The Annals of Applied Probability
1997, Vol. 7, No. 4, 972–995

SOME WIDTH FUNCTION ASYMPTOTICS FOR WEIGHTED TREES

By Mina Ossiander, Ed Waymire and Qing Zhang

Oregon State University

Consider a rooted labelled tree graph τ_n having a total of n vertices. The *width function* counts the number of vertices as a function of the distance to the root ϕ. In this paper we compute large n asymptotic behavior of the width functions for two classes of tree graphs (both random and deterministic) of the following types: (i) Galton–Watson random trees τ_n conditioned on total progeny and (ii) a class of deterministic self-similar trees which include an "expected" Galton-Watson tree in a sense to be made precise. The main results include: (i) an extension of Aldous's theorem on "search-depth" approximations by Brownian excursion to the case of weighted Galton–Watson trees; (ii) a probabilistic derivation which generalizes previous results by Troutman and Karlinger on the asymptotic behavior of the expected width function and provides the fluctuation law; and (iii) width function asymptotics for a class of deterministic self-similar trees of interest in the study of river network data.

1. Introduction and statements of results. Let $\mathbb{T}$ be the space of *labelled tree graphs rooted* at ϕ. An element τ of $\mathbb{T}$ may be coded as a *set* of finite sequences of positive integers $\langle i_1, i_2, \dots, i_n\rangle \in \tau$ such that the following hold:

(i) $\phi \in \tau$ is coded as the empty sequence;

(ii) if $\langle i_1, \dots, i_k\rangle \in \tau$, then $\langle i_1, \dots, i_j\rangle \in \tau \ \forall\ 1 \le j \le k$;

(iii) if $\langle i_1, i_2, \dots, i_n\rangle \in \tau$, then $\langle i_1, \dots, i_{n-1}, j\rangle \in \tau \ \forall\ 1 \le j \le i_n$.

If $\langle i_1, \dots, i_n\rangle \in \tau$, then $\langle i_1, \dots, i_{n-1}\rangle \in \tau$ is referred to as the *parent* vertex to $\langle i_1, \dots, i_n\rangle$. A pair of vertices are connected by an edge (adjacent) if and only if one of them is *parent* to the other. In this way edges may also be identified with the (unique) nonparental or *descendant* vertex. This specifies the (planar) graph structure of τ and makes τ a rooted connected graph without cycles. The space $\mathbb{T}$ may be viewed as a complete metric space with the ultrametric

$$\rho(\tau, \gamma) = \frac{1}{\sup\{n\colon \gamma|n-1 = \tau|n-1\}},$$

and $\tau|n = \{\langle i_1, \dots, i_k\rangle \in \tau\colon k \le n\}$. This and the countable dense subset $\mathbb{T}_0$ of finite labelled tree graphs rooted at ϕ make $\mathbb{T}$ a Polish space. An important class of probability distributions on the Borel sigma field of $\mathbb{T}$ for this paper is the Galton–Watson distribution with single progenitor and offspring

Received April 1996; revised May 1997.

AMS 1991 *subject classifications*. 60J80, 60J85, 60J55, 60J70.

Key words and phrases. Branching process, width function, Brownian excursion, local time, occupation time, self-similar tree.

distribution p_k, $k = 0, 1, \ldots$, for which the probability assigned to a ball $B(\tau, 1/(N+1))$, $\tau \in \mathbb{T}_0$, $N \in \{1, 2, \ldots\}$ is

$$P\left(B\left(\tau, \frac{1}{N+1}\right)\right) = \prod_{v \in \tau|(N-1)} p_{l(v)},$$

where $l(v) = \#\{j: \langle v, j\rangle \in \tau|N\}$. The *height of a vertex* $v = \langle i_1, i_2, \ldots, i_n\rangle \in \tau$ is $h(v) = n+1$. The *height of a tree* τ is

$$h(\tau) = \sup\{h(v): v \in \tau\}. \tag{1.1}$$

Let $\nu = \nu(\tau)$ be the total number of vertices in τ. Let $\langle e\rangle$ be an edge (or link) of τ. We associate with each edge $\langle e\rangle$ of the tree τ a *random weight* $W(e)$, where $\{W(e): e = \langle i_1, \ldots, i_n\rangle,\ n \geq 1,\ i_n \geq 1\}$ is a (denumerable) collection of iid positive random variables, independent of τ and, without loss of generality, assumed to have mean 1. We let $Z_n = Z_n(\tau)$ be the number of vertices in τ at height $n+1$. Then $Z_0(\tau) = 1$ counts only the root. Vertices at level 1 are labeled as $\langle i_1\rangle$, $1 \leq i_1 \leq Z_1$, where $Z_1 = Z_1(\tau)$ is the total number of vertices in τ at height 2. We arrange the order of these vertices from left to right; that is, $\langle 1\rangle$ is on the far left, next (if it exists) is $\langle 2\rangle$ and so on, and $\langle 0, Z_1\rangle$ is on the far right. Recursively, the label $\langle i_1, i_2, \ldots, i_{k+1}\rangle$ with $i_j \geq 1$ and $1 \leq j \leq k+1$ is assigned to one of the vertices adjacent to $\langle i_1, i_2, \ldots, i_k\rangle$. In general, a label of the form $\langle i_1, i_2, \ldots, i_k\rangle$ is a vertex of τ which is connected to the root ϕ by a *self-avoiding path* in τ of exactly k edges. For convenience, one may add an edge (stem) to the root ϕ so that root can be regarded as connected to a ghost vertex by a stem. This also increases the vertex degree of the root. Figure 1 is an example of a labeled tree as described above.

The *weighted height of vertex* $\langle i_1, i_2, \ldots, i_k\rangle$ is defined as

$$H_W(i_1, i_2, \ldots, i_k) = \sum_{j=0}^{k} W(i_1, i_2, \ldots, i_j). \tag{1.2}$$

The *weighted tree height* of τ is defined as

$$H_W(\tau) = \sup_{\langle e\rangle \in \tau} H_W(e). \tag{1.3}$$

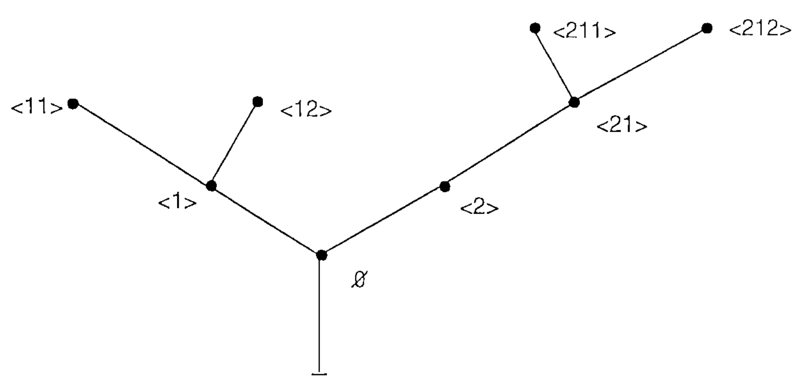

FIG. 1. *A rooted labelled tree.*

The *contour size* of the weighted tree τ at height h, also called the (*weighted*) *width function* of τ, is defined as

$$Z(h) = \begin{cases} 1, & \text{for } 0 \le h < W_\phi, \\ Z(h,\tau) = \#\{\langle e\rangle \in \tau\colon H_W(e') \le h < H_W(e)\}, & \text{else,} \end{cases} \tag{1.4}$$

where $\langle e'\rangle$ is the parent of $\langle e\rangle$ and # denotes the cardinality of the set. In the special case in which all $W(e) = W(i_1, i_2, \ldots, i_k) = 1$, the width function is given by

$$Z(h) = Z_k, \qquad k \le h < k+1,\ k = 0, 1, \ldots, \tag{1.5}$$

where Z_k is the total number of vertices at height $k+1$. This is the *search-depth* local time process in computer science applications (see [2]). Hydrologists interpret the weights as geomorphologic parameters such as lengths, elevation drops and so on, associated with river networks (see also [24]). Under the conditions on the weights to be imposed in this paper the width function and the search-depth local time will be (suitably scaled) asymptotically equivalent in distribution.

Troutman and Karlinger [23] considered the expected width function in the case of Galton–Watson random trees τ_n conditioned on total size n. In the hydrologic context one imagines a rain of particles uniformly distributed over τ_n and traveling at constant rate v. Then the (hydrograph) proportion of particles which reach the outlet (ϕ) in time t is represented by $Z(vt)/n$. The expected value is the best prediction in the mean square sense given the size of the network. This simple idealization illustrates some basic ideas in prediction problems based on maps of river basins (in place of stream gauge networks). Graphical data for a familiar sample river basin in Kentucky extracted from 30-m resolution Digital Elevation Maps using RiverTools software developed by Scott Peckham is given in Figure 2.

The width function for the Kentucky river network depicted in Figure 2 is given in Figure 3.

