Elect. Comm. in Probab. **14** (2009), 151–164

ELECTRONIC COMMUNICATIONS in PROBABILITY

AN ELEMENTARY PROOF OF HAWKES'S CONJECTURE ON GALTON-WATSON TREES

THOMAS DUQUESNE[1]
Laboratoire de Probabilités et Modèles Aléatoires; Université Paris 6, 4 Place Jussieu, Boite courrier 188, 75252 PARIS Cedex 05, FRANCE
email: thomas.duquesne@upmc.fr

Submitted November 12, 2008, *accepted in final form* March 23, 2009

AMS 2000 Subject classification: 60J80, 28A78.
Keywords: Galton-Watson tree; exact Hausdorff measure; boundary; branching measure; size-biased tree.

Abstract

In 1981, J. Hawkes conjectured the exact form of the Hausdorff gauge function for the boundary of supercritical Galton-Watson trees under a certain assumption on the tail at infinity of the total mass of the branching measure. Hawkes's conjecture has been proved by T. Watanabe in 2007 as well as other precise results on fractal properties of the boundary of Galton-Watson trees. The goal of this paper is to provide an elementary proof of Hawkes's conjecture under a less restrictive assumption than in T. Watanabe's paper, by use of size-biased Galton-Watson trees introduced by Lyons, Pemantle and Peres in 1995.

1 Introduction.

Fractal properties of the boundary of supercritical Galton-Watson trees have been intensively studied since the seminal paper by R.A. Holmes [9], who first studied the exact Hausdorff measure of a specific case, and since the paper by J. Hawkes [8] who determined the growth number of a Galton-Watson tree in the general case and proved that the Hausdorff dimension of its boundary is the logarithm of the mean of the offspring distribution (see also Lyons [14] for a simple proof). The problem of finding an exact Hausdorff function in the general setting has been studied by Q. Liu [11] who considered a large class of offspring distributions. Packing measure, thin and thick points as well as multifractal properties of the branching measure have been also investigated by Q. Liu [13, 12], P. Moerters and N. R. Shieh [16, 17] and T. Watanabe [20].
In the cases studied by Q. Liu [11], the corresponding Hausdorff measure coincides with the branching measure. This was predicted by J. Hawkes [8] who conjectured the general form of the Hausdorff gauge function for the boundary of supercritical Galton-Watson trees under a natural assumption on the right tail of the distribution of the total mass of the branching measure. This long standing conjecture has been recently solved by T. Watanabe in [21]: in this paper,

[1] RESEARCH SUPPORTED BY ANR-08-BLAN-0190

T. Watanabe provides necessary and sufficient conditions for the existence of an exact Hausdorff measure that is absolutely continuous with respect to the branching measure as well as precise results on several important examples. The goal of our paper is to provide an elementary proof of Hawkes's conjecture that holds true under a less restrictive assumption than in Watanabe's paper and that relies on size-biased Galton-Watson trees that have been introduced in [15] by R. Lyons, R. Pemantle and Y. Peres.

Let us briefly state Hawkes's conjecture (we refer to Section 2 for more formal definitions). To simplify notation, we assume that all the random variables that we consider are defined on the same probability space $(\Omega, \mathscr{F}, \mathbf{P})$ that is assumed to be complete and sufficiently large to carry as many independent random variables as we need. Let $\xi = (\xi(k), k \in \mathbb{N})$ be a probability distribution on $\mathbb{N}$ that is viewed as the offspring distribution of a Galton-Watson tree $\mathscr{T}$. Informally, $\mathscr{T}$ is the family-tree of a population stemming from one ancestor and evolving randomly as follows: each individual of the population has an independent random number of children distributed in accordance with the offspring distribution ξ. We assume that the ancestor is at generation 0, its children are at generation 1, their children are at generation 2... and so on; if we denote by $Z_n(\mathscr{T})$ the number of individuals at generation n, then $(Z_n(\mathscr{T})\,;\, n \geq 0)$ is a Galton-Watson Markov chain with offspring distribution ξ that starts at state 1. To avoid trivialities, we assume that $\xi(1) < 1$. Let us set

$$\mathrm{m} = \sum_{k \in \mathbb{N}} k\, \xi(k) \in [0, \infty] \quad \text{and} \quad f(r) = \sum_{k \in \mathbb{N}} \xi(k) r^k \,,\; r \in [0,1] \,. \tag{1}$$

The function f is the generating function of ξ. It is convex and the equation $f(r) = r$ has at most two roots in $[0,1]$. We denote by q the smallest one (1 being obviously the largest one). Standard results on Galton-Watson Markov chains imply that $\mathbf{P}(\#\mathscr{T} < \infty) = q$. Morever $q = 1$ iff $\mathrm{m} \leq 1$. Therefore, if $\mathrm{m} > 1$, $\mathbf{P}(\#\mathscr{T} = \infty) > 0$. We shall make the following assumptions on ξ:

$$\mathrm{m} \in (1, \infty) \qquad \text{and} \qquad \sum_{k \geq 2} k \log(k)\, \xi(k) \,<\, \infty \,. \tag{2}$$

We suppose $\mathrm{m} > 1$ because we want to study $\mathscr{T}$ on the event $\{\#\mathscr{T} = \infty\}$. The second assumption on ξ is motivated by the following result known as Kesten-Stigum's Theorem that asserts that, under (2), we have

$$\lim_{n \to \infty} \frac{Z_n(\mathscr{T})}{\mathrm{m}^n} = W \quad \text{a.s. and in } \mathrm{L}^1(\Omega, \mathscr{F}, \mathbf{P}) \,. \tag{3}$$

Consequently, W has unit expectation: $\mathbf{E}[W] = 1$. Moreover, $\mathbf{1}_{\{W>0\}} = \mathbf{1}_{\{\#\mathscr{T}=\infty\}}$ almost surely and the exact rate of growth of $\mathscr{T}$ is $n \mapsto \mathrm{m}^n W$. We refer to the original paper by Kesten and Stigum [10] and to the paper by R. Lyons, R. Pemantle and Y. Peres [15] for an elementary proof of (3). The distribution of W is of great interest and it is known to have a positive continuous density on $(0, \infty)$ and an atom at 0 of mass q (see Athreya and Ney [3], Chapter II).

