| \int_a^s \frac{dt}{f_n(t)} - \int_a^s \frac{dt}{f(\alpha_n(t))} \right| \to 0.$$

Further, substituting $\alpha_n(t) = w$

$$\int_a^s \frac{dt}{f(\alpha_n(t))} = \int_a^{\alpha_n(s)} \frac{dw}{\alpha_n'(t) f(w)} \to \int_a^s \frac{dw}{f(w)}$$

uniformly for all $s \in [a, b]$. □

Define now for $f \in D_+$ a function $g = \phi(f) \in D'$ by

$$g(u) = \sup\{s \le 1 \mid \int_0^s \frac{dt}{f(t)} \le u\}.$$

Thus as above

$$u = \int_0^{g(u)} \frac{dt}{f(t)} \text{ for } u < h, \text{ and } g(u) = 1 \text{ for } u \ge h,$$

with

$$h = \int_0^1 \frac{dt}{f(t)}.$$

g is continuous and increasing. It is everywhere differentiable from the right (since f is continuous from the right), and the derivative is given by

$$\frac{d^+}{du} g(u) = f(g(u)) \text{ for } u < h \text{ and } \frac{d^+}{du} g(u) = 0 \text{ for } u \ge h.$$

We denote

$$\psi(f) = \frac{d^+}{du} \phi(f).$$

Lemma 3 *Suppose* $f_n \to f$ *in* D_+, $\sup_u |\phi(f_n)(u) - \phi(f)(u)| \to 0$ *and* $\int_0^s dt/f(t) < \infty$ *for all* $s < 1$. *Then* $\psi(f_n) \to \psi(f)$ *in* D'.

Proof Denote $g = \phi(f)$, $g_n = \phi(f_n)$ and $h_n = \int_0^1 dt/f_n(t)$. Let $\alpha_n(t)$ be as above. The required bijections $\beta_n : \mathbb{R}_+ \to \mathbb{R}_+$ are defined as

$$\beta_n(u) = \begin{cases} g^{-1}(\alpha_n(g_n(u))) & \text{for } u \le h_n, \\ \beta_n(h_n) + (u - h_n) & \text{for } u > h_n. \end{cases}$$

Then

$$\begin{aligned} \sup_{u \le h_n} |g(\beta_n(u)) - g_n(u)| &= \sup_{u \le h_n} |\alpha_n(g_n(u)) - g_n(u)| \\ &\le \sup_t |\alpha_n(t) - t| \to 0. \end{aligned}$$

By assumption it follows

$$\sup_{u \le h_n} |g(\beta_n(u)) - g(u)| \to 0.$$

If $h = \int_0^1 dt/f(t) < \infty$, then $g^{-1}(s) = \int_0^s dt/f(t)$ is uniformly continuous on $[0,1]$, and it follows

$$\sup_u |\beta_n(u) - u| = \sup_{u \le h_n} |\beta_n(u) - u| \to 0.$$

If on the other hand $h = \infty$, then $h_n \to \infty$ and g^{-1} is uniformly continuous on every intervall $[0, s]$ with $s < 1$. In this case we may conclude

$$\sup_{u \le v} |\beta_n(u) - u| \to 0$$

for every $v > 0$. This is one of the desired properties.

Next for $u > h_n$ we have $1 = g_n(h_n) = g_n(u)$ and $1 = \alpha_n(1) = \alpha_n(g_n(h_n)) = g(\beta_n(h_n)) = g(\beta_n(u))$, therefore

$$\begin{aligned} &\sup_u |f_n(g_n(u)) - f(g(\beta_n(u)))| \\ &= \sup_{u \le h_n} |f_n(g_n(u)) - f(g(\beta_n(u)))| \\ &= \sup_{u \le h_n} |f_n(g_n(u)) - f(\alpha_n(g_n(u)))| \\ &= \sup_t |f_n(t) - f(\alpha_n(t))| \to 0. \end{aligned}$$

Since $\psi(f) = f \circ g$ and $\psi(f_n) = f_n \circ g_n$, also

$$\sup_u |\psi(f_n)(u) - \psi(f)(\beta_n(u))| \to 0,$$

which proves the claim. $\square$

We are now ready to prove the main result of the section. Let Y, Y^n be stochastic processes with paths in D_+, and define processes C, C^n, H and H^n by

$$C_u = \phi(Y)(u), \quad C_u^n = \phi(Y^n)(u),$$
$$H_u = \psi(Y)(u), \quad H_u^n = \psi(Y^n)(u).$$

Theorem 4 *Assume* $\inf_{\delta \le s \le 1-\delta} Y_s > 0$ *a.s. for all* $\delta > 0$ *and, as* $\epsilon \to 0$,

$$\limsup_n \mathrm{P}^n(C_u^n \le \epsilon) \to 0$$

for all $u > 0$*. Then* $C_u > 0$ *a.s. for all* $u > 0$*, i.e.* H *is non-degenerate, and* $C^n \to C$, $H^n \to H$ *in distribution, as* $n \to \infty$*.*

Proof Due to a wellknown theorem of Skorokhod (see [35], chapter IV.3, Theorem 13) we may assume that the processes Y_n and Y are defined on a single probability space $(\Omega, \mathcal{A}, \mathrm{P})$, and that $Y_n \to Y$ a.s. in D_+. Because of semicontinuity (Lemma 2 i)) and the definition of C it follows $C_u \ge \limsup_n C_u^n$ a.s.. From Fatou's Lemma

$$\mathrm{P}(C_u < \epsilon) \le \mathrm{P}(\limsup_n C_u^n < \epsilon) \le \limsup_n \mathrm{P}(C_u^n < \epsilon).$$

Thus our assumptions imply $C_u > 0$ a.s. for all $u > 0$.