Let $\mu_n(h)$ be the conditional expectation of the width function evaluated at h given total progeny $\nu = n$; that is,

$$\mu_n(h) = E[Z(h)|\nu = n]. \tag{1.6}$$

Let K_n be the normalization constant defined by

$$K_n = \int_0^\infty \mu_n(h)\, dh, \tag{1.7}$$

and define a probability measure F_n with density $K_n^{-1}\mu_n$, suitably scaled. That is,

$$\frac{d}{dh} F_n(h) = a_n K_n^{-1} \mu_n(a_n h), \qquad h \ge 0, \tag{1.8}$$

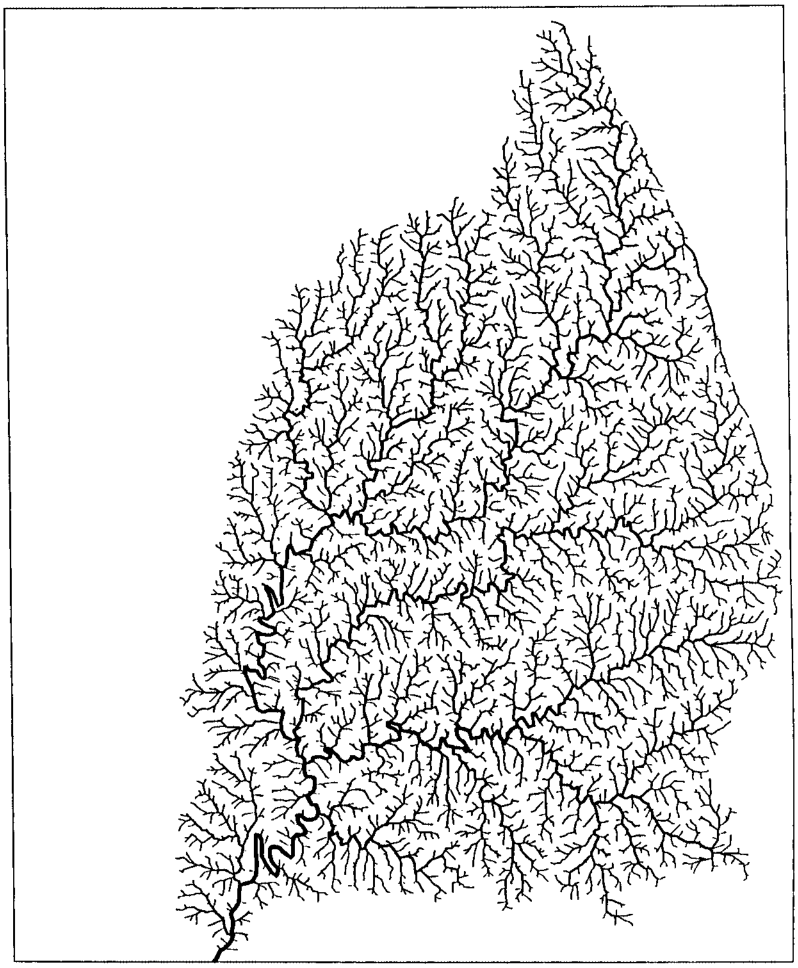

FIG. 2. *Kentucky river network.*

where a_n is a positive scale parameter. If we take $a_n = \sqrt{n}$ in (1.8), then for a critical Galton–Watson tree τ_n and under the existence of a moment generating function in a neighborhood of the origin for the weights, Troutman and Karlinger [23] show that $F_n \Rightarrow F$, where $F'(h) = (h/2)\exp(-h^2/4)$, a *Rayleigh density*, and "$\Rightarrow$" denotes *convergence in distribution*.

In Section 2 we will explain how the "expected" Galton–Watson critical binary tree considered by Troutman and Karlinger [23] may be viewed within a broader class of deterministic *self-similar trees*. This class of trees seems to have originated in the analysis of river network data (e.g., see [22] and [19]). To define this class of trees requires a notion of *network order* introduced by Horton [14] and refined by Strahler [21], according to the following algorithm. First the vertices of either degree 1 or 2 will be called *nonbranching*, where the

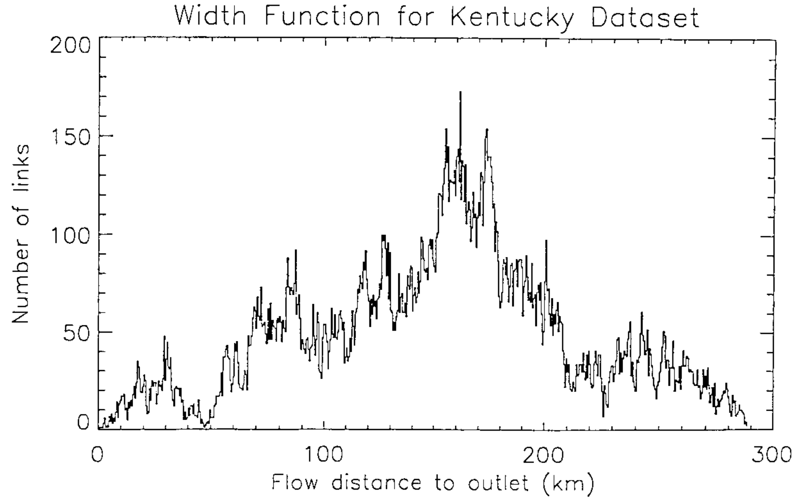

FIG. 3. *Kentucky river network width function.*

degree of a vertex counts the number of adjacent vertices including the parent. Those of degree one are called *leaves.* All leaves and adjacent paths of adjacent nonbranching vertices are assigned order 1. The orders of all other vertices (or associated edges) are recursively defined as the maximum of orders of the offspring vertices when these are not all equal, else it is the common order of the offspring incremented by 1. A contiguous path of edges of equal order is called a *stream* of the said order. The highest vertex belonging to a stream of a given order is called the *terminal* vertex for the stream. The order of the root ϕ defines the *order of the network* τ and is denoted $\omega(\tau)$. This scheme provides an "order or scale of resolution" in which the given tree is regarded to be at the finest scale of resolution and the next level of coarsening is obtained by removing the order 1 streams. The next level of "nonbranching" vertices in the pruned tree is assigned order 2. The next level of coarsening is obtained by pruning off the (lowest) order 2 streams and so on. The same algorithm for network order also occurs as an optimization parameter in binary arithmetic register allocation problems in computer science [8, 9].

Now, given the notion of order, a finite tree graph for which the number T_{ij} of order j subtrees rooted at nonterminal vertices of an order i stream is (a) the same for each order i stream in the network and (b) a function of i, j only through $i-j,\ j \le i$, is called *topologically self-similar*; that is, the matrix of generators is Toeplitz.

REMARK. In river network analysis one computes the sample average $\overline{T}_{ij}$ of the number of order j subtrees supported by the various streams of order i in the network. The generators for the Kentucky river network given in Figure 2 have the "approximate" Toeplitz form given by the matrix in Table 1.

TABLE 1
Sample generators for Kentucky river network

Tree generator matrix						
1.13	0.00	0.00	0.00	0.00	0.00	0.00
2.87	1.13	0.00	0.00	0.00	0.00	0.00
6.78	2.70	1.05	0.00	0.00	0.00	0.00
14.71	5.89	2.76	1.16	0.00	0.00	0.00
51.12	20.00	10.75	3.88	1.88	0.00	0.00
84.50	28.50	14.50	6.00	2.50	1.50	0.00
86.00	24.00	8.00	3.00	2.00	1.00	0.00

It is well-known in the hydrology literature (see [20]) that in the case of a critical binary Galton–Watson process one has

$$\mathbb{E}T_{ij} = \tfrac{1}{2}2^{i-j}. \tag{1.9}$$

A derivation of (1.9) will be given in Section 2. We will consider the width functions of the class of topologically self-similar trees for which

$$T_{ij} = \frac{(b-1)}{2} \cdot 2^{i-j}, \tag{1.10}$$

where $b \geq 2$ is an integer.

The paper is organized as follows. In Section 2 some technical preliminaries are provided which will be used to establish the main results. A principal tool is that of the associated polygonal walk, referred to as the *search-depth process*, (see [1] and [17]), together with our extension of a recent result of Aldous [2] on weak convergence of the search-depth process to Brownian excursion for the case of weighted trees; see [4], [12] and [7] for some related preliminaries and results on Brownian excursion. The weak convergence of the width function (local time) is then obtained using the well-known existence of a density for Brownian excursion (e.g., see [15]). Also included in Section 2 are statements and proofs of a few related results which are well known in the hydrology literature but less likely to be known among probabilists. The main results are stated and proved in Section 3.

2. Preliminaries.

The *search-depth process* is the polygonal path process obtained by following the contour of the tree (see [2]). More precisely, if τ_n is a labelled rooted tree with n vertices, then define

$$V_0(\tau_n) = \phi \tag{2.1}$$

and, given $V_k(\tau_n) = \langle i_1, \dots, i_m\rangle$, define $V_{k+1}(\tau_n) = \langle i_1, \dots, i_m, j\rangle$, where $j = \min\{i\colon \langle i_1, \dots, i_m, i\rangle \in \tau_n$ and $\langle i_1, \dots, i_m, i\rangle \neq V_l(\tau_n)\ \forall\ l < k+1\}$, provided the latter set is non-empty, in which case $V_{k+1}(\tau_n) = \langle i_1, \dots, i_{m-1}\rangle$. The usual search-depth process is then defined by

$$S_k(\tau_n) = h(V_k(\tau_n)), \qquad k = 0, 1, \dots, 2n, \tag{2.2}$$

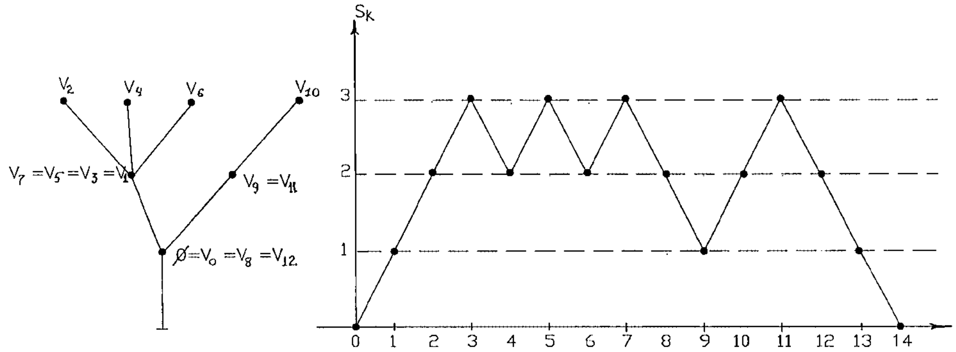

FIG. 4. *Standard search-depth process corresponding to τ_n.*

for the height function h defined in Section 1. Use linear interpolation between values to define the *scaled search-depth* $\{S_t^{(n)}(\tau_n)\colon 0 \le t \le 1\}$ such that $S_t^{(n)}(\tau_n) := S_k(\tau_n)/\sqrt{n}$ at $t = k/2n$; see Figure 4.