We view $\mathscr{T}$ as an ordered rooted tree. Namely, we take the ancestor of the population as the root and we associate a rank of birth with any individual. In this way we can label each individual by a finite word of positive integers $(i_1, \ldots, i_n)$: the length n of the word is the generation of the labelled individual and i_k represents the birth-rank of its ancestor at generation k. The set of words created in this way completely encodes $\mathscr{T}$ which can be therefore viewed as a random subset of the set of finite words written with positive integers. We denote by $\mathbb{U}$ the set of finite words written by positive integers. We are interested in the infinite boundary of $\mathscr{T}$ that is the set of all infinite lines of descent in $\mathscr{T}$. We denote by $\partial\mathscr{T}$ the boundary of $\mathscr{T}$. Let us assume that $\partial\mathscr{T}$ is non-empty (which is equivalent to assume that $\#\mathscr{T}$ is infinite); we denote by $\mathbb{N}^* = \mathbb{N}\backslash\{0\}$ the set of positive integers; to each infinite line of descent we associate an infinite $\mathbb{N}^*$-valued sequence

$(i_n\,;\,n \geq 1)$ such that for any $n \geq 1$, the finite word $(i_k\,;\,1 \leq k \leq n)$ is the label of the individual at generation n belonging to the infinite line of descent. Therefore, we can view $\partial\mathscr{T}$ as a random subset of the set of $\mathbb{N}^*$-valued and $\mathbb{N}^*$-indexed sequences that we denote by $\partial\mathbb{U}$. We equip $\partial\mathbb{U}$ with the metric δ defined as follows: if $\mathbf{u} = (i_k\,;\,k \geq 1)$ and $\mathbf{v} = (j_k\,;\,k \geq 1)$ are two elements of $\partial\mathbb{U}$, then $\delta(\mathbf{u},\mathbf{v}) = \exp(-n)$, where n is the largest integer m such that $\mathbf{u}$ and $\mathbf{v}$ agree on the first m terms (n is taken as 0 if $i_1 \neq j_1$). The resulting metric space $(\partial\mathbb{U},\delta)$ is complete and separable and $\partial\mathscr{T}$ is actually a random compact subset of $\partial\mathbb{U}$.

Hawkes's conjecture concerns the problem of finding a gauge function g such that the g-Hausdorff measure of $\partial\mathscr{T}$ is positive and finite. One natural candidate for such a measure is the branching measure M that is a finite random measure on $\partial\mathbb{U}$ associated with $\partial\mathscr{T}$ such that $M(\partial\mathscr{T}) = W$ a.s. Roughly speaking, M is the most spread out measure on $\partial\mathscr{T}$ (we refer to Section 2 for more details on M). Before stating more precisely the conjecture, let us briefly recall standard definitions about Hausdorff measures: we restrict our attention to g-Hausdorff measures with sufficiently regular gauge function g; more precisely, we say that g is a regular function if firstly, there exists an interval $(0,r_0)$ on which g is right-continuous non-decreasing, secondly, $\lim_0 g = 0$ and thirdly, there exists $C > 1$ such that $g(2r) \leq Cg(r)$, for any $r \in (0,r_0/2)$ (this last assumption is often called the "doubling condition" though some authors use the term of "blanketed" Hausdorff function). Then, the g-Hausdorff measure on $(\partial\mathbb{U},\delta)$ is defined as follows: for any $A \subset \partial\mathbb{U}$ and for any $\varepsilon \in (0,r_0)$, we first set

$$\mathscr{H}_g^{(\varepsilon)}(A) = \inf\left\{ \sum_{n\in\mathbb{N}} g\left(\mathrm{diam}(C_n)\right) \ ;\, A \subset \bigcup_{n\in\mathbb{N}} C_n \ \text{ and } \ \mathrm{diam}(C_n) \leq \varepsilon\,,\, n \in \mathbb{N} \right\},$$

where $\mathrm{diam}(C) = \sup\{\delta(\mathbf{u},\mathbf{v})\,;\,\mathbf{u},\mathbf{v} \in C\}$ stands for the diameter of a subset $C \subset \partial\mathbb{U}$; then the g-Hausdorff measure is given by $\mathscr{H}_g(A) = \lim_{\varepsilon\to 0} \mathscr{H}_g^{(\varepsilon)}(A) \in [0,\infty]$.

• *Hawkes's conjecture.* Let ξ be a probability measure on $\mathbb{N}$ that satisfies (2). Let $\mathscr{T}$ be a Galton-Watson tree with offpring distribution ξ. Let W be defined by (3). We set

$$F(x) := -\log \mathbf{P}(W > x)\,. \tag{4}$$

We first assume that F is regularly varying at ∞; more precisely, we suppose that F is of the following form:

$$F(x) = x^b \ell(x) \tag{5}$$

where $b > 0$ and ℓ is slowly varying function at ∞. We denote by F^{-1} the right-continuous inverse of F: $F^{-1}(x) = \inf\{y \geq 0 \;:\; F(y) > x\}$ and we set

$$g(r) = r^{\log \mathrm{m}} F^{-1}\big(\log\log 1/r\big)\,, \quad r \in (0,e^{-1})\,. \tag{6}$$

Then, Hawkes [8] p.382 conjectured that under (5), there exists $c_\xi \in (0,\infty)$ that only depends on ξ such that

$$\mathscr{H}_g(\partial\mathscr{T}) = c_\xi W \quad \text{a.s.} \tag{7}$$

As already mentioned, this result has been solved by T. Watanabe (2007) in Theorem 1.6 [21] under the following assumption that is weaker than (5): for any sufficiently large x

$$A^{-1}x^b\ell(x) \leq F(x) \leq Ax^b\ell(x)\,, \tag{8}$$

where A is a constant larger than one. The paper by T. Watanabe [21] contains other general results. (Actually, Watanabe's proof of Hawkes's conjecture is a consequence of Theorem 1.2 [21] that provides a general criterion to decide whether the branching measure is an exact Hausdorff measures with a regular gauge function.) The goal of this paper is to give an alternative short proof of Hawkes's conjecture under a less restrictive assumption than (8) and by use of different techniques that we claim to be elementary.