Next let $C_u < 1$. Then $C_u^n < 1$ for large n, therefore $u = \int_0^{C_u} dt/Y_t = \int_0^{C_u^n} dt/Y_t^n$. It follows

$$\begin{aligned} |C_u^n - C_u| &\le \max Y^n \left| \int_0^{C_u^n} \frac{dt}{Y_t^n} - \int_0^{C_u} \frac{dt}{Y_t^n} \right| \\ &= \max Y^n \left| \int_0^{C_u} \frac{dt}{Y_t} - \int_0^{C_u} \frac{dt}{Y_t^n} \right|. \end{aligned}$$

Therefore for any $a < 1$ and b such that $C_b = a$

$$\begin{aligned} \sup_u |C_u^n - C_u| &\le \sup_{u \le b} |C_u^n - C_u| + (1 - C_b^n) + (1 - C_b) \\ &\le 2 \sup_{u \le b} |C_u^n - C_u| + 2(1-a) \\ &\le 2 \max Y^n \sup_{s \le a} \left| \int_0^s \frac{dt}{Y_t} - \int_0^s \frac{dt}{Y_t^n} \right| + 2(1-a). \end{aligned}$$

Now $\max Y^n \to \max Y$ a.s.. In view of Lemma 2 ii) it follows for every $\eta > 0$ there is a $v > 0$ (chose a sufficiently close to 1), such that for all $\epsilon > 0$

$$\begin{aligned} &\limsup_n \ \mathrm{P}(\sup_u |C_u^n - C_u| > \eta) \\ &\le \ \eta + \limsup_n \ \mathrm{P}(\left| \int_0^\epsilon \frac{dt}{Y_t} - \int_0^\epsilon \frac{dt}{Y_t^n} \right| > v) \\ &\le \ \eta + \mathrm{P}(C_v < \epsilon) + \limsup_n \ \mathrm{P}(C_v^n < \epsilon) \\ &\le \ \eta + 2 \ \limsup_n \ P(C_v^n < \epsilon). \end{aligned}$$

By assumption the righthand side can be made arbitrarily small, such that $\sup_u |C_u - C_u^n| \to 0$ in probability. In view of Lemma 3 $H^n \to H$ in probability, and the proof is finished. □

4 Size-biased Galton–Watson Trees

In this section we verify one of the conditions of Theorem 4 for height profiles of a $CGW(n)$-tree. We shall make use of size-biased trees, as they have been constructed by Geiger for binary branching trees resp. splitting trees (compare [21]). Let

$$\varphi(\lambda) = \sum_{x=0}^{\infty} \lambda^x p_x$$

be the generating function of the offspring distribution (p_x). Fix $\lambda \in (0, 1)$ and consider the probability weights

$$q_x = \lambda^x p_x / \varphi(\lambda).$$

Then we have for a Galton Watson tree T the formula

$$\mathrm{Q}(T = t) = \prod_i q_{d(i)} = \lambda^{s(t)-1} \varphi(\lambda)^{-s(t)} \prod_i p_{d(i)} = \lambda^{s(t)-1} \varphi(\lambda)^{-s(t)}\, \mathrm{P}(T = t),$$

where $d(i)$ denotes the number of children of individual i in the (nonrandom) tree t, and P and Q denote the probability measures, corresponding to the offspring distributions (p_x) resp. (q_x). It follows the known fact (compare [25]) that two $CGW(n)$-tree with offspring distributions (p_x) and (q_x) are equal in distribution. Passing over to (q_x) has the effect that the total size of T becomes finite in mean ([15], chapter XII.5):

$$\mathrm{E}_{\mathrm{Q}} s(T) = (1 - \mu)^{-1} < \infty,$$

where

$$\mu = \sum_x x q_x < 1.$$

Thus (differently from P) we may bias Q by introducing the probability measure

$$\widehat{\mathrm{Q}}(T = t) = (1 - \mu) s(t)\, \mathrm{Q}(T = t).$$

Note: Conditioning on size $s(T) = n$, it makes no difference, whether we consider Q or $\widehat{\mathrm{Q}}$. If moreover M is chosen purely at random from the vertices of T, i.e. from $\{1, 2, \ldots, s(T)\}$, then

$$\widehat{\mathrm{Q}}(T = t, M = m) = (1 - \mu)\mathrm{Q}(T = t).$$

Let $(\widehat{q}_x)$ be the distribution obtained from (q_x) by size-biasing, i.e.

$$\widehat{q}_x = xq_x/\mu,$$

then we may rewrite the above formula as

$$\widehat{\mathrm{Q}}(T = t, M = m) = (1-\mu)\mu^g \prod_{i<m} \widehat{q}_{d(i)} \frac{1}{d(i)} \prod_{i \not< m} q_{d(i)},$$

where g denotes the generation of individual m, and $i < m$ means that individual i is a predecessor of m. Following Geiger [21], this formula gives raise to a probabilistic construction for the size-biased Galton Watson tree:

Construction of the tree $\widehat{T}$

- Let G be a random variable with geometric distribution and parameter μ. Choose independent random variables $\widehat{\xi}_1, \ldots, \widehat{\xi}_G$ with distribution $(\widehat{q}_x)$. Let ζ_j be random numbers, taken independently and uniformly from $\{1, 2, \ldots, \widehat{\xi}_j\}$, $j = 1, \ldots, G$.