The following proposition is well known and simple to prove.

PROPOSITION 2.1. *Let τ_n be distributed as a Galton–Watson tree conditioned to have total progeny n. Then $\{S_k(\tau_n)\}_{k=0}^{2n}$ is distributed as a simple random walk conditioned to be positive over $k = 1, \dots, 2n-1$ and to hit zero for the first time at $k = 2n$ if and only if the offspring distribution is geometric.*

We refer to the conditioned random walk in Propositon 2.1 as a *random walk excursion*. It follows from results of Kaigh [16] and Durrett and Iglehart [6] that the random walk excursion suitably scaled converges to the continuous state *Brownian excursion* $\{W_0^+(t)\colon 0 \le t \le 1\}$. The following recent result of Aldous [2] fills the gap for more general offspring distributions.

THEOREM 2.2 [2]. *Let τ_n be a Galton–Watson tree conditioned to have total progeny n and whose offspring distribution L satisfies $\mathbb{E}L = 1$, $0 < \operatorname{Var}(L) = \sigma^2 < \infty$, $\gcd\{j\colon P(L = j) > 0\} = 1$. Then $\{S_t^{(n)}(\tau_n)\colon 0 \le t \le 1\} \Rightarrow \{2\sigma W_0^+(t)\colon 0 \le t \le 1\}$ as $n \to \infty$.*

For the applications of interest to us here we require an extension of Aldous's theorem to weighted trees. This is given in the next section (see Theorem 3.1).

The (weighted) search-depth process is defined by replacing h by H_W in (2.2). That is,

$$\hat{S}_k(\tau_n) = H_W(V_k(\tau_n)). \tag{2.3}$$

Again use linear interpolation between values to define the *weighted scaled search depth process* $\{\hat{S}_t^{(n)}\colon 0 \le t \le 1\}$ such that $\hat{S}_t^{(n)}(\tau_n) := \hat{S}_k(\tau_n)/\sqrt{n}$ at $t = k/2n$; see Figure 5.

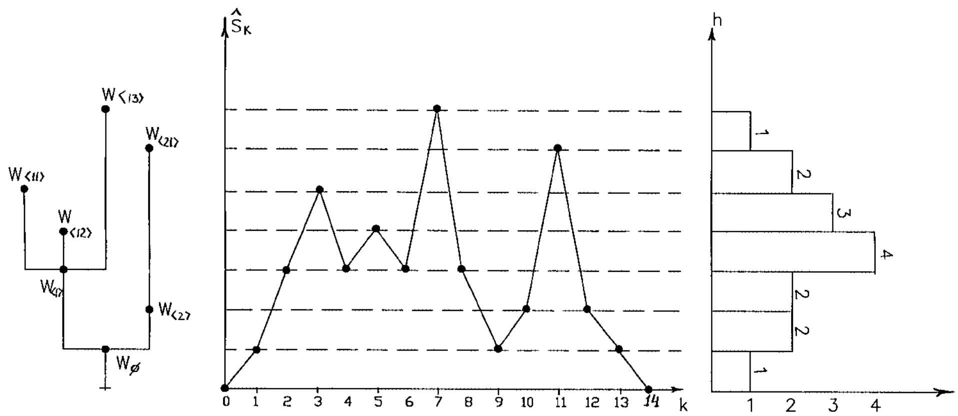

FIG. 5. *Weighted search-depth process corresponding to τ_n.*

For a Borel subset J of $\mathbb{R}$, define the (total) *weighted search-depth occupation time* by

$$\Gamma_n(J) = \tfrac{1}{2}\int_0^1 \mathbb{I}_J(\hat{S}_t^{(n)})\,dt \tag{2.4}$$

The *weighted search-depth local time* is then defined as the Radon–Nikodym derivative (with respect to Lebesgue measure λ on $\mathbb{R}$) accordingly as

$$\gamma_n(h) = \frac{d\Gamma_n}{d\lambda}(h). \tag{2.5}$$

Then observe that, on τ_n,

$$\gamma_n(h) = \frac{\sqrt{n}}{2n}\left[Z(\sqrt{n}h) + \sum_{k:\,\hat{S}_k \le \sqrt{n}h < \hat{S}_{k+1}} (W_{k+1}^{-1} - 1)\right]. \tag{2.6}$$

In particular, both the weighted search-depth local time and the weighted width function (i.e., Banach indicatrix) $Z(\sqrt{n}h)$ of the weighted search-depth process on τ_n are constant over identical intervals. In the case of constant weights the second term in (2.6) vanishes, while more generally we will see that it is $o(1)$ in probability as $n \to \infty$; see Theorem 3.3.

In Section 3 we shall show that the fluctuation law for the weighted search-depth occupation time and weighted width function, viewed as the density of a random measure, follows that of the occupation time for the Brownian excursion in the large total progeny limit. While neither the local time nor occupation time is a continuous functional, we will see that the occupation time is a.s. continuous. Of particular importance to us in this computation is the following proposition (see [15] for a proof).

PROPOSITION 2.3. *The distribution of $W_0^+(t)$ is absolutely continuous with respect to Lebesgue measure for each $t \in [0, 1]$.*

The expected value results of [23] will follow (see Corollary 3.4) from the fluctuation law calculations and a computation by Chung [5] of the expected occupation time of $\{W_0^+(t)\}$ contained in the following result.

THEOREM 2.4 [5]. *Let*

$$S([a,b)) = \int_0^1 \mathbb{I}(a \le W_0^+(t) < b)\,dt$$

for $0 < a \le b < \infty$. *Then*

$$\mathbb{E}S([a,b)) = \int_a^b 4h\exp(-2h^2)\,dh.$$

In the remainder of this section we will consider two classes of special deterministic self-similar trees, the so called *Peano trees* and *uniform b-ary trees*. We will calculate the width function asymptotics for the Peano trees in this section and that of the uniform b-ary trees in Section 3.

First let us recall the general notion of self-similarity introduced in Section 1. Consider a class of finite trees of orders $m = 2, 3, \dots$. Suppose that τ is such a tree of order m and let $T_{i,j}$ denote the number of subtrees of τ of order j rooted at nonterminal vertices of a stream of order i, where $2 \le i \le m$. The array $((T_{i,j}))$ is an $(m-1) \times (m-1)$ lower triangular matrix, referred to as the *generator matrix*.

DEFINITION 2.1. *Self-similar trees* are the trees with generators $\{T_{i,j}\}$ with the property that $T_{i,i-j} = T_j$, where T_j counts the number of subtrees of order $i - j$ rooted at the nonterminal vertices of a stream of order i.

Note that in terms of the above matrices, self-similar trees are defined by the condition that the generator matrices be *Toeplitz*, that is, have constant values along diagonals. In view of Proposition 2.1 the expected critical Galton–Watson binary tree is self-similar with generator 2^{i-j-1}.

Figure 6 shows several examples of self-similar trees of order 3. Note that two trees can have the same generators and still be different, owing to two freedoms in adding the edges. One is that edges can enter a tree from the left or right side. This is illustrated by comparing Figure 6a and 6b. The second freedom arises from the many possible ways in which subtrees of different orders can be interspersed, which is illustrated in Figure 6b and 6c.

The Peano tree is represented by a class of self-similar trees with branching number $b = 3$ and generators $\{T_1 = 0, T_k = 2^{k-1}\colon k = 2, 3, \dots\}$. Figure 7 provides examples of the Peano trees with order 2, 3 and 4. The dashed lines are subtrees which are put in according to the generators.

In order to compute the numbers of the kth generation of a Peano tree of order m, it is also convenient to draw cluster forms of the Peano trees corresponding to Figure 8 so that the recursive relations can be established. For example, a Peano tree with order 3 consists of four Peano trees of order 2. Three of them are located in parallel on the same level, and the other one is

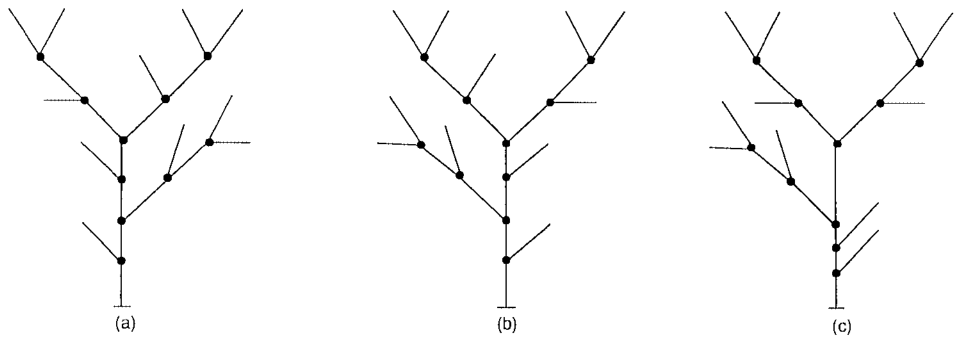

FIG. 6. *Self-similar trees:* $m = 3,\ b = 2,\ T_k = 2^{k-1}$.