Theorem 1.1. *Let ξ be a probability measure on $\mathbb{N}$ which satisfies (2). Let $\mathscr{T}$ be a Galton-Watson tree with offpring distribution ξ. Let W be defined by (3); let F be defined by (4) and let g be defined by (6). We assume*

$$\sup_{x\in[1,\infty)} \frac{F^{-1}(2x)}{F^{-1}(x)} < \infty \,. \tag{9}$$

Then, there exists $c_\xi \in (0,\infty)$ that only depends on ξ such that

$$\mathbf{P}\text{–a.s. for } M\text{–almost all } \mathbf{u}: \quad \limsup_{r\to 0} \frac{M(B(\mathbf{u},r))}{g(r)} = c_\xi^{-1} \,, \tag{10}$$

where $B(\mathbf{u},r)$ stands for the open ball in $(\partial\mathbb{U},\delta)$ with center $\mathbf{u}$ and radius r. Furthermore, we have

$$\mathbf{P}\text{–a.s} \qquad \mathscr{H}_g(\,\cdot\cap\partial\mathscr{T}) = c_\xi\cdot M \,, \tag{11}$$

where M stands for the branching measure associated with $\partial\mathscr{T}$.

Remark 1.1. *Let us prove that (9) is a weaker assumption than (8). Indeed, assume that (9) is not satisfied, then there exists a sequence $x_n \in [1,\infty)$, $n\in\mathbb{N}$, increasing to ∞ such that $F^{-1}(2x_n) \geq nF^{-1}(x_n)$, for any $n\in\mathbb{N}$. Recall that W has a positive density on $(0,\infty)$, so that F is continuous and increasing. Thus, if one sets $y_n = F^{-1}(x_n)$, one easily gets*

$$2F(y_{n+p}) \geq F(ny_{n+p})\,, \quad n,p\in\mathbb{N}.$$

Suppose that F satisfies (8). Then, $(y_n, n\in\mathbb{N})$ tends to ∞ and the previous inequality entails

$$2A\cdot y_{n+p}^b\,\ell(y_{n+p}) \geq A^{-1}\cdot n^b y_{n+p}^b\,\ell(ny_{n+p})\,, \quad n,p\in\mathbb{N}.$$

Fix n such that $n^b > 2A^2$. Then, the previous inequality rises a contradiction since

$$\lim_p \ell(y_{n+p})/\ell(ny_{n+p}) = 1,$$

by definition of slow variation functions. □

Remark 1.2. *Let us briefly discuss (9) for further uses. It is easy to prove that (9) is equivalent to the following:*

$$\exists a>0\,, \quad F^{-1}(sx) \leq 2^a s^a F^{-1}(x)\,, \quad s,x\in[1,\infty)\,. \tag{12}$$

Therefore, F satisfies

$$2^{-1}s^{1/a}F(x) \leq F(sx)\,, \quad s\geq 2^a, x\geq F^{-1}(1)\,. \tag{13}$$

Consequently, $\mathbf{P}(W>y) \leq \exp(-Cy^{1/a})$, for any $y \geq 2^a F^{-1}(1)$, where $C = 2^{-1}(F^{-1}(1))^{-1/a}$. This implies that $\mathbf{E}[W^n] < \infty$ for any $n\geq 1$, which is equivalent to $\sum_{k\geq 0} k^n\xi(k) < \infty$ for any $n\geq 1$, by a standard result (see Rem. 3, p. 33 in [3] or Theorem 0 in Bingham and Doney [4]). Thus, (9) is much stronger than (2). □

Remark 1.3. *There is no known necessary and sufficient condition expressed in terms of ξ for F to satisfy (9) (neither for (8) nor (5)). However specific cases have been considered by Q. Liu and T. Watanabe: see [11] and [21]. Let us also mention that when (11) holds true, there is no simple general closed formula giving c_ξ in terms of ξ. The avaible results characterizing c_ξ are either quite involved or they require the knowledge of the distribution of W: see Theorem 1.1 in [21] for a general characterization of c_ξ; see Liu (Theorem 1[11]) or Watanabe (Theorem 1.6 [21]) when the support of ξ is bounded; see also Watanabe (Theorem 1.6 [21]) when F satisfies (8).* □

2 Notation and basic definitions.

Let us start with basic notation: we denote by $\mathbb{N}$ the set of nonnegative integers and by $\mathbb{N}^*$ the set of positive integers; let $\mathbb{U} = \bigcup_{n\in\mathbb{N}}(\mathbb{N}^*)^n$ be the set of finite words written with positive integers, with the convention $(\mathbb{N}^*)^0 = \{\varnothing\}$. Let $u = (i_k\,;\, 1 \leq k \leq n) \in \mathbb{U}$; we set $|u| = n$ that is the length of u, with the convention $|\varnothing| = 0$. Words of unit length are identified with positive integers. For any $m \in \mathbb{N}$ we set $u_{|m} = (i_k\,;\, 1 \leq k \leq n \wedge m)$, with the convention $u_{|0} = \varnothing$; observe that $u_{|m} = u$ if $m \geq n$. Let $v = (j_k; 1 \leq k \leq m) \in \mathbb{U}$, we define $u * v \in \mathbb{U}$ by the word $(\ell_k\,;\, 1 \leq k \leq n+m)$ where $\ell_k = i_k$ if $k \leq n$ and $\ell_k = j_{k-n}$ if $k > n$: the word $u * v$ is the concatenation of u and v (observe that $\varnothing * u = u * \varnothing = u$). We next introduce the genealogical order $\preceq$ by writing $u \preceq v$ iff $v_{||u|} = u$. For any $u, v \in \mathbb{U}$, we denote by $u \wedge v$ the $\preceq$-maximal word w such that $w \preceq u$ and $w \preceq v$. For any $u \in \mathbb{U}$, we denote by $\mathbb{U}_u$ the $\preceq$-successors of u. Namely, $\mathbb{U}_u$ is the set of words $u * v$ where v varies in $\mathbb{U}$. On $\mathbb{U}_u$, we define the u-shift θ_u by $\theta_u(u * v) = v$.