- The trunk of $\widehat{T}$: Build up a line of G consecutive individuals, where the j'th individual has $\widehat{\xi}_j$ children. Let the first one be the tree's root and the $(j+1)$'th one be the ζ_j'th child (in the order of birth) of the j'th individual. Give the label M to the ζ_G'th child of the G'th individual.

- The tree top of $\widehat{T}$: Besides these G individuals the trunk contains $y = \widehat{\xi}_1 + \ldots + \widehat{\xi}_G + 1 - G$ additional individuals. They propagate in the usual Galton-Watson manner. This means: In order to complete the tree $\widehat{T}$ we attach independent Galton-Watson trees $T_1, \ldots, T_y$ with offspring distribution (q_x) to the trunk.

In this manner we obviously obtain a size-biased tree together with an individual M, taken at random from the tree, and belonging to generation G:

$$\mathrm{Q}(\widehat{T} = t, M = m) = \widehat{\mathrm{Q}}(T = t, M = m).$$

Note that for this construction it is only required that (p_x) has mean 1. In the limiting case $\lambda = 1$ (or equivalently $\mu = 1$) G takes a.s. the value ∞, and $\widehat{q}_x$ is equal to

$$\widehat{p}_x = xp_x.$$

Then we get an *infinite size-biased tree* $\widetilde{T}$, which already appeared in the work of Grimmett [22], Kesten [27], Aldous [2] and others. We use this tree to describe the asymptotic shape of the lower part of a $CGW(n)$-tree. For any tree t let $t(k)$ be the tree, which results by cutting off all individuals in t belonging to a generation greater than k. Thus, $t(k)$ has in generation k' the generation sizes $z_{k'}$ for $k' \le k$, and 0 for $k' > k$. Further let $\mathcal{T}(k)$ be the set of trees with height at most k.

Theorem 5 *Let T be a Galton-Watson tree with an offspring distribution (p_x) fulfilling assumption A, and let (k_n) be a sequence of natural numbers such that $k_n = o(n/a_n)$. Then for $n \to \infty$*

$$\sup_{B \subset \mathcal{T}(k_n)} |\mathrm{P}(T(k_n) \in B \mid s(T) = n) - \mathrm{P}(\widetilde{T}(k_n) \in B)| \to 0.$$

In the special case of a Poisson offspring distribution this was proved by Aldous in [2]. We prepare the proof by gathering some facts of an analytical character. Let

$$v(x) = \sum_{y \le x} y(y-1)p_y.$$

In order that (p_x) belongs to the domain of attraction of a stable law with index α, it is necessary and sufficient, that $v(x)$ varies regularly at infinity with exponent $2 - \alpha$. Then a_n can be chosen as any sequence with the property

$$a_n^2/v(a_n) = cn(1 + o(1))$$

with some $c > 0$ (compare [16], chapter XVII.5). It follows that a_n varies regularly with exponent $1/\alpha$. We have a local limit law at our disposal, which in our case reads as follows.

Proposition 6 *Let $\xi_1, \xi_2, \ldots$ be i.i.d. random variables with common distribution (p_x), satisfying assumption A. Then uniformly in $x \in \mathbb{Z}$*

$$\mathrm{P}(\xi_1 + \ldots + \xi_n - n = x) = a_n^{-1} g(x/a_n)(1 + o(1)),$$

where g denotes the (continuous and strictly positive) density of the limit law ν.

The proof can be found in [18], chapter 9, up to an exceptional case, and in [33] in full generality.

Lemma 7 *Let $\widehat{\xi}_1, \widehat{\xi}_2, \ldots$ be independent random variables with distribution $(\widehat{p}_x)$. Then assumption A implies convergence of $a_n^{-1}(\widehat{\xi}_1 + \ldots + \widehat{\xi}_{[n/a_n]})$ in distribution.*

Proof If $v(\infty) = \sigma^2 < \infty$, then a_n is asymptotically proportional to $n^{1/2}$ and $(\widehat{p}_x)$ has finite mean, and the claim follows from the ordinary law of large numbers. The case $\alpha = 2, v(\infty) = \infty$ is similar. Then $v(x) = \sum_{y \le x}(y-1)\widehat{p}_y$ is a slowly varying function. It follows $x\mathrm{P}(\widehat{\xi} \ge x) = o(v(x))$, as $x \to \infty$ (compare [16], chapter VIII.9, Theorem 2). This allows to apply a generalized law of large numbers, as given f.e. in [13], chapter 1.5, exercise 5.11.

If $\alpha < 2$, then $x \sum_{y \geq x} \widehat{p}_y \sim (2-\alpha)(\alpha-1)^{-1} v(x)$, as follows from the cited theorem in [16]. Therefore $\mathrm{P}(\widehat{\xi} \geq x)$ is regular varying with exponent $1-\alpha$, such that $\widehat{\xi}$ belongs in distribution to the domain of attraction of a positive stable law. Moreover $\mathrm{P}(\widehat{\xi} \geq a_n) \sim c a_n / n$ for some $c > 0$, from which the claim follows by standard results on convergence in distribution to stable laws. $\square$

Lemma 8 *Let $T, T_1, T_2, \ldots$ be independent Galton-Watson trees with offspring distribution (p_k), satisfying assumption A. Then it follows:*
i) $\mathrm{P}(s(T) = n) = n^{-1} a_n^{-1} g(0)(1 + o(1))$, as $n \to \infty$.
ii) $n^{-1}(s(T_1) + \ldots + s(T_{[a_n]}))$ converges in distribution.