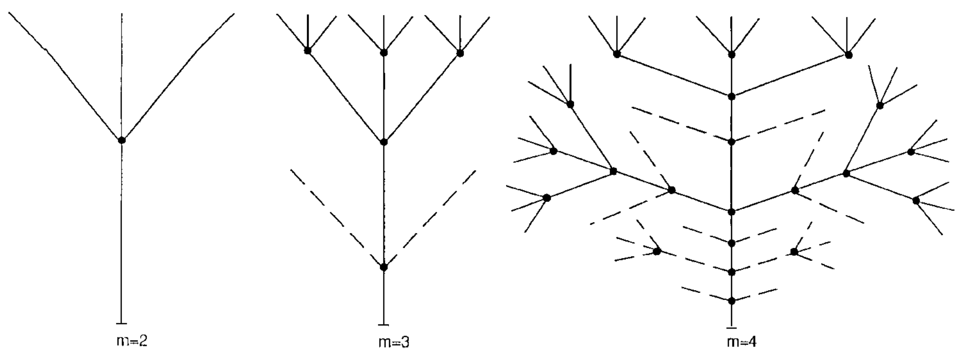

FIG. 7. *Peano trees:* $m = 2, 3, 4,\ b = 3,\ T_1 = 0,\ T_k = 2^{k-1}$.

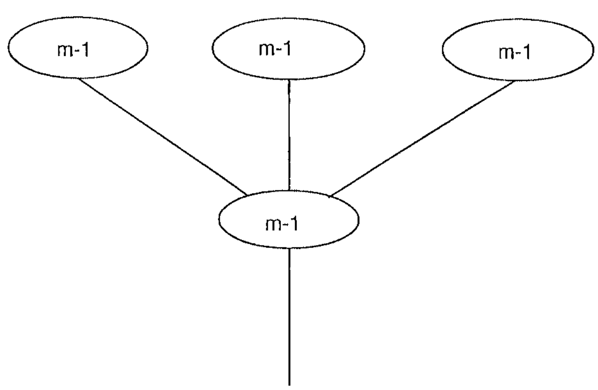

FIG. 8. *Cluster forms of Peano trees:* $m, b = 3,\ T_1 = 0,\ T_k = 2^{k-1}$.

below. In general, a Peano tree of order m has four sub-Peano trees of order $m-1$.

Let $Z_{m,k}$ be the numbers of the kth generation of a Peano tree of order m. The width functions of a self-similar tree of order m and generator $\{T_1 = 0,\ T_k = 2^{k-1};\ k = 1, 2, \dots\}$ is

$$Z_m(h) = \sum_{k=0}^{2^{m-1}-1} Z_{m,k+1} \cdot \mathbb{I}(k \le h < k+1), \tag{2.7}$$

where $m = 2, 3, \dots$. A simple induction argument shows that a Peano tree of order m has height 2^{m-1} and total progeny 4^{m-1}. We define a normalized width density f_m as

$$f_m(h) = \frac{Z_m(h)}{N_m}, \tag{2.8}$$

where $h \in [0, 2^{m-1})$ and $N_m = 4^{m-1}$. Then f_m defines a measure μ_m with distribution function

$$\mu_m[0, h] = \int_0^{2^{m-1}h} f_m(y)\,dy = \int_0^h 2^{m-1} f_m(2^{m-1}y)\,dy \tag{2.9}$$

for $h \in [0, 1)$. The following theorem formalizes observations of Marani, Rigon and Rinaldo [18].

THEOREM 2.5. *The measure μ_m converges weakly to a continuous singular probability measure on* $[0, 1]$ *as* $m \to \infty$, *namely, the induced infinite product measure* $(\frac{1}{4}\delta_0 + \frac{3}{4}\delta_1)^{\mathbb{N}}$ *under the map* $\phi(x) = (x_1, x_2, \dots),\ x = \sum_{i=1}^{\infty} x_i 2^{-i},\ x \in [0, 1]$.

PROOF. The width function of a Peano tree is

$$Z_m(h) = \sum_{k=0}^{2^{m-1}-1} Z_{m,k+1} \cdot \mathbb{I}(k \le h < k+1)$$

with corresponding width measure

$$f_m(h) = \frac{Z_m(h)}{N_m}.$$

Observe by induction that $Z_m(2^{m-1}h)/N_m,\ 0 \le h < 1$, is constant over each of the intervals

$$\Delta(x_1, \dots, x_{2^{m-1}}) = \left[\sum_{i=1}^{2^{m-1}} x_i 2^{-i}, \sum_{i=1}^{2^{m-1}} x_i 2^{-i} + 2^{-(m-1)}\right], \qquad x_i \in \{0, 1\},$$

with constant value

$$\left(\tfrac{1}{4}\right)^{\sum_{i=1}^{2^{m-1}} x_i} \times \left(\tfrac{3}{4}\right)^{2^{m-1} - \sum_{i=1}^{2^{m-1}} x_i}.$$

In view of tightness there is at least one limit probability measure and since any limit measure agrees with the asserted induced product measure on the semialgebra of cylinder sets of the form $\Delta(x_1, \dots, x_{2^{m-1}}),\ x_i \in \{0, 1\}$, there can be only one limit point and the proof is complete. □

REMARK. The multiscaling (cascade) structure of the width function (landforms) illustrated in Theorem 2.5 is of special interest to hydrologists in view of the observed multiscaling structure (random cascades) of rainfall (see [11], [13] and [25]). This scaling structure provides a basis for distributing rainfall over drainage basins commensurate with the observed variability and intermittancy properties of precipitation and provides a framework in which one can compute flood exponents in terms of the scaling characteristics of landforms and precipitation. A simple example to illustrate this point is worked out in [10] using the Peano tree.

Before introducing the next class of self-similar trees we sketch a proof of (1.9), a fact well known in hydrology and originally due to Shreve [20]. Shreve's original proof is based on generating function recursions while the proof sketched below is probabilistic. For this we first introduce a map π on the subset $\mathbb{T}_0$ of finite trees in $\mathbb{T}$ by $\pi(\{\phi\}) = \phi$, else $\pi(\tau)$ is the tree graph obtained by pruning the lowest order streams from τ. Also define $\overline{\tau}$ as the tree graph obtained by identifying adjacent vertices of degrees 1 or 2 with a single vertex. Then the order $\omega(\tau)$ of the tree may be expressed as

$$\omega(\tau) = \inf\{n\colon \pi^{(n-1)}(\tau) = \{\phi\}\}. \tag{2.10}$$

PROPOSITION 2.6 [20]. *Suppose τ is a critical binary Galton–Watson tree. Let T_{ij} denote the number of order j subnetworks in a randomly selected order i stream. Then*

$$\mathbb{E}T_{ij} = \tfrac{1}{2}2^{i-j}.$$

PROOF. By a simple induction one observes the lack of memory property for $\omega(\tau)$, from which it follows that

$$P(\omega(\tau) = k) = 2^{-k}, \qquad k = 1, 2 \dots .$$

The order of a subnetwork rooted at a vertex of order j therefore has order distributed as a truncated (at $j-1$) geometric distribution, that is, $2^{j-i-1}/(2^{j-1}-1),\ i = 1, 2, \dots, j-1$. The number of vertices in a stream of order j is geometrically distributed with expected value 2^{j-1}. Therefore, the expected number of subtrees of order i at an internal vertex in a stream of order j is $(2^{j-1}-1) \times 2^{j-i-1}/(2^{j-1}-1)$. □

REMARK. One may check by induction that, given $Z_1 = 2$, the (conditional) distribution of $\overline{\pi}(\tau)$ coincides with the (unconditional) distribution of τ. We refer to the invariance under the composite map $\overline{\pi}$ of the distribution of a finite random tree τ, conditional on $Z_0 > 0$, as a *stochastic self-similarity*.