Definition 2.1. *A subset $T \subset \mathbb{U}$ is a tree iff it satisfies the following conditions.*

- Tree(1)*: If $u \in T$, then $u_{|m} \in T$, for any $m \in \mathbb{N}$ (in particular $\varnothing$ belongs to T).*
- Tree(2)*: For any $u \in \mathbb{U}$, there exists $k_u(T) \in \mathbb{N} \cup \{-1\}$ such that the following properties hold true.*
  - *If $k_u(T) = -1$, then $u \notin T$.*
  - *If $k_u(T) = 0$, then $T \cap \mathbb{U}_u = \{u\}$.*
  - *If $k_u(T) \geq 1$, then the set of words $\{u * i\,;\, 1 \leq i \leq k_u(T)\}$ is exactly the set of words $v \in T$ such that $|v| = |u| + 1$ and $u \preceq v$.* □

Note that $u \mapsto k_u(T)$ is uniquely determined by T. If we view T as the family tree of a population whose $\varnothing$ is the ancestor, then $k_u(T)$ represents the number of children of $u \in T$. We denote by $\mathbb{T}$ the class of subsets of $\mathbb{U}$ satisfying Tree(1) and Tree(2). More precisely, this definition provides a canonical coding of finite-degree ordered rooted trees. For sake of simplicity, any element T in $\mathbb{T}$ shall be called a tree.
For any $m \in \mathbb{N}$ and any $T \in \mathbb{T}$, we set $T_{|m} = \{u \in T \;:\; |u| \leq m\}$. Observe that $T_{|m}$ is a finite tree. For any word $u \in \mathbb{U}$ and any tree T, we define the u-shift of T by

$$\theta_u T = \theta_u\left(T \cap \mathbb{U}_u\right) = \{v \in \mathbb{U} \;:\; u * v \in T\}.$$

We see that $\theta_u T$ is empty iff $u \notin T$. If $u \in T$, then $\theta_u T$ represents the subtree that starts at u (or the set of the descendents of u). Observe that in any case, $\theta_u T$ is a tree according to Definition 2.1. For any $u \in T$, we define the tree T cut at vertex u as the following subset of $\mathbb{U}$:

$$\mathrm{Cut}_u T = T \backslash (\mathbb{U}_u \backslash \{u\}).$$

Observe that $u \in \mathrm{Cut}_u T$ and that $\mathrm{Cut}_u T$ is a tree; $\mathrm{Cut}_u T$ represents the set of individuals that are not strict descendents of u. Next, for any $T \in \mathbb{T}$ and any $n \in \mathbb{N}$, we set

$$Z_n(T) = \#\{u \in T \ : \ |u| = n\} \ \in \mathbb{N}.$$

In the graph-terminology $Z_n(T)$ is the number of vertices of T at distance n from the root. If we view T as the family tree of a population whose ancestor is $\varnothing$ and whose genealogical order is $\preceq$, then $Z_n(T)$ is the number of individuals at the n-th generation.

We denote by $\partial\mathbb{U}$ the set $(\mathbb{N}^*)^{\mathbb{N}^*}$ of the $\mathbb{N}^*$-valued and $\mathbb{N}^*$-indexed sequences. Let $\mathbf{u} = (i_k\,;\, k \geq 1)$ be in $\partial\mathbb{U}$. For any $m \geq 0$, we set $\mathbf{u}_{|m} = (i_k\,;\, 1 \leq k \leq m) \in \mathbb{U}$, with convention $\mathbf{u}_{|0} = \varnothing$. If $\mathbf{v} \in \partial\mathbb{U}\backslash\{\mathbf{u}\}$, then we denote by $\mathbf{u} \wedge \mathbf{v}$ the $\preceq$-maximal finite word w such that $\mathbf{u}_{|_{|w|}} = \mathbf{v}_{|_{|w|}} = w$; we also set $\mathbf{u} \wedge \mathbf{u} = \mathbf{u}$. We equip $\partial\mathbb{U}$ with the following ultrametric δ given by

$$\delta(\mathbf{u}, \mathbf{v}) = \exp\left(-|\mathbf{u} \wedge \mathbf{v}|\right) \, .$$

The resulting metric space $(\partial\mathbb{U}, \delta)$ is separable and complete and we denote by $\mathscr{B}(\partial\mathbb{U})$ its Borel sigma-field. For any $r \in (0, \infty)$ and for any $\mathbf{u} \in \partial\mathbb{U}$, we denote by $B(\mathbf{u}, r)$ the open δ-ball with center $\mathbf{u}$ and radius r. We shall often use the notation

$$n(r) = \lfloor (-\log(r))_+ \rfloor + 1 \, , \tag{14}$$

where $(\cdot)_+$ stands for the positive part function and $\lfloor \cdot \rfloor$ for the integer part function. Observe that $B(\mathbf{u}, r) = \partial\mathbb{U}$ if $r > 1$; if $r \in (0, 1]$, $B(\mathbf{u}, r)$ is the set of $\mathbf{v}$ such that $\mathbf{v}_{|n(r)} = \mathbf{u}_{|n(r)}$. This has several consequences. *Firstly*, any open ball is also a closed ball. *Secondly*, we have $B(\mathbf{v}, r) = B(\mathbf{u}, r)$ for any $\mathbf{v}$ in $B(\mathbf{u}, r)$; therefore there is only a countable number of balls with positive radius; more precisely for any $u \in \mathbb{U}$, we set

$$B_u = \left\{ \mathbf{v} \in \partial\mathbb{U} \ : \ \mathbf{v}_{|_{|u|}} = u \right\}. \tag{15}$$

Then,

$$\{B(\mathbf{u}, r)\,;\, r \in (0, \infty), \mathbf{u} \in \partial\mathbb{U}\} = \left\{B_u\,;\, u \in \mathbb{U}\right\} \, .$$

Thirdly, for any pair of balls either they are disjoint or one is contained in the other. We shall further refer to these properties as to the *specific properties of balls in* $\partial\mathbb{U}$.

Let $T \subset \mathbb{T}$. We define the boundary ∂T by the set $\left\{ \mathbf{u} \in \partial\mathbb{U} \ : \ \mathbf{u}_{|n} \in T \, , \, n \in \mathbb{N}^* \right\}$. Obviously, ∂T is empty iff T if finite; moreover, since $T_{|m}$ is finite for any $m \in \mathbb{N}$, an easy diagonal-extraction argument implies that ∂T is a compact set of $(\partial\mathbb{U}, \delta)$.