Proof i) is an immediate consequence of Proposition 6 and the following classical formula (see Dwass [10] and Kolchin [28], Lemma 2.1.3):

$$\mathrm{P}(s(T) = n) = \frac{1}{n} \mathrm{P}(\xi_1 + \ldots + \xi_n = n - 1).$$

Since a_n is regularly varying with exponent $1/\alpha$, it follows $\mathrm{P}(s(T) \geq n) \sim (1 - \alpha^{-1}) g(0) a_n^{-1}$. Consequently $s(T)$ belongs in distribution to the domain of attraction of a positive stable law, and ii) follows from standard results. $\square$

We come to the *Proof of Theorem 5.* Together with T we consider an individual M, chosen at random from T. We already pointed out that

$$\mathrm{P}(T = t, M = m \mid s(T) = n) = \widehat{\mathrm{Q}}(T = t, M = m \mid s(T) = n).$$

Therefore instead of T we may consider the size-biased tree $\widehat{T}$, as obtained in the above construction. Let $x_1, \ldots, x_k$ be offspring numbers of the lower k individuals in the trunk (including the root). Let the $(j+1)$'th individual in the trunk be the d_j'th child of the j'th individual, with $d_j \leq x_j$. Let $t_1, \ldots, t_z$, $z = x_1 + \ldots + x_k - k$ be the trees, growing out of these offspring to the right and left of the trunk. Because of the lack of memory of a geometric distribution

$$\begin{aligned}
&\mathrm{Q}(\widehat{\xi}_1 = x_1, \ldots, \widehat{\xi}_k = x_k, \\
&\qquad \zeta_1 = d_1, \ldots, \zeta_k = d_k, T_1 = t_1, \ldots, T_z = t_z, s(\widehat{T}) = n) \\
&= \quad \mathrm{Q}(G \geq k)\, \widehat{q}_{x_1} \ldots \widehat{q}_{x_k} x_1^{-1} \ldots x_k^{-1} \\
&\qquad \times \mathrm{Q}(T_1 = t_1) \ldots \mathrm{Q}(T_z = t_z)\, \mathrm{Q}(s(\widehat{T}) = n - k - s) \\
&= \quad \mu^k \widehat{q}_{x_1} \ldots \widehat{q}_{x_k} x_1^{-1} \ldots x_k^{-1}\, \mathrm{Q}(T_1 = t_1) \ldots \mathrm{Q}(T_z = t_z) \\
&\qquad \times (1 - \mu)(n - k - s) \mathrm{Q}(s(T) = n - k - s),
\end{aligned}$$

with $s = s(t_1) + \ldots + s(t_z)$. As to the dependence on λ it is easy to check, that the probabilities on the righthand side contain altogether the factor $\mu^{-k}\lambda^{n-1}\varphi(\lambda)^n$. Therefore there are numbers $c(n,\lambda)$ such that

$$\begin{aligned} & \mathrm{Q}(\widehat{\xi}_1 = x_1, \ldots, \widehat{\xi}_k = x_k, \\ & \qquad \zeta_1 = d_1, \ldots, \zeta_k = d_k, T_1 = t_1, \ldots, T_z = t_z, s(\widehat{T}) = n) \\ &= \quad c(n,\lambda)\, \widehat{p}_{x_1} \ldots \widehat{p}_{x_k} x_1^{-1} \ldots x_k^{-1}\, \mathrm{P}(T_1 = t_1) \ldots \mathrm{P}(T_z = t_z) \\ & \qquad \times (n-k-s)\mathrm{P}(s(T) = n-k-s), \end{aligned}$$

in particular

$$\mathrm{Q}(s(\widehat{T}) = n) = c(n,\lambda)\, n\, \mathrm{P}(s(T) = n).$$

From Lemma 8 i) we get

$$\begin{aligned} & \mathrm{P}(\widehat{\xi}_1 = x_1, \ldots, \widehat{\xi}_k = x_k, \\ & \qquad \zeta_1 = d_1, \ldots, \zeta_k = d_k, T_1 = t_1, \ldots, T_z = t_z \mid s(T) = n) \\ &\sim \quad \widehat{p}_{x_1} \ldots \widehat{p}_{x_k} x_1^{-1} \ldots x_k^{-1}\, \mathrm{P}(T_1 = t_1) \ldots \mathrm{P}(T_z = t_z), \end{aligned}$$

as long as $k + s(t_1) + \ldots + s(t_z) = o(n)$. Now it is assumed that $k = k_n = o(n/a_n)$. From Lemma 7 it follows, that in probability $z = \widehat{\xi}_1 + \ldots + \widehat{\xi}_k - k_n = o(a_n)$. Therefore in view of Lemma 8 $s(T_1) + \ldots + s(T_z) = o(n)$. Thus our claim follows. $\square$

We use now Theorem 5 to verify the condition of Theorem 4 for a height process.

Lemma 9 *Let C^n be the cumulative height process of a CGW(n)-tree, as defined in section 2. If assumption A is satisfied, then as $\epsilon \to 0$,*

$$\limsup_n \mathrm{P}(C_u^n \leq \epsilon) \to 0$$

for all $u > 0$.