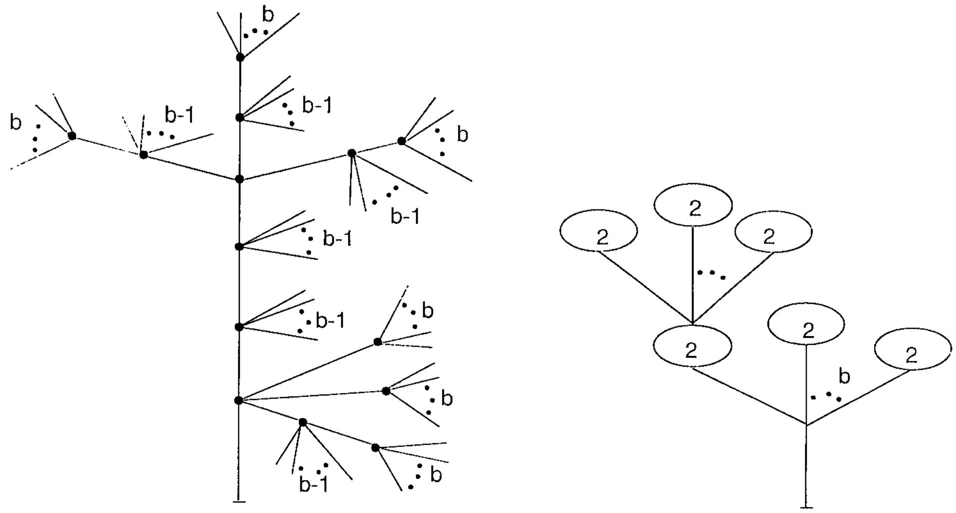

FIG. 9. *Cluster forms of uniform b-ary trees:* $m = 3,\ b = 3,\ T_k = (b-1)2^{k-1},\ k \ge 1$.

More generally, the second class of problems which we consider is that of computing the width function for the class of deterministic self-similar trees defined by generators $\{T_k(b) = (b-1)2^{k-1},\ k = 1, 2, \dots\}$ (see Fig. 9). We set the problem up in this section but provide the main result in Section 3.

As noted above in the case of Peano trees, self-similar trees can have the same generators and still be graph theoretically distinct due to the freedom in adding subtrees. Since we are only interested in the width functions, trees having same width functions are referred to as *trees without distinction.* We will construct those trees according to the following rules:

RULE 1. All edges are added to the trees from right sides.

RULE 2. Subtrees of order $m-1, m-2, \dots, 2, 1$ will be added to the *principal path* (a chain of edges connecting root to the top of the tree) of order m, so that two parallel subtrees of order $m-1$ are constructed in the principal path, together with the other two parallel subtrees of order $m-1$ in the upper level.

The *uniform b-ary self-similar trees* are defined as the trees without distinction having generators $\{T_k(b) = (b-1)2^{k-1},\ k = 1, 2, \dots\}$ In general, for trees of order m we can get the following cluster form of the uniform b-ary trees. The recursive equations of $\{Z_{m,k},\ k = 1, 2, \dots, 2^m - 1\}$ are

(2.11a) $$Z_{m,1} = 1,$$

(2.11b) $$Z_{m,j+1} = bZ_{m-1,j},$$

(2.11c) $$Z_{m,j+2^{n-1}} = bZ_{m-1,j},$$

for $j = 1, 2, \dots, 2^{m-1} - 1$. The total progeny is

$$N_m = 2bN_{m-1} + 1 = (2b)^{m-1} + (2b)^{m-2} + \cdots + 1 = \frac{(2b)^m - 1}{2b - 1}, \tag{2.12}$$

where $b = 2, 3, \dots,$ $n = 2, 3, \dots,$ and $Z_{1,1} = N_1 = 1$.

We will show in Section 3 that the width functions of a b-ary tree of given order, as a normalized probabilty measure, converge weakly to a uniform distribution function over [0, 1]. The following theorem of Troutman and Karlinger [23] is of interest for comparison of these results to expected behavior of critical Galton–Watson trees.

THEOREM 2.7 [23]. *Let $\tau_{n,m}$ be distributed as a Galton–Watson binary tree conditioned to have order m and total progency n. Then the expected width function converges weakly to a uniform distribution as $n \to \infty$.*

3. Main results.

THEOREM 3.1. *Let τ_n be a Galton–Watson tree conditioned to have total progeny n and whose offspring distribution L satisfies $\mathbb{E}L = 1$, $0 < \operatorname{Var} L = \sigma^2 < \infty$, $\gcd\{j\colon P(L = j) > 0\} = 1$. Suppose that the iid weights $\{W_{\langle e\rangle}\}$, positive and independent of τ_n, have mean 1 and variance s^2, and assume $\lim_{x\to\infty}(x \log x)^2 P(|W_\phi - 1| > x) = 0$. Then the scaled weighted search-depth process $\{\hat{S}_t^{(n)}(\tau_n)\colon 0 \le t \le 1\}$ converges in distribution to $\{2\sigma W_0^+(t)\colon 0 \le t \le 1\}$ as $n \to \infty$.*

PROOF. We will show that the weighted search depth process $\hat{S}^{(n)}$ and the unweighted search depth process $S^{(n)}$, defined on the same random tree τ_n with total progeny $\nu = n$, are asymptotically equivalent. The result then follows from Theorem 4.1 of [2].

Fix $\varepsilon > 0$. From Aldous's Theorem 2.2 we see that $S^{(n)}$ is tight. Choose M so that

$$\sup_n P\Bigl(\sup_t S_t^{(n)} > M \big| \nu = n\Bigr) < \varepsilon.$$

Let $A_M = [\sup_t S_t^{(n)} > M]$, $B_n = \{\max_{\langle e\rangle\in\tau_n} |W_{\langle e\rangle} - 1| \le b_n\}$, where b_n is a constant depending on n, $C_j = \{k\colon S_k = j\}$ and $|C_j| = \#\{k \in C_j\}$. Let $\mu_{n,j} := jE(W_\phi - 1)\mathbb{I}[|W_\phi - 1| \le b_n]$. By stratifying and using Bernstein's inequality (e.g., see [3]), for any $b_n > 0$,

$$\begin{aligned} &P\Bigl(\sup_t \bigl|\hat{S}_t^{(n)} - S_t^{(n)}\bigr| > \varepsilon \big| \nu = n\Bigr) \\ &\quad\le P\Bigl(A_M^c \cap B_n \cap \Bigl[\sup_t \bigl|\hat{S}_t^{(n)} - S_t^{(n)}\bigr| > \varepsilon\Bigr] \Big| \nu = n\Bigr) \\ &\qquad + P(A_M | \nu = n) + P(B_n^c | \nu = n) \end{aligned}$$

$$\begin{aligned}
&\le \sum_{1\le j\le M\sqrt{n}} P\Big(B_n \cap \Big[\max_{k\in C_j}|\hat{S}_k - S_k| > \varepsilon\sqrt{n}\Big]\Big|\nu = n\Big) \\
&\quad + \varepsilon + nP(|W_\phi - 1| > b_n) \\
&\le \sum_{1\le j\le M\sqrt{n}} \mathbb{E}\{|C_j|\,|\nu = n\} P\bigg(\bigg[\bigg|\sum_1^j (W_{\langle e_i\rangle} - 1)\bigg| > \varepsilon\sqrt{n}\bigg] \\
&\qquad\qquad \cap \Big[\max_{1\le i\le j}|W_{\langle e_i\rangle} - 1| \le b_n\Big]\bigg) \\
&\quad + \varepsilon + nP(|W_\phi - 1| > b_n) \\
&\le 2\sum_{1\le j\le M\sqrt{n}} \mathbb{E}\{|C_j|\,|\nu = n\} \exp\bigg\{-\frac{(\varepsilon\sqrt{n} - \mu_{n,j})^2}{2j(s^2 + 2(\varepsilon\sqrt{n} - \mu_{n,j})b_n/3j)}\bigg\} \\
&\quad + \varepsilon + nP(|W_\phi - 1| > b_n) \\
&\le 2n \max_{1\le j\le M\sqrt{n}} \exp\bigg\{-\frac{\sqrt{n}(\varepsilon - \mu_{n,j}/\sqrt{n})^2}{2M(s^2 + 2(\varepsilon - \mu_{n,j}/\sqrt{n})b_n/3M)}\bigg\} \\
&\quad + \varepsilon + nP(|W_\phi - 1| > b_n).
\end{aligned}$$

Take $b_n = b\sqrt{n}/\ln n$, where $b < 3\varepsilon/16$, to see that $\lim_n(\ln n/\ln b_n) = 2$, $\max_{1\le j\le M\sqrt{n}}|\mu_{n,j}| \le Mb^{-1}s^2\ln n$, and for all n sufficiently large

$$\begin{aligned}
P\Big(\sup_t|\hat{S}_t^{(n)} - S_t^{(n)}| > \varepsilon\Big|\nu = n\Big) &\le 2n\exp\{-2\ln n\} + \varepsilon \\
&\quad + \left(\frac{\ln n}{b_n \ln b_n}\right)^2 (b\ln b_n)^2 P(|W_\phi - 1| > b_n).
\end{aligned}$$

Now let $n \to \infty$ to see that $\hat{S}^{(n)}$ and $S^{(n)}$ are asymptotically equivalent. □

REMARK. In [7] it is shown that $x^2P(W_\phi > x) \to 0$ as $x \to \infty$ makes the weighted height of the tree $O(\sqrt{n})$. One may expect it should be possible to relax the moment condition on the weights in Theorem 3.1 a bit by removing the slowly varying factor $(\log x)^2$, but apart from this the moment condition should be best possible for the given scaling.

THEOREM 3.2. *Under the conditions given in Theorem 3.1, let $\gamma_n(h)$ denote the weighted search-depth local time density defined in (2.5). Then, as $n \to \infty$,*

$$\int_a^b \gamma_n(h)\,dh \Rightarrow \int_0^1 \mathbb{I}(a \le 2\sigma W_0^+(t) < b)\,dt$$

for any $0 < a < b$, where $\{W_0^+(t): 0 \le t \le 1\}$ is (standard) Brownian excursion.