We equip $\mathbb{T}$ with the sigma field $\mathscr{G}$ generated by the subsets

$$A_u := \{T \in \mathbb{T} \ : \ u \in T\} \ , \quad u \in \mathbb{U} \, . \tag{16}$$

For any $u \in \mathbb{U}$, it is easy to check that the function $T \mapsto k_u(T)$ is $\mathscr{G}$-measurable. Moreover $T \mapsto \theta_u T$, $u \mapsto \mathrm{Cut}_u(T)$ are $(\mathscr{G}, \mathscr{G})$-measurable and for any $n \in \mathbb{N}$, $T \mapsto Z_n(T)$ is $\mathscr{G}$-measurable.

Recall that we suppose throughout the paper that all the random variables we need are defined on the same probability space $(\Omega, \mathscr{F}, \mathbf{P})$. Then, a random tree $\mathscr{T}$ is a function from Ω to $\mathbb{T}$ that is $(\mathscr{F}, \mathscr{G})$-measurable. We shall concentrate our attention on a particular class of random trees called Galton-Watson trees that can be recursively defined as follows.

Definition 2.2. *Let $\xi = (\xi(k)\,;\, k \in \mathbb{N})$ be a probability on $\mathbb{N}$. A random tree $\mathscr{T}$ is said to be a Galton-Watson tree with offspring distribution ξ (a GW(ξ)-tree for short) if its distribution on $(\mathbb{T}, \mathscr{G})$ is characterized by the following conditions.*

- GW(1)*: the* $\mathbb{N}$*-valued random variable* $k_\varnothing(\mathscr{T})$ *is distributed in accordance with* ξ*.*
- GW(2)*: If* $\xi(k)>0$*, then under* $\mathbf{P}\left(\,\cdot\,|k_\varnothing(\mathscr{T})=k\right)$*, the first generation subtrees* $\theta_1\mathscr{T},\dots,\theta_k\mathscr{T}$ *are independent with the same distribution as* $\mathscr{T}$ *under* $\mathbf{P}$*.* □

This recursive definition induces a unique distribution on $(\mathbb{T},\mathscr{G})$: we refer to Neveu [18] for a proof.

Recall that we assume (2) and recall (3). Then, for any $u\in\mathbb{U}$ we set $W_u:=\limsup \mathrm{m}^{-n}Z_n(\theta_u\mathscr{T})$. Definition 2.2 easily implies that for any $u\in\mathbb{U}$, the random variables $\left(W_{u*v}\,;\,v\in\mathbb{U}\right)$ under $\mathbf{P}(\,\cdot\,|\,u\in\mathscr{T})$ have the same distribution as the random variables $\left(W_v\,;\,v\in\mathbb{U}\right)$ under $\mathbf{P}$. Therefore, we almost surely have

$$\forall\, u\in\mathbb{U}\quad W_u=\lim_{n\to\infty}\mathrm{m}^{-n}Z_n(\theta_u\mathscr{T})<\infty\quad\text{and}\quad \mathbf{1}_{\{W_u>0\}}=\mathbf{1}_{\{\partial(\theta_u\mathscr{T})\neq\emptyset\}}. \tag{17}$$

Moreover, since for any $n\geq 1$ and for any $u\in\mathbb{U}$, $Z_n(\theta_u\mathscr{T})$ is the sum of the $Z_{n-1}(\theta_{u*i}\mathscr{T})$'s over $i\in\mathbb{N}^*$, we a.s. have

$$\forall u\in\mathbb{U}\,,\quad W_u=\mathrm{m}^{-1}\sum_{i\in\mathbb{N}^*}W_{u*i}\,. \tag{18}$$

Denote by $\mathscr{M}_f(\partial\mathbb{U})$ the set of finite measures on $(\partial\mathbb{U},\mathscr{B}(\partial\mathbb{U}))$; equip $\mathscr{M}_f(\partial\mathbb{U})$ with the topology of weak convergence. Then, a random finite measure on $(\partial\mathbb{U},\mathscr{B}(\partial\mathbb{U}))$ is an function from Ω to $\mathscr{M}_f(\partial\mathbb{U})$ that is measurable with respect to $\mathscr{F}$ and to the Borel sigma-field of $\mathscr{M}_f(\partial\mathbb{U})$. Thanks to (18), the collection of random variables $(W_u\,;\,u\in\mathbb{U})$ allows to define a random finite measure M on $(\partial\mathbb{U},\mathscr{B}(\partial\mathbb{U}))$ that is characterized by the following properties.

- BM(1): Almost surely, M is diffuse and its topological support is $\partial\mathscr{T}$.
- BM(2): Almost surely, for any $u\in\mathbb{U}$, $M(B_u)=\mathrm{m}^{-|u|}W_u$, where we recall notation B_u from (15).

This defines the branching measure associated with $\mathscr{T}$. It is in some sense the most spread out measure on $\partial\mathscr{T}$; it is a natural candidate to be a Hausdorff measure on ∂T. Actually, Theorem 1.1 is proved by applying to M the following Hausdorff-type comparison results of measures.

Proposition 2.1. *Let* $\mu\in\mathscr{M}_f(\partial\mathbb{U})$*. Let* $g:(0,r_0)\to(0,\infty)$ *be a right-continuous and non-decreasing function such that* $\lim_0 g=0$ *and such that there exists* $C>1$ *that satisfies* $g(2r)\leq Cg(r)$*, for any* $r\in(0,r_0/2)$*. Then, for any Borel subset* $A\subset\partial\mathbb{U}$*, the following assertions hold true.*

- *(i) If* $\limsup_{r\to 0}\frac{\mu(B(\mathbf{u},r))}{g(r)}\leq 1$ *for any* $\mathbf{u}\in A$*, then* $\mathscr{H}_g(A)\geq C^{-1}\mu(A)$*.*
- *(ii) If* $\limsup_{r\to 0}\frac{\mu(B(\mathbf{u},r))}{g(r)}\geq 1$ *for any* $\mathbf{u}\in A$*, then* $\mathscr{H}_g(A)\leq C\mu(A)$*.*

This is a standard result in Euclidian spaces: see Lemmas 2 and 3 of Rogers and Taylor [19] for the original proof. We refer to [6] for a general version in metric spaces (more precisely, we refer to Theorem 4.15 [6] in combination with Proposition 4.24 [6]). We shall actually need the following more specific result.