For the proof we need an estimate on the height of a Galton-Watson tree.

Lemma 10 *Let $h(T) = \max\{k \mid Z_k > 0\}$ be the height of a Galton-Watson tree. Then assumption A implies $\mathrm{P}(h(T) > n/a_n) = c'\, a_n^{-1}(1 + o(1))$ for some $c' > 0$.*

Proof Since $v(x)$ is regularly varying, by Karamata's Tauberian Theorem (compare [16], chapter XIII.5) $\varphi''(\lambda)$ varies regularly at 1- with exponent $\alpha - 2$, more precisely

$$\varphi''(\lambda) = \sum_x x(x-1)\lambda^{x-2} p_x \sim \Gamma(3-\alpha)\, v((1-\lambda)^{-1}).$$

It follows that $\varphi'(\lambda) - 1$ and $\varphi(\lambda) - \lambda$ vary regularly at 1-, too (compare [16], chapter VIII.9), and

$$\varphi(\lambda) - \lambda \sim \frac{1}{\alpha(\alpha-1)}(1-\lambda)^2 \varphi''(\lambda).$$

This allows us to apply Lemma 2 from Slack [37], which in our notation says

$$\frac{\varphi(\mathrm{P}(h(T) \leq n)) - \mathrm{P}(h(T) \leq n)}{\mathrm{P}(h(T) > n)} \sim \frac{1}{(\alpha - 1)n}.$$

Combining these estimates and replacing n by n/a_n we get

$$\mathrm{P}(h(T) > n/a_n)\, v(\mathrm{P}(h(T) > n/a_n)^{-1}) \sim \frac{\alpha\, a_n}{\Gamma(3-\alpha)\, n}.$$

Comparing this with $v(a_n)/a_n \sim ca_n/n$, our claim follows. $\square$

Proof of Lemma 9 We have to show that there are sufficiently many individuals in the bottom of the trees under consideration. This will be done first for the infinite size-biased tree $\widetilde{T}$. Let $T_1, \ldots, T_z$ be those Galton-Watson trees, which grow out of the trunk of $\widetilde{T}$ at an individual, belonging to a generation less than n/a_n. Then, given $\eta > 0$, in view of Lemma 7 there are numbers c_1, c_2, such that $c_1 a_n \leq z \leq c_2 a_n$ with probability at least $1 - \eta/3$. In view of Lemma 10 the number of trees $T_i, i \leq c_2 a_n$ with $h(T_i) > n/a_n$ has asymptotically a Poisson distribution. Therefore there is a number l, such that

$$\limsup_n \mathrm{P}(h(T_i) > n/a_n \text{ for at least } l \text{ of the } i \leq z) \leq \eta/2.$$

On the other hand because of Lemma 8 the number of trees $T_i, i \leq c_1 a_n$ such that $s(T_i) > \delta n$ is asymptotically Poisson distributed, with a parameter going to ∞, as δ goes to 0. Therefore there is a $\delta > 0$ such that

$$\limsup_n \mathrm{P}(s(T_i) > \delta n \text{ for less than } l \text{ of the } i \leq z) \leq \eta/2.$$

Altogether we conclude: For any $\eta > 0$ there is a $\delta > 0$, such that with probability at least $1 - \eta$ there is a tree $T_i, i \leq z$, such that $h(T_i) \leq n/a_n$ and $s(T_i) > \delta n$. Since T_i stems from an individual in a generation less than n/a_n, we see, that for large n the number of individuals in the first $2n/a_n$ generations of $\widetilde{T}$ is bigger than δn with probability at least $1 - \eta$.

Now let $\epsilon_n > 0$ be a sequence, converging to 0. Choose $\delta > 0$ and define $m = m(n)$ by $\epsilon_n n = \delta m$. Then from the definition of C^n in section 2 for $u > 0$, if n is large enough,

$$\begin{aligned} \mathrm{P}(C_u^n \leq \epsilon_n) &= \mathrm{P}(Z_0 + Z_1 + \ldots + Z_{[nu/a_n]} \leq \delta m \mid s(T) = n) \\ &\leq \mathrm{P}(Z_0 + Z_1 + \ldots + Z_{[2m/a_m]} \leq \delta m \mid s(T) = n). \end{aligned}$$

Since $m = o(n)$, we may in view of Theorem 5 switch over from T to $\widetilde{T}$, therefore

$$\limsup_n \mathrm{P}(C_u^n \leq \epsilon_n) \leq \eta$$

for all $\eta > 0$. This holds for any sequence ϵ_n, thus, as is not difficult to see, the claim of the lemma follows. $\square$

5 Convergence of Excursions

In this section we prove convergence in distribution of the rescaled random walk S, conditioned on the event $S(n) = 0, S(i) > 0$ for $i < n$. The usual way to define a random walk excursion is, to condition the random walk on the event $S(n) \leq 0, S(i) > 0$ for $i < n$. Since we deal with a random walk, skipfree to the left (i.e. steps to the left cannot be bigger than 1), this makes no difference. Convergence of normalized random walk excursions seems to be studied only in the finite variance case, see Kaigh [24]. In this section we derive the required generalization. Let as in section 2

$$S^n(s) = a_n^{-1} S([ns]),\ 0 \leq s \leq 1.$$

Theorem 11 *Under assumption A the processes S^n, conditioned on the event $S(n) = 0, S(i) > 0$ for $i < n$, converge in distribution to the normalized excursion Y of the corresponding Lévy-process. It holds $Y_0 = Y_1 = 0$ and $\inf_{\delta \leq s \leq 1-\delta} Y_s > 0$ a.s. for all $\delta > 0$.*

Before proving this result let us first complete the proof of our main theorem.