PROOF. We have, for any $0 < a < b$,

$$\int_a^b \gamma_n(h)\,dh = \frac{1}{2}\int_0^1 \mathbb{I}\left(\frac{a}{\sigma} \le \frac{\hat{S}_t^{(n)}}{\sigma} < \frac{b}{\sigma}\right)dt.$$

For fixed $0 < a < b$ define F: $C[0,1] \to R$ by

$$F(w) = \int_0^1 \mathbb{I}(w(t) \in \Delta)\,dt$$
$$= \lambda\{t\colon w(t) \in \Delta\}$$

for $w \in C[0,1]$ and $\Delta = [a/\sqrt{2}, b/\sqrt{2})$; λ denotes Lebesgue measure. Then note that F is clearly not continuous on $C[0,1]$. Let D_F be the set of discontinuities of F. Therefore, in view of the above and Theorem 3.1, we need only show $W_0^+(D_F) = 0$, where we also use W_0^+ to denote the distribution of the process $\{W_0^+(t)\colon 1 \le t \le 1\}$. As we will see, the key to this proof is the fact that $\{W_0^+(t)\}$ has an absolutely continuous distribution with respect to Lebesgue measure for each t (cf. Proposition 2.3). Beyond this we apply a standard Fubini argument. Let $\mathbb{I}_\Delta(\omega, t) = \mathbb{I}(w(t) \in \Delta)$. Then for each $\omega \in C[0,1]$ the section $t \to \mathbb{I}_\Delta(\omega, t)$ is continuous Lebesgue-a.e. Note that since

$$F(\omega) = \int_0^1 \mathbb{I}_\Delta(w, t)\,dt,$$

it follows from Fubini's theorem that F is measurable. In view of Proposition 2.3, for $A = \{(\omega, t)\colon \omega(t) \in \partial\Delta\}$, $\partial\Delta = \{a/\sqrt{2}, b/\sqrt{2}\}$, one obviously has

$$W_0^+\big(\{\omega \in C[0,1]\colon (\omega, t) \in A\}\big) = \int_{\partial\Delta} p(t, y)\,dy = 0.$$

Therefore, by Fubini's theorem,

$$\int_{C[0,1]} \lambda\big(\{t \in [0,1]\colon (\omega, t) \in A\}\big) W_0^+(d\omega)$$
$$= \int_{[0,1]} W_0^+\big(\{\omega \in C[0,1]\colon (\omega, t) \in A\}\big)\lambda(dt)$$
$$= 0.$$

Therefore,

$$\lambda\big(\{t \in [0,1]\colon (\omega, t) \in A\}\big) = 0$$

for W_0^+-a.a. ω. Let

$$B = \big\{\omega \in C[0,1]\colon \lambda(\{t \in [0,1]\colon (\omega, t) \in A\}) = 0\big\}.$$

Then $W_0^+(B^c) = 0$ and if $\omega \in B$, then, for $\omega_n \in C[0,1]$ such that $\omega_n \to \omega$ in $C[0,1]$,

$$\mathbb{I}_\Delta(\omega_n, t) \to \mathbb{I}_\Delta(\omega, t) \quad \text{a.e. } t,$$

and therefore using Lebesgue's dominated convergence theorem one has

$$F(\omega_n) = \int_0^1 \mathbb{I}_\Delta(\omega_n, t)\,dt \to \int_0^1 \mathbb{I}_\Delta(\omega, t)\,dt = F(\omega).$$

Thus F is continuous except on a set of W_0^+-probability 0 and the proof is complete. □

THEOREM 3.3. *Assume the conditions of Theorems 3.1 and 3.2. Then*

$$\mathbb{E}\left\{\left(\left|2\int_0^h \gamma_n(y)\,dy - \frac{1}{\sqrt{n}}\int_0^h Z(\sqrt{n}y)\right|\right)\middle|\nu = n\right\} \to 0$$

as $n \to \infty$.

PROOF. First note that

$$\begin{aligned}
K_n &= \int_0^\infty \mu_n(h)\,dh \\
&= \int_0^\infty \mathbb{E}(Z(h)|\nu = n)\,dh \\
&= \int_0^\infty \mathbb{E}\left(\sum_{k:\,\hat{S}_k \le h < \hat{S}_{k+1}} 1 \middle| \nu = n\right) dh \\
&= \mathbb{E}\left(\int_0^\infty \sum_k 1[\hat{S}_k \le h < \hat{S}_{k+1}]\,dh \middle| \nu = n\right) \\
&= \mathbb{E}\left(\sum_k (\hat{S}_{k+1} - \hat{S}_k) 1[\hat{S}_{k+1} - \hat{S}_k > 0] \middle| \nu = n\right) \\
&= \mathbb{E}\left(\sum_k W_{k+1} 1[W_{k+1} > 0] \middle| \nu = n\right) \\
&= \mathbb{E}\left(\sum_{\langle e\rangle} W_{\langle e\rangle} \middle| \nu = n\right) \\
&= n.
\end{aligned}$$

It is not difficult to verify for the polygonal paths that

$$\gamma_n(h) = \frac{1}{2\sqrt{n}} \sum_{k:\,\hat{S}_k \le \sqrt{n}h < \hat{S}_{k+1}} W_{k+1}^{-1}.$$

Since also

$$Z(\sqrt{n}h) = \sum_{k:\,\hat{S}_k \le \sqrt{n}h < \hat{S}_{k+1}} 1,$$

one has

$$Z(\sqrt{n}h) = 2\sqrt{n}\left\{\gamma_n(h) - \sum_{k:\,\hat{S}_k \le \sqrt{n}h < \hat{S}_{k+1}} (W_{k+1}^{-1} - 1)\right\}.$$

Now,

$$F_n(h) = \int_0^h \sqrt{n}\frac{\mu_n(\sqrt{n}y)}{K_n}\,dy.$$

Therefore,

$$
\begin{aligned}
F_n(h) &= \frac{\sqrt{n}}{K_n}\int_0^h \mathbb{E}(Z(\sqrt{n}y)|\nu = n)\,dy \\
&= \frac{\sqrt{n}}{K_n}\int_0^h \mathbb{E}\biggl(\frac{2n}{\sqrt{n}}\int_0^h \gamma_n(y)\,dy - \int_0^h \sum_{k:\,\hat{S}_k\le\sqrt{n}y<\hat{S}_{k+1}} (W_{k+1}^{-1}-1)\,dy|\nu = n\biggr) \\
&= \frac{2n}{K_n}\mathbb{E}(\Gamma_n(0,h]|\nu = n) \\
&\quad - \frac{\sqrt{n}}{K_n}\mathbb{E}\biggl(\sum_{k=0}^{2n-1}\int_0^h 1[\hat{S}_k < \hat{S}_{k+1}]1\biggl[\frac{\hat{S}_k}{\sqrt{n}}, \frac{\hat{S}_{k+1}}{\sqrt{n}}\biggr)(y) \\
&\qquad\qquad \times (W_{k+1}^{-1}-1)\,dy\biggm|\nu = n\biggr).
\end{aligned}
$$

To complete the proof we will first show that the second term is $o(1)$. For this recall that the $W_{\langle e\rangle}$'s are positive, iid and independent of τ_n. One has

$$
\begin{aligned}
&\biggl|\frac{\sqrt{n}}{K_n}\mathbb{E}\biggl(\sum_{k=0}^{2n-1}\int_0^h 1[\hat{S}_k < \hat{S}_{k+1}]1\biggl[\frac{\hat{S}_k}{\sqrt{n}}, \frac{\hat{S}_{k+1}}{\sqrt{n}}\biggr)(y)(W_{k+1}^{-1}-1)\,dy\biggm|\nu = n\biggr)\biggr| \\
&\quad = \biggl|\frac{\sqrt{n}}{K_n}\mathbb{E}\biggl(\sum_{k=0}^{2n-1}\biggl\{\frac{(1-W_{k+1})}{\sqrt{n}}1\biggl[\frac{\hat{S}_k}{\sqrt{n}} < \frac{\hat{S}_{k+1}}{\sqrt{n}} \le h\biggr] \\
&\qquad\qquad + (W_{k+1}^{-1}-1)\biggl(h - \frac{\hat{S}_k}{\sqrt{n}}\biggr)1\biggl[\frac{\hat{S}_k}{\sqrt{n}} < h < \frac{\hat{S}_{k+1}}{\sqrt{n}}\biggr]\biggr\}\biggm|\nu = n\biggr)\biggr| \\
&\quad = \biggl|\frac{1}{\sqrt{n}}\mathbb{E}\biggl(\sum_{k=0}^{2n-1}\frac{(1-W_{k+1})}{\sqrt{n}}1\bigl[W_{k+1} > 0, \hat{S}_k < h\sqrt{n}\bigr]\biggm|\nu = n\biggr) \\
&\qquad - \frac{1}{\sqrt{n}}\mathbb{E}\biggl(\sum_{k=0}^{2n-1}\biggl\{\frac{(1-W_{k+1})}{\sqrt{n}} - (W_{k+1}^{-1}-1)\biggl(h - \frac{\hat{S}_k}{\sqrt{n}}\biggr)\biggr\} \\
&\qquad\qquad \times 1\biggl[\frac{\hat{S}_k}{\sqrt{n}} < h < \frac{\hat{S}_{k+1}}{\sqrt{n}}\biggr]\biggm|\nu = n\biggr)\biggr| \\
&\quad = \biggl|0 - \frac{1}{n}\mathbb{E}\biggl(\sum_{k=0}^{2n-1}(1-W_{k+1})\biggl(1 - \frac{\sqrt{n}h - \hat{S}_k}{W_{k+1}}\biggr)1\biggl[\frac{\hat{S}_k}{\sqrt{n}} < h < \frac{\hat{S}_{k+1}}{\sqrt{n}}\biggr]\biggm|\nu = n\biggr)\biggr| \\
&\quad \le \frac{1}{n}\mathbb{E}\biggl(\sum_{k=0}^{2n-1}|1-W_{k+1}|1\biggl[\frac{\hat{S}_k}{\sqrt{n}} < h < \frac{\hat{S}_{k+1}}{\sqrt{n}}\biggr]\biggm|\nu = n\biggr) \\
&\quad \le \frac{s}{n}\sum_{k=0}^{2n-1}P^{1/2}(\hat{S}_k < \sqrt{n}h < \hat{S}_{k+1}|\nu = n) \\
&\quad \le 2s\sqrt{\frac{\sum_{k=0}^{2n-1}P(\hat{S}_k < \sqrt{n}h < \hat{S}_{k+1}|\nu = n)}{2n}}.
\end{aligned}
$$