Lemma 2.2. *Let* $\mu\in\mathscr{M}_f(\partial\mathbb{U})$*. Let* $g:(0,r_0)\to(0,\infty)$ *be a right-continuous and non-decreasing function such that* $\lim_0 g=0$ *and such that there exists* $C>1$ *that satisfies* $g(2r)\leq Cg(r)$*, for any* $r\in(0,r_0/2)$*. First assume that there is* $\kappa\in(0,\infty)$ *such that the following holds true.*

$$\text{For }\mu\text{–almost all }\mathbf{u}\,,\quad \limsup_{r\to 0}\frac{\mu(B(\mathbf{u},r))}{g(r)}=\kappa\,. \tag{19}$$

Next, assume that there is $\kappa_0 \in (0,\kappa)$ such that:

$$\mathscr{H}_g\left(\left\{\mathbf{u}\in\partial\mathbb{U}\ :\ \limsup_{r\to 0} g(r)^{-1}\mu(B(\mathbf{u},r))\leq\kappa_0\right\}\right)=0 \tag{20}$$

Then, $\mathscr{H}_g(\,\cdot\cap\operatorname{supp}\mu)=\kappa^{-1}\mu$, where $\operatorname{supp}\mu$ stands for the topological support of μ.

Remark 2.1. *Such a result holds true thanks to the specific properties of the balls of $(\partial\mathbb{U},\delta)$; this can be extended to general Polish spaces if μ satisfies a Strong Vitali Covering Property: see Edgar [6] for a discussion of this topic. Lemma 2.2 is probably known, however the author is unable to find a reference. That is why a brief proof is provided below.* □

Proof of Lemma 2.2: recall from (15) notation B_u, $u\in\mathbb{U}$; if A is a non-empty subset of $\partial\mathbb{U}$ that is not reduced to a point, then there exists $u\in\mathbb{U}$ such that $A\subset B_u$ and $\operatorname{diam}(B_u)=e^{-|u|}=\operatorname{diam}(A)$. This allows to take the set of balls as the covering set in the definition of $\mathscr{H}_g$, which then coincides with the so-called spherical g-Hausdorff measure. Let us denote by E the set of all $\mathbf{u}\in\partial\mathbb{U}$ where the limsup in (19) holds. Then (19) implies that $\mu(\partial\mathbb{U}\backslash E)=0$. Let K be a compact subset of $\operatorname{supp}\mu$. Since two balls of $\partial\mathbb{U}$ are either disjoint or one contains the other, then for any $p\geq 1$, there exists a finite sequence of pairwise disjoint balls $B_{u_1^p}$, ... , $B_{u_{n_p}^p}$ with respective diameters $r_i^p=\exp(-|u_i^p|)\leq p^{-1}$, $1\leq i\leq n_p$, such that

$$\forall 1\leq i\leq n_p\ ,\quad K\cap B_{u_i^p}\neq\emptyset\ ,\quad K\subset\bigcup_{i=1}^{n_p}B_{u_i^p}\quad\text{and}\quad\lim_{p\to\infty}\sum_{i=1}^{n_p}g(r_i^p)=\mathscr{H}_g(K)\,.$$

Since $K\subset\operatorname{supp}\mu$ and $K\cap B_{u_i^p}\neq\emptyset$, then $\mu(B_{u_i^p})>0$, and it makes sense to define a function f_p on $\partial\mathbb{U}$ by

$$f_p(\mathbf{u})=\sum_{i=1}^{n_p}\mathbf{1}_{B_{u_i^p}}(\mathbf{u})\frac{g(r_i^p)}{\mu(B_{u_i^p})}\,.$$

Observe that $\int f_p\,d\mu=\sum_{i=1}^{n_p}g(r_i^p)$. Next, for any $\mathbf{u}\in K\cap E$, denote by $j^p(\mathbf{u})$ the unique index $i\in\{1,\ldots,n_p\}$ such that $\mathbf{u}\in B_{u_i^p}$. Observe that $f_p(\mathbf{u})$ is equal to $g(r^p_{j^p(\mathbf{u})})/\mu(B(\mathbf{u},r^p_{j^p(\mathbf{u})}))$. Moreover, we have

$$\forall\mathbf{u}\in K\cap E\ ,\quad\liminf_{p\to\infty}\frac{g(r^p_{j^p(\mathbf{u})})}{\mu(B(\mathbf{u},r^p_{j^p(\mathbf{u})}))}\geq\liminf_{r\to 0}\frac{g(r)}{\mu(B(\mathbf{u},r))}=\kappa^{-1}\,.$$

Then, Fatou's lemma implies

$$\kappa^{-1}\mu(K\cap E)\leq\int_{K\cap E}\liminf_{p\to\infty}f_p(\mathbf{u})\,\mu(d\mathbf{u})\leq\liminf_{p\to\infty}\int_{\partial\mathbb{U}}f_p(\mathbf{u})\mu(d\mathbf{u})=\mathscr{H}_g(K)\,,$$

which entails

$$\kappa^{-1}\mu(K)\leq\mathscr{H}_g(K)\,. \tag{21}$$

Let us prove the converse inequality. Thanks to the specific properties of the balls of $(\partial\mathbb{U},\delta)$, for any $\eta\in(0,1)$, we can find a sequence of pairwise disjoint balls $(B(\mathbf{u}_n,r_n)\,;\,n\in\mathbb{N})$ whose diameters are smaller than η, such that $\mathbf{u}_n\in K\cap E$ for any $n\in\mathbb{N}$ and such that the following holds true:

$$K\cap E\ \subset\ \bigcup_{n\in\mathbb{N}}B(\mathbf{u}_n,r_n)\quad\text{and}\quad g(r_n)\leq(1+\eta)\kappa^{-1}\,\mu(B(\mathbf{u}_n,r_n))\,,\ n\in\mathbb{N}.$$