Proof of Theorem 1 Recall the notation introduced in section 2. In view of Theorem 4 and Lemma 9 it remains to show that also the processes Y^n converge in distribution. From the random walk representation of section 2 it follows

$$\sup_k |C^n_{(k+1)a_n/n} - C^n_{ka_n/n}| = \frac{1}{n} \sup_k Z_k \leq \frac{a_n}{n} \sup_t S^n(t).$$

Because of Theorem 11 the righthand supremum converges in distribution. Since $a_n = o(n)$, we obtain

$$\sup_k |t^n_{k+1} - t^n_k| = o(1)$$

in probability, with $t^n_k = C^n_{ka_n/n}$. Without loss we may again assume that this convergence as well as the convergence of S^n to Y takes place in the a.s. sense. This means, that there are a.s. functions $\alpha_n : [0,1] \to [0,1]$ such that

$$\sup_t |S^n(t) - Y(\alpha_n(t))| \to 0, \ \sup_t |\alpha_n(t) - t| \to 0.$$

The role of the α_n is, as is well-known, to match the jump points $s_1, s_2, \dots$ of Y to those jumps of S^n, which asymptotically are not negligible. These are given by $\alpha_n(s_j)$. Define $\beta_n(s_j) = t^n_{k+1}$, if $\alpha_n(s_j) \in [t^n_k, t^n_{k+1})$ and $j \leq d_n$. If d_n is going sufficiently slowly to ∞, we obtain by linear interpolation bijections $\beta_n : [0,1] \to [0,1]$, which match the jump points of Y to those of Y^n. This implies, as is not difficult to see, that a.s.

$$\sup_t |Y^n(t) - Y(\beta_n(t))| \to 0, \ \sup_t |\beta_n(t) - t| \to 0.$$

Thus $Y^n \to Y$ a.s. in the Skorohod sense, and our claim follows. □

The plan of the proof of Theorem 11 is, to reduce the theorem to convergence of random walks and random walk bridges. The first step, namely to generalize Donsker's theorem to the infinite variance situation, is fairly obvious. It is probabely known, though we could not find it in the literature (compare however Bloznelis [7] and the papers cited therein).

Proposition 12 *Under assumption A the unconditioned processes S^n converge in distribution to the Lévy-process X fulfilling $X_1 = \nu$ in distribution.*

Proof Since S^n and X have independent, stationary increments, convergence of the finite-dimensional distributions follows immediately from assumption A. Further $|S^n(\tau_n + \theta_n) - S^n(\tau_n)| \stackrel{d}{=} |S^n(\theta_n)| \to 0$ in probability for any sequence of positive numbers $\theta_n > 0$, going to 0, and any sequence τ_n of stopping-times, bounded uniformly from above. Now tightness follows from a criterion due to Aldous [1]. □

Next we discuss convergence of random walk bridges. For the construction of Lévy-bridges we refer the reader to [5], chapter VIII.3. Random walk bridges are treated in the next proof in quite a similar spirit.

Proposition 13 *Under assumption A the processes S^n, conditioned on the events $S(n) = 0$, converge in distribution to the Lévy-process X, conditioned on the event $X = 0$.*

Proof Fix $t \in (0,1)$ and let $\kappa : D \to \mathbb{R}$ be a continuous functional, such that $\kappa(f), f \in D$ does only depend on the values of $f(s), s \leq t$. Then by the Markov property

$$\mathrm{E}(\kappa(S^n) \, | S(n) = 0) = \mathrm{E}(\kappa(S^n) h_n(S([tn])))$$

with $h_n(x) = \mathrm{P}(S(n) - S([tn]) = -x)/\mathrm{P}(S(n) = 0)$. By Proposition 6

$$\begin{aligned} h_n(x) &= \frac{g(-x/a_{n-[tn]})}{a_{n-[tn]}} \Big/ \frac{g(0)}{a_n}(1+o(1)) \\ &= (1-t)^{-1/\alpha} g(-x/a_{n-[tn]})/g(0)\,(1+o(1)) \end{aligned}$$

uniformly in x. Therefore

$$\mathrm{E}(\kappa(S^n)\,|S(n)=0) = \frac{\mathrm{E}(\kappa(S^n)g(-(1-t)^{-1/\alpha}S^n(t)))}{(1-t)^{1/\alpha}g(0)}(1+o(1)).$$

Proposition 12 implies

$$\mathrm{E}(\kappa(S^n)\,|S(n)=0) \to \frac{\mathrm{E}(\kappa(X)g(-(1-t)^{-1/\alpha}X(t)))}{(1-t)^{1/\alpha}g(0)}.$$

This proves convergence in distribution of the process $(S^n(s))_{s\leq t}$, conditioned on the event $S(n) = 0$, for any $0 < t < 1$. Finally we have the duality relation $(S^n(1-s))_s \stackrel{d}{=} (S^n(0) - S^n(s-))_s$, thus also convergence of the processes $(S^n(s))_{s\geq 1-t}$ follows for every $0 < t < 1$. Combining these results our claim follows. $\square$