The zero term in the last equality of the above calculation is simply a consequence of the fact that $[W_{k+1} > 0]$ is equivalent to $[S_{k+1} > S_k]$ and the weights are assumed positive, mean 1 and independent of τ_n. In particular,

$$\begin{aligned}&\mathbb{E}\big((1 - W_{k+1})1[W_{k+1} > 0]1[\hat{S}_k < h\sqrt{n}]\big|\tau_n\big)\\ &\quad = \mathbb{E}\big((1 - W_{k+1})1[W_{k+1} > 0]\big)\mathbb{E}\big(1[\hat{S}_k < h\sqrt{n}]\big|\tau_n\big) = 0.\end{aligned}$$

Now let $B_n = \{\max_{\langle e\rangle\in\tau_n} W_{\langle e\rangle} \le \sqrt{n}/\log n\}$. In view of the tail condition on the weights one has $P(B_n^c) \to 0$ as $n \to \infty$. Also

$$1\big[\hat{S}_k < \sqrt{n}h < \hat{S}_{k+1}\big] \le 2n\int_{k/2n}^{(k+1)/2n} 1\left[h - \frac{W_{k+1}}{\sqrt{n}} < \hat{S}_t^{(n)} < h + \frac{W_{k+1}}{\sqrt{n}}\right] dt.$$

Thus, once W_{k+1} is bounded by $\sqrt{n}/\log n$ on B_n,

$$\begin{aligned}&1\big[\{\hat{S}_k < \sqrt{n}h < \hat{S}_{k+1}\} \cap B_n\big]\\ &\quad \le 2n\int_{k/2n}^{(k+1)/2n} 1\left[\left\{h - \frac{W_{k+1}}{\sqrt{n}} < \hat{S}_t^{(n)} < h + \frac{W_{k+1}}{\sqrt{n}}\right\} \cap B_n\right] dt\\ &\quad \le 2n\int_{k/2n}^{(k+1)/2n} 1\left[h - \frac{1}{\log n} < \hat{S}_t^{(n)} < h + \frac{1}{\log n}\right] dt.\end{aligned}$$

This gives

$$\frac{1}{2n}\sum_{k=0}^{2n-1} P(\hat{S}_k < \sqrt{n}h < \hat{S}_{k+1}) \le P(B_n^c) + \mathbb{E}\left(2\Gamma_n\left(h - \frac{1}{\log n}, h + \frac{1}{\log n}\right)\right) \to 0,$$

as $n \to \infty$. □

COROLLARY 3.4. *Assume the conditions of Theorems 3.1 and 3.2. Let*

$$F_n(h) = \int_0^{\sqrt{n}h} \frac{\mu_n(y)}{K_n}\, dy$$

as defined in (1.6) and (1.7). Then $F_n \Rightarrow F$, where $F'(h) = 2h\exp(-h^2),\ h \ge 0$.

PROOF. In view of Theorems 3.2 and 3.3 we have

$$\begin{aligned}\lim_{n\to\infty} F_n(x) &= \int_0^1 P\left[0 \le W_0^+(u) < \frac{x}{\sqrt{2}}\right] du\\ &= \int_0^1 \mathbb{E}\left[1\left(0 \le W_0^+(u) < \frac{x}{\sqrt{2}}\right)\right] du\\ &= \mathbb{E}\left[S\left(\left[0, \frac{x}{\sqrt{2}}\right)\right)\right].\end{aligned}$$

Thus, using Theorem 2.4, we have

$$\lim_{n\to\infty} F_n(x) = \int_0^{x/\sqrt{2}} 4y \exp(-2y^2)\, dy$$
$$= \int_0^x 2y \exp(-y^2)\, dx.$$

This concludes the proof. □

REMARK. The Rayleigh distribution in Theorem 3.3 differs from that in [23] by a scale factor of $\frac{1}{2}$ because τ_n in that paper is a binary tree.

Let us now consider the width function of the uniform b-ary trees of order m with generator $\{T_k(b) = (b-1)2^{k-1}, k = 1, 2, \ldots\}$. In particular, let

$$Z_m(x) = \sum_{k=0}^{2^m-2} Z_{m,\,k+1} \cdot \mathbb{I}(k \le x < k+1),$$

where $m = 2, 3, \ldots,$ and let

$$f_m(x) = \frac{Z_m(x)}{N_m}.$$

The height of b-ary tree of order m is given by $H_m = 2^m - 1$, so that the distribution function is

$$F_m(x) = \int_0^{H_m x} f_m(y)\, dy = \int_0^{H_m x} \frac{Z_m(y)}{N_m}\, dy = \int_0^x \frac{Z_m(H_m y)}{N_m/H_m}\, dy.$$

If we let F_m denote the distribution function defined by the density function

$$\frac{Z_m(H_m x)}{N_m/H_m},$$

then we have the following theorem.

THEOREM 3.5. *The distribution function F_m converges weakly to the uniform distribution on $[0, 1]$ as $m \to \infty$.*

The proof of above theorem will follow from the following two lemmas.

LEMMA 3.1. *Let $\phi_m(s)$ be the moment generating function of F_m,*

$$\phi_m(s) = \int_0^1 e^{sx} \frac{Z_m(H_m x)}{N_m/H_m}\, dx.$$

Then $\phi(s) = \lim_{m\to\infty} \phi_m(s)$ exists, for $s \in [0, 1]$.

PROOF. Let

$$\psi_m(s) = \int_0^{H_m} e^{sx} Z_m(x)\,dx.$$

Since

$$\begin{aligned}
Z_m(x) &= \sum_{k=0}^{2^m-2} Z_{m,\,k+1} \cdot \mathbb{I}(k \le x < k+1) \\
&= Z_{m,\,1} \cdot \mathbb{I}(0 \le x < 1) + \sum_{k=1}^{2^{m-1}-1} Z_{m,\,k+1} \cdot \mathbb{I}(k \le x < k+1) \\
&\quad + \sum_{k=2^{m-1}}^{2^m-2} Z_{m,\,k+1} \cdot \mathbb{I}(k \le x < k+1) \\
&= \mathbb{I}(0 \le x < 1) + \sum_{k=1}^{2^{m-1}-1} Z_{m,\,k+1} \cdot \mathbb{I}(k \le x < k+1) \\
&\quad + \sum_{k=1}^{2^{m-1}-1} Z_{m,\,k+2^{m-1}} \cdot \mathbb{I}(k \le x - 2^{m+1} + 1 < k+1),
\end{aligned}$$

we have from the recursive relation that

$$\begin{aligned}
Z_m(x) &= Z_{m-1,\,1} \cdot \mathbb{I}(0 \le x < 1) + b \cdot \sum_{k=1}^{2^{m-1}-1} Z_{m-1,\,k} \cdot \mathbb{I}(k \le x < k+1) \\
&\quad + b \cdot \sum_{k=1}^{2^{m-1}-1} Z_{m-1,\,k} \cdot \mathbb{I}(k \le x - 2^{m-1} + 1 < k+1) \\
&= Z_{m-1,1} \cdot \mathbb{I}(0 \le x < 1) + b \cdot \sum_{k=0}^{2^{m-1}-2} Z_{m-1,\,k+1} \cdot \mathbb{I}(k \le x - 1 < k+1) \\
&\quad + b \cdot \sum_{k=0}^{2^{m-1}-2} Z_{m-1,\,k+1} \cdot \mathbb{I}(k \le x - 2^{m-1} < k+1) \\
&= Z_{m-1,\,1} \cdot \mathbb{I}(0 \le x < 1) + bZ_{m-1}(x-1) \cdot \mathbb{I}(1 \le x < 2^{m-1}) \\
&\quad + bZ_{m-1}(x - 2^{(m-1)}) \cdot \mathbb{I}(2^{m-1} \le x < 2^m - 1).
\end{aligned}$$

Therefore,

$$\begin{aligned}
\psi_m(s) &= \int_0^{H_m} \exp(sx)\big[Z_{m-1,\,1} \cdot \mathbb{I}(0 \le x < 1) + bZ_{m-1}(x-1) \cdot \mathbb{I}(1 \le x < 2^{m-1}) \\
&\qquad\qquad + bZ_{m-1}(x - 2^{m-1}) \cdot \mathbb{I}(2^{m-1} \le x < 2^m - 1)\big]\,dx \\
&= \int_0^1 \exp(sx)\,dx + b\int_1^{2^{m-1}} \exp(sx) Z_{m-1}(x-1)\,dx \\
&\quad + b\int_{2^{m-1}}^{2^m-1} \exp(sx) Z_{m-1}(x - 2^{m-1})\,dx
\end{aligned}$$

$$= \frac{(\exp s) - 1}{s} + b \int_0^{2^{m-1}-1} \exp(s(x+1)) Z_{m-1}(x)\,dx + b \int_0^{2^{m-1}-1} \exp(s(x+2^{m-1})) Z_{m-1}(x)\,dx.$$

Finally we obtain the recursions

$$\psi_m(s) = \frac{(\exp s) - 1}{s} + b((\exp s) + \exp(2^{m-1}s))\psi_{m-1}(s),$$
$$\psi_0(s) = 0,$$
$$\psi_1(s) = \frac{(\exp s) - 1}{s},$$

where $m = 1, 2, \ldots$.