This implies that for any $\eta \in (0,1)$,

$$\mathscr{H}_g^{(\eta)}(K \cap E) \leq (1+\eta)\kappa^{-1}\,\mu(K^\eta), \tag{22}$$

where $K^\eta = \{\mathbf{u} \in \partial\mathbb{U} \ : \ \delta(\mathbf{u}, K) \leq \eta\}$. By letting η go to 0, standard arguments entail

$$\mathscr{H}_g(K \cap E) \leq \kappa^{-1}\,\mu(K)\,. \tag{23}$$

We next have to prove that $\mathscr{H}_g(\operatorname{supp}\mu \cap (\partial\mathbb{U}\backslash E)) = 0$: let $b > a > 0$; we set $E_{a,b} = \{\mathbf{u} \in \operatorname{supp}\mu \ : \ \limsup_{r\to 0} g(r)^{-1}\mu(B(\mathbf{u},r)) \in [a,b)\}$. Proposition 2.1 (ii) implies that $\mathscr{H}_g(E_{a,b}) \leq Ca^{-1}\mu(E_{a,b})$. Therefore $\mathscr{H}_g(E_{a,b}) = 0$ if $a > \kappa$ or if $0 < a < b < \kappa$ and (20) easily entails $\mathscr{H}_g(\operatorname{supp}\mu \cap (\partial\mathbb{U}\backslash E)) = 0$. This, combined with (21) and (23), entails that for any compact set K, $\mathscr{H}_g(K \cap \operatorname{supp}\mu) = \kappa^{-1}\mu(K)$ and standard arguments complete the proof. ■

3 Size biased trees.

We now introduce random trees with a distinguished infinite line of descent that are called size-biased trees and that have been introduced by R. Lyons, R. Pemantle and Y. Peres in [15] to provide a simple proof of Kesten-Stigum theorem. To that end, let us first set some notation: let $T \in \mathbb{T}$ be such that $\partial T \neq \emptyset$. Let $\mathbf{u} \in \partial T$. We set

$$\mathrm{Gr}(\mathrm{T}, \mathbf{u}) = \left\{ v \in T\backslash\{\varnothing\} \ : \ v \neq \mathbf{u}_{|_{|v|}} \text{ and } v_{|_{|v|-1}} = \mathbf{u}_{|_{|v|-1}} \right\}\,.$$

$\mathrm{Gr}(\mathrm{T}, \mathbf{u})$ is the set of individuals of T whose parent belongs to the infinite line of descent determined by $\mathbf{u}$ but who do not themself belong to the $\mathbf{u}$-infinite line of descent (namely, such individuals have a sibling on the $\mathbf{u}$-infinite line of descent). In other words, $\mathrm{Gr}(\mathrm{T}, \mathbf{u})$ is the set of the vertices where are grafted the subtrees stemming from the $\mathbf{u}$-infinite line of descent.

Let ξ be an offspring distribution that satisfies (2). The size-biased offspring distribution ξ is the probability measure $\widehat{\xi}$ given by $\widehat{\xi}(k) = k\xi(k)/\mathrm{m}$, $k \geq 0$. We define a probability distribution ρ on $\mathbb{N}^* \times \mathbb{N}^*$ by

$$\rho(k,\ell) = \mathbf{1}_{\{k \leq \ell\}}\mathrm{m}^{-1}\xi(\ell) = \mathbf{1}_{\{k \leq \ell\}}\ell^{-1}\widehat{\xi}(\ell)\,, \quad k,\ell \geq 1\,,$$

and we call ρ the repartition distribution associated with ξ.

Definition 3.1. *Let $(\mathscr{T}^*, \mathbf{U}^*) : \Omega \to \mathbb{T} \times \partial\mathbb{U}$ be a $(\mathscr{F}, \mathscr{G} \otimes \mathscr{B}(\partial\mathbb{U}))$-measurable function; $(\mathscr{T}^*, \mathbf{U}^*)$ is a ξ-size-biased Galton-Watson tree (a $\widehat{\mathrm{GW}}(\xi)$-tree for short) iff the following holds.*

- Size-bias(1): *for any $n \geq 0$, $\mathbf{U}^*_{|n} \in \mathscr{T}^*$. Moreover, if we set $\mathbf{U}^* = (i^*_n\,;\, n \in \mathbb{N}^*)$, then the sequence of $\mathbb{N}^* \times \mathbb{N}^*$-valued random variables $(i^*_n\,;\, k_{\mathbf{U}^*_{|n-1}}(\mathscr{T}))$, $n \geq 1$ is i.i.d. with distribution ρ.*
- Size-bias(2): *conditional on the sequence $((i^*_n\,;\, k_{\mathbf{U}^*_{|n-1}}(\mathscr{T}))\,;\, n \geq 1)$, the subtrees $\theta_u\mathscr{T}$, where u ranges in $\mathrm{Gr}(\mathscr{T}^*, \mathbf{U}^*)$, are i.i.d. $\mathrm{GW}(\xi)$-trees.* □

The size-biased tree $\mathscr{T}^*$ can be informally viewed as the family-tree of a population containing two kinds of individuals: the mutants and the non-mutants; each individual has an independent offspring: the non-mutants's one is distributed according to ξ and the mutants's one according to $\widehat{\xi}$; moreover, the ancestor is a mutant and a mutant has exactly one child who is a mutant (its other children being non-mutants); the rank of birth of the mutant child is chosen uniformly at random among the progeny of its mutant genitor; so there is only one mutant per generation and $\mathbf{U}^*$ represents the ancestral line of the mutants.

Size-biased trees have been introduced Lyons, Pemantle and Peres [15]. In this paper, the authors mention related constructions: we refer to their paper for a detailed bibliographical account. When $\mathrm{m} \leq 1$, the size-biased tree $\mathscr{T}^*$ has one single infinite line of descent and it is called a sintree after Aldous [2]. Such a biased tree is related to (sub)critical Galton-Watson trees conditioned on non-extinction: see Grimmett [7], Aldous and Pitman [1] and also [5] for more details.
The name "size-biased tree" is explained by the following elementary result (whose proof is left to the reader): let $G_1 : \mathbb{T} \times \mathbb{U} \to [0,\infty)$ and $G_2 : \mathbb{T} \to [0,\infty)$ be two measurable functions. Then, for any $n \geq 0$, we have

$$\mathbf{E}\left[\sum_{u\in\mathscr{T}:|u|=n} G_1\left(\mathrm{Cut}_u\mathscr{T}\,;\,u\right)G_2(\theta_u\mathscr{T})\right] = \mathrm{m}^n\,\mathbf{E}\left[G_1\left(\mathrm{Cut}_{\mathbf{U}^*_{|n}}\mathscr{T}^*\,;\,\mathbf{U}^*_{|n}\right)\right]\mathbf{E}\left[G_2(\mathscr{T})\right], \tag{24}$$

with the convention that a sum over an empty set is null. This immediately entails

$$\mathbf{P}\left(\mathscr{T}^*_{|n}\in dT\right) = \frac{Z_n(T)}{\mathrm{m}^n}\,\mathbf{P}\left(\mathscr{T}_{|n}\in dT\right)\ ,$$

which explains the name "size-biased" tree. Recall notation M for the branching measure. We derive from (24) the following key-lemma.