The proof also shows that for any $t < 1$ the distribution of the process $(X_s)_{s\leq t}$, given $X_1 = 0$, is absolute continuous with respect to the distribution of the unconditioned process $(X_s)_{s\leq t}$ (which follows as well from the construction in [5]). This allows to transfer properties. We need the following one: The Lévy-process X attains on $[0,1]$ its minimal value a.s. at exactly one point, namely

$$T = \inf\{s \leq 1 \mid X_t \geq \min(X_s, X_{s-}) \text{ for all } t \geq s\}.$$

Also $T < 1$ a.s., and X is a.s. continuous at T. This follows from Propositions 2.1 to 2.4 in Millar [32] (for our process 0 is regular for both $(-\infty, 0)$ and $(0,\infty)$). By absolute continuity this carries over to the process $(X_s)_{s\leq 1}$, conditioned on $X_1 = 0$.

There is an easy recipe using cyclic permutation of a path, which allows to pass over from a bridge $1 = S(0), S(1), \ldots, S(n-1), S(n) = 0$ to an excursion $1 = S(0), S(1) > 0, \ldots, S(n-1) > 0, S(n) = 0$, and which has been utilized by different people (see f.e. [38]). Let

$$T^n = \min\{i \leq n \mid S(j) \geq S(i) \text{ for all } j = i, i+1, \ldots, n\}$$

be the moment, when S takes its first minimal value before time n. Then $\overline{S}$, given by

$$\overline{S}(i) = \begin{cases} S(T^n + i) - S(T^n) + 1, & \text{if } i \leq n - T^n \\ S(T^n + i - n) - S(T^n), & \text{if } i \geq n - T^n, \end{cases}$$

is an excursion for any bridge S. In this manner we associate to each excursion $n+1$ different bridges, one being the given excursion itself.

This motivates to associate for any $f \in D$ an $\overline{f} \in D$, given by

$$\overline{f}(t) = \begin{cases} f(t+T(f)) - f(T(f)) + f(0), & \text{if } t \leq 1-T(f) \\ f(t+T(f)-1) - f(T(f)) + f(1), & \text{if } t \geq 1-T(f), \end{cases}$$

where

$$T(f) = \min\{s \leq 1 \mid f(t) \geq \min(f(s), f(s-)) \text{ for all } t \geq s\}.$$

Thus $\overline{f}(0) = f(0)$, $\overline{f}(1) = f(1)$.

Lemma 14 *Suppose, that $f \in D$ attains its minimum at no other point than $T(f)$. Then $f_n \to f$ implies $\overline{f}_n \to \overline{f}$ in the space D.*

Proof $f_n \to f$ means that there are bijections α_n, such that

$$\sup_t |f_n(t) - f(\alpha_n(t))| \to 0, \ \sup_t |\alpha_n(t) - t| \to 0.$$

From the uniqueness of the minimum of f it follows

$$T(f_n) \to T(f).$$

Now let $\eta_n > 0$ be numbers going to zero and define

$$\beta_n(t) = \begin{cases} \alpha_n(t+T(f_n)) - T(f), & \text{if } \eta_n \leq t \leq 1-T(f_n) \\ \alpha_n(t+T(f_n)-1) + 1 - T(f), & \text{if } 1-T(f_n) \leq t \leq 1-\eta_n. \end{cases}$$

Also put $\beta_n(0) = 0$, $\beta_n(1) = 1$ and continue β_n by linear interpolation on the whole interval $[0,1]$. If η_n goes to zero slowly enough, β_n is a bijection of $[0,1]$, and

$$\sup_t |\beta_n(t) - t| \to 0.$$

Further

$$\begin{aligned} &\sup_t |\overline{f}_n(t) - \overline{f}(\beta_n(t))| \\ &\quad \leq 2\sup_t |f_n(t) - f(\alpha_n(t))| \\ &\qquad + 2 \sup_{T(f_n)-\eta_n \leq s,t \leq T(f_n)+\eta_n} |f_n(s) - f(\alpha_n(t))| \\ &\quad \leq 4\sup_t |f_n(t) - f(\alpha_n(t))| \\ &\qquad + 2 \sup_{T(f_n)-\eta_n \leq s,t \leq T(f_n)+\eta_n} |f(\alpha_n(s)) - f(\alpha_n(t))|. \end{aligned}$$

If $T(f)$ is a point of continuity of f, then the righthand terms all go to zero, and it follows the claim $\overline{f}_n \to \overline{f}$. The case that f has a jump at $T(f)$ (which will not be considered in the sequel) is treated similarly; then $T(f_n) = T(f)$, for n large enough. □

Proof of Theorem 11 Proposition 13 shows that the rescaled random walk bridges S^n converge in distribution to a Lévy-bridge X. Due to Skorohod's theorem we may assume without loss of generality, that $S^n \to X$ a.s. in the Skorohod topology. Since X attains its minimum a.s. at a unique point, which is a point of continuity, it follows from Lemma 14 $\overline{S^n} \to \overline{X}$ a.s.. By construction $Y = \overline{X}$ has the stated property. □

This method of deriving a the normalized excursion Y from the bridge X, is well-known for Brownian motion, and in the general case discussed in detail in Chaumont [9].