It now follows that, since

$$\phi_m(s) = \frac{1}{N_m} \psi_m\left(\frac{s}{H_m}\right),$$

we have

$$\phi_m(s) = \frac{1}{N_m} \frac{\exp(s/H_m) - 1}{s/H_m} + b\left(\exp \frac{s}{H_m} + \exp \frac{2^{m-1}s}{H_m}\right) \frac{N_{m-1}}{N_m} \phi_{m-1}\left(\frac{H_{m-1}}{H_m} s\right), \tag{3.1}$$
$$\phi_0(s) = 0,$$
$$\phi_1(s) = \frac{(\exp s) - 1}{s},$$

where $H_m = 2^m - 1$ and

$$N_m = \frac{(2b)^m - 1}{2b - 1}$$

for $m = 1, 2, \ldots$.

If we let $P_m(\cdot)$ be the probability measure defined by $\phi_m(s)$, let $\delta_{\{0\}}(\cdot)$ be the Dirac measure at 0 and let $\lambda(\cdot)$ be the uniform distribution on [0,1]. Then

$$P_m(\cdot) = \frac{1}{N_m}\lambda(H_m \cdot) + \frac{N_{m-1}}{N_m} b\left(\delta_{\{1/H_m\}}(\cdot) + \delta_{\{2^{m-1}/H_m\}}(\cdot)\right) * P_{m-1}\left(\frac{H_{m-1}}{H_m}\cdot\right),$$

where $*$ denotes convolution. Since $\phi_1(s) = (e^s - 1)/s$, $P_1(\cdot)$ is the uniform distribution on $[0, 1]$ and, by induction, $P_m(\cdot)$ has compact support $[0, H_{m-1}/H_m] \subset [0, 1]$. We also note that

$$\lim_{m\to\infty} \frac{N_{m-1}}{N_m} b = \lim_{m\to\infty} \frac{(2b)^{m-1} - 1}{(2b)^m - 1} b = \frac{1}{2}.$$

Thus by tightness $P_m(\cdot)$ has a weakly convergent subsequence to a probability measure $P(\cdot)$ which in view of the above recursion is a unique limit point. In particular, $\lim_{m\to\infty} \phi_m(s)$ exists by *Helly's selection theorem*. □

LEMMA 3.2. *Let* $\phi(s) = \lim_{m\to\infty} \phi_m(s)$. *Then*

$$\phi(s) = \frac{e^s - 1}{s}.$$

In particular, $\phi(s)$ *is the moment generating function of the uniform distribution on* $[0, 1]$.

PROOF. If we take limits on both sides of (3.1), then we have

$$\phi(s) = \frac{(1 + e^{s/2})}{2}\phi\left(\frac{s}{2}\right).$$

Let X be the random variable distributed on $[0, 1]$ having moment generating function $\phi(s)$. If we let $\{U_i\}$ be iid symmetric Bernoulli 0–1 valued, then the moment generating function of U_i is

$$\frac{1 + e^s}{2}.$$

In view of the recursion we have for any $n \geq 1$ that

$$X = \frac{U_1 + X}{2} = \sum_{i=1}^{n} \frac{U_i}{2^i} + \frac{X}{2^n},$$

where "=" is equality in distribution. Since $\{U_i/2^i\}$ are independent and $\sum_i \mathbb{E}(U_i/2^i) < \infty$ and $\sum_i \mathrm{Var}(U_i/2^i) < \infty$, it follows by Kolmogorov's theorem that $U = \sum_{i=1}^{\infty} U_i/2^i$ converges a.s. and therefore in distribution. For arbitrary $\varepsilon_i \in \{0, 1\}$ and $k \geq 1$,

$$P\left(\sum_{i=1}^{k} \frac{\varepsilon_i}{2^i} \leq U < \sum_{i=1}^{k} \frac{\varepsilon_i}{2^i} + \frac{1}{2^k}\right) = P(U_1 = \varepsilon_1, \ldots, U_k = \varepsilon_k)$$
$$= \frac{1}{2^k} = \left|\left[\sum_{i=1}^{k} \frac{\varepsilon_i}{2^i},\ \sum_{i=1}^{k} \frac{\varepsilon_i}{2^i} + \frac{1}{2^k}\right)\right|.$$

Since the binary rationals are dense in $(0, 1)$, it follows that $P(a \leq U < b) = b - a$, $0 \leq a < b \leq 1$. Therefore U is uniformly distributed over $[0, 1]$. Now $X/2^n$ goes to 0 in distribution and therefore X is uniformly distributed over $[0, 1]$. This concludes the proof of the theorem. □

Acknowledgment. This research was partially supported by grants from NSF and NASA.

REFERENCES

[1] ALDOUS, D. (1991). The continuum random tree II: an overview. In *Durham Symposium on Stochastic Analysis 1990* 23–70. Cambridge Univ. Press.
[2] ALDOUS, D. (1993). The continuum random tree III. *Ann. Probab.* 21 248–289.
[3] BENNETT, G. (1962). Probability inequalities for the sum of independent random variables. *J. Amer. Statist. Assoc.* 57 33–45.
[4] BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

[5] CHUNG, K. L. (1976). Excursions in Brownian motion. *Ark. Mat.* 14 155–177.
[6] DURRETT, R. T. and IGLEHART, D. L. (1977). Functionals of Brownian meander and Brownian excursion. *Ann. Probab.* 5 130–135.
[7] DURRETT, R., KESTEN, H. and WAYMIRE, E. (1991). On weighted heights of random trees. *J. Theoret. Probab.* 4 223–237.
[8] ERSHOV, A. P. (1958). On programming of arithmetic operations. *Comm. ACM* 1 3–6.
[9] FLAJOLET, P. and PRODINGER, H. (1986). Register allocation for unary–binary trees. *SIAM J. Comput.* 15 629–640.
[10] GUPTA, V. K., CASTRO, S. and OVER, T. M. (1996). Scaling exponents of peak flows from spatial rainfall and channel networks. Preprint.
[11] GUPTA, V. K. and WAYMIRE, E. (1995). Multiplicative cascades and spatial variability in rainfall, river networks and floods. In *Reduction and Predictability of Natural Disasters*. Sante Fe Institute, Santa Fe, NM.
[12] HARRIS, T. E. (1956). First passage and recurrence distribution. *Trans. Amer. Math. Soc.* 73 471–486.
[13] HOLLEY, R. and WAYMIRE, E. (1992). Multifractal dimensions and scaling exponents for strongly bounded random cascades. *Ann. Appl. Probab.* 2 819–845.
[14] HORTON, R. E. (1945). Erosional development of streams and their drainage basins: hydrophysical approach to quantitative morphology. *Geol. Soc. Am. Bull.* 56 275–370.
[15] ITÔ, K. and MCKEAN, H. (1965). *Diffusion Processes and Their Sample Paths*. Springer, Berlin.
[16] KAIGH, W. D. (1976). An invariance principal for random walk conditioned by a late return to zero. *Ann. Probab.* 4 115–121.
[17] LOUCHARD, G. (1986). Brownian motion and algorithmic complexity. *BIT* 26 17–34.
[18] MARANI, A., RIGON, R. and RINALDO, A. (1991). A note on fractal channel networks. *Water Resources Research* 27 3041–3049.
[19] PECKHAM, S. (1995). New results for self-similar trees with applications to river networks. *Water Resources Res.* 31 1023–1029.
[20] SHREVE, R. L. (1966). Statistical law of stream numbers. *Journal of Geology* 74 17–37.
[21] STRAHLER, A. N. (1957). Quantitative analysis of watershed geomorphology. *EOS Transactions of the American Geophysical Union* 38 913–920.
[22] TOKUNAGA, E. (1984). Consideration on the composition of drainage networks and their evolution. *Transactions of the Japanese Geomorphology Union* 5 71–77.
[23] TROUTMAN, B. and KARLINGER, M. R. (1984). On the expected width function for topologically random channel networks. *J. Appl. Probab.* 21 836–849.
[24] WAYMIRE, E. (1992). On network structure function computations. In *New Directions in Time Series Analysis* 2 363–373. Springer, Berlin.
[25] WAYMIRE, E. and WILLIAMS, S. C. (1996). A cascade decomposition theory with applications to Markov and exchangeable cascades. *Trans. Amer. Math. Soc.* 348 585–632.

DEPARTMENT OF MATHEMATICS
OREGON STATE UNIVERSITY
CORVALLIS, OREGON 97331-4605
E-MAIL: ossiand@math.orst.edu
waymire@math.orst.edu
zhang@math.orst.edu