References

[1] Aldous, D.: Stopping times and tightness. *Ann. Probab.* **6**, 335-340 (1978)

[2] Aldous, D.: The continuum random tree I. *Ann. Probab.* **19**, 1-28 (1991)

[3] Aldous, D.: The continuum random tree II: An overview. *Stochastic Analysis*, M. Barlow, N. Bingham (Eds.), 23-70 (1991)

[4] Bennies, J. and Kersting G.: A random walk approach to Galton-Watson trees. Preprint (1997)

[5] Bertoin, J.: *Lévy Processes.* Cambridge University Press, Cambridge 1996

[6] Biane, P.: Relations entre pont et excursion du mouvement Brownien réel. *Ann. Inst. Henri Poincaré* **22**, 1-7 (1986)

[7] Bloznelis, M.: Central limit theorem for stochastically continuous processes. Convergence to Stable Limit. *J. Theor. Probab.* **9** 541-560 (1996)

[8] Borovkov, K., Vatutin, V.: On distribution tails and expectations of maxima in critical branching processes. *J. Appl. Prob.* **33**, 614-622 (1996)

[9] Chaumont, L.: Excursion normalisée, méandre et pont pour des processus de Lévy stables. *Bull. Sci. Math.* **121**, 377-403 (1997)

[10] Dwass, M.: The total progeny in a branching process and a related random walk. *J. Appl. Prob.* **6**, 682-686 (1969)

[11] Drmota, M. and Gittenberger, B.: On the profile of random trees. *Random Struct. Alg.* **10**, 421-451 (1997)

[12] Drmota, M. and Gittenberger, B.: On the strata of random mappings - a combinatorical approach. Preprint (1997)

[13] Durrett, R.: *Probability: Theory and Examples.* Wadsworth, Belmont 1991

[14] Ethier, S. and Kurtz, T.: *Markov Processes, Characterisation and Convergence.* Wiley, New York 1986

[15] Feller, W.: *An Introduction to Probability Theory and its Applications.* Vol. 1. Wiley, New York 1950

[16] Feller, W.: *An Introduction to Probability Theory and its Applications.* Vol. 2. Wiley, New York 1971

[17] Flajolet, P., Odlyzko, A.: The average height of binary trees and other simple trees. *J. Comput. System Sci.* **25**, 171-213 (1982)

[18] Gnedenko, B.V. and Kolmogorov, A.N.: *Limit Distributions for Sums of Independent Random Variables.* Addison-Wesley, Cambridge 1954.

[19] Jeulin, T.: Application de la théorie du grossissement de filtrations a l'étude des temps locaux du mouvement Brownienne. *Lecture Notes in Math.* **1118**, 197-305.

[20] LeGall J.F., LeYan Y.: Branching processes in Lévy-processes: The exploration process. *Ann. Probab.* **26**, 213-252 (1998)

[21] Geiger, J.: Size-biased and conditioned random splitting trees. *Stoch. Proc. Appl.* **65**, 187-207 (1996)

[22] Grimmett, G.R.: Random labeled trees and their branching network. *J. Austral. Math. Soc. Ser. A* **30**, 229-237 (1980)

[23] Harris, T.E.: First passage and recurrence times. *Trans. Amer. Math. Soc.* **73**, 471-486 (1952)

[24] Kaigh, W.D.: An invariance principle for random walk conditioned by a late return to zero. *Ann. Probab.* **4**, 115-121 (1976)

[25] Kennedy, D.P.: The Galton-Watson-process, conditioned on the total progeny. *J. Appl. Prob.* **12**, 800-806 (1975)

[26] Kersting, G.: On the profile of a conditioned Galton-Watson process. Manuscript (1996)

[27] Kesten, H.: Subdiffusive behavior of random walk on a random cluster. *Ann. Inst. H. Poincaré Sect. B* **22**, 425-487 (1987)

[28] Kolchin, V.K.: *Random Mappings.* Optimization Software, New York 1986

[29] Lamperti, P.: Continuous-state branching processes. *Bull. Amer. Math. Soc.* **73**, 382-386 (1967)

[30] Lyons, R., Pemantle R., Peres Y.: Conceptual proofs of $L \log L$ criteria for mean behaviour of branching processes. *Ann. Probab.* **23**, 1125-1138 (1995)

[31] Meir, A., Moon, J.W.: On the altitude of nodes in random trees. *Canad. J. Math.* **30**, 997-1015 (1978)

[32] Millar, P.W.: Zero-one laws and the minimum of a Markov process. *Trans. Amer. Math. Soc.* **226**, 365-391 (1977)

[33] Mukhin, A.B.: On some necessary and sufficient conditions of validity of local limit theorems. *Dokl. AN Uzbek. SSR, Ser. Fiz.-Mat.* **2**, 7-8 (1986) (in Russian)

[34] Pitman, J.: The SDE solved by local times of a Brownian excursion or bridge derived from the height profile of a random tree or forest. Technical Report No. 503, Berkeley (1997)

[35] Pollard, D.: *Convergence of Stochastic Processes.* Springer, Berlin 1984

[36] Rogers, L.C.G., Williams, D.: *Diffusions, Markov Processes, and Martingales.* Wiley, Chichester 1987

[37] Slack, R.S.: A branching process with mean one and possibly infinite variance. *Z.Wahrsch.verw.Geb.* **9**, 139-145 (1968).

[38] Takács, L.: *Combinatorial Methods in the Theory of Stochastic Processes.* Kreiger Publ., New York 1977

[39] Takács, L.: Limit distributions for queues and random rooted trees. *J. Appl. Math. Stoch. Anal.* **6**, 189-216 (1993)