Lemma 3.1. *Let ξ be an offspring distribution that satisfies (2). Let $(\mathscr{T}^*,\mathbf{U}^*)$ be a $\widehat{\mathrm{GW}}(\xi)$-tree and let $\mathscr{T}$ be a $\mathrm{GW}(\xi)$-tree. Let $G:\mathbb{T}\times\partial\mathbb{U}\to[0,\infty)$ be $\mathscr{G}\otimes\mathscr{B}(\partial\mathbb{U})$-measurable. Then,*

$$\mathbf{E}\left[\int_{\partial\mathbb{U}} M(d\mathbf{u})\,G(\mathscr{T};\mathbf{u})\right] = \mathbf{E}\left[\,G(\mathscr{T}^*;\mathbf{U}^*)\,\right]. \tag{25}$$

Proof: let us first assume that $G(\mathscr{T};\mathbf{u}) = G(\mathscr{T}_{|n};\mathbf{u}_{|n})$. Recall that $M(\theta_u\mathscr{T}) = m^{-n}W_u$. Then, observe that

$$\int_{\partial\mathbb{U}} M(d\mathbf{u})\,G(\mathscr{T};\mathbf{u}) = \sum_{u\in\mathscr{T}:|u|=n} G(\mathscr{T}_{|n},\mathbf{u}_{|n})m^{-n}W_u\ .$$

Since $\mathbf{E}[W_u \mid u\in\mathscr{T}] = 1$, (24) implies (25). Recall notation A_u, $u\in\mathbb{U}$ from (16). The previous result implies that (25) holds true for functions G of the form $\mathbf{1}_C$ where C belong to $\mathscr{P}$:

$$\mathscr{P} = \{(A_{u_1}\cap\ldots\cap A_{u_p})\times B_v\ ;\ p\in\mathbb{N}^*,\ v,u_1,\ldots,u_p\in\mathbb{U}\}\ ,$$

which is a pi-system generating $\mathscr{G}\otimes\mathscr{B}(\partial\mathbb{U})$. Then, a standard monotone-class argument entails the desired result. ■

Let us apply Lemma 3.1 to our purpose: let $\mathscr{T}$ be a $\mathrm{GW}(\xi)$-tree and let $\mathbf{u}\in\partial\mathscr{T}$; recall notation $\mathrm{Gr}(\mathscr{T},\mathbf{u})$ from the beginning of the section. It is easy to prove that for any $r\in(0,1]$,

$$M(B(\mathbf{u},r)) = \sum_{p\geq 0}\ \sum_{\substack{v\in\mathrm{Gr}(\mathscr{T},\mathbf{u})\\ |v|=n(r)+1+p}} \mathrm{m}^{-p-n(r)-1}W_v\ , \tag{26}$$

where we recall that $n(r) = \lfloor -\log(r)\rfloor + 1$. Let $(\mathscr{T}^*,\mathbf{U}^*)$ be a $\widehat{\mathrm{GW}}(\xi)$-tree. For any $v\in\mathrm{Gr}(\mathscr{T}^*,\mathbf{U}^*)$, we set

$$W^*_v = \limsup_{n\to\infty}\frac{Z_n(\theta_v\mathscr{T}^*)}{\mathrm{m}^n}\ .$$

By Size-bias(2), observe that conditional on the $\mathbb{N}^* \times \mathbb{N}^*$-valued sequence $\big((i^*_n \, ; \, k_{\mathbf{U}^*_{|n-1}}(\mathscr{T}^*)) \, ; \, n \geq 1 \big)$, the random variables $(W^*_v \, ; v \in \mathrm{Gr}(\mathscr{T}^*, \mathbf{U}^*))$ are i.i.d. with the same distribution as W. Next, for any $n \geq 1$, we set

$$Y_n = \sum_{\substack{v \in \mathrm{Gr}(\mathscr{T}^*, \mathbf{U}^*) \\ |v| = n}} W^*_v \quad \text{and} \quad X_n = \sum_{p \geq 0} m^{-p} Y_{p+n} \, .$$

Lemma (3.1) combined with (26) implies that for any measurable nonnegative function G on the set of real valued cadlag functions on $(0, 1]$, one has

$$\mathbf{E}\left[\int_{\partial \mathbb{U}} M(d\mathbf{u}) \, G\Big(\big(M(B(\mathbf{u}, r)) \big)_{r \in (0,1]} \Big) \right] = \mathbf{E}\left[G\Big(\big(m^{-n(r)-1} X_{n(r)+1} \big)_{r \in (0,1]} \Big) \right] \, . \tag{27}$$

Observe that $B(\mathbf{u}, r)$ is empty and $n(r) = \infty$ if $r = 0$, that is why the radius r ranges in $(0, 1]$ in (27). Note that $(Y_n \, ; \, n \geq 1)$ is an i.i.d. sequence of random variables; thus, the X_n's have the same distribution. We provide two (rough) bounds of the tail at ∞ of Y_1 and X_1 that are needed in the proof section: recall notation F from the introduction; the first bound is a straightforward consequence of the definitions:

$$\mathbf{P}(X_1 > x) \geq \mathbf{P}(Y_1 > x) \geq C_0 \, \mathbf{P}(W > x) = C_0 \, \exp(-F(x)) \, . \tag{28}$$

where $C_0 := (1 - \widehat{\xi}(1)) = \mathbf{P}(k_{\{\varnothing\}}(\mathscr{T}^*) \geq 2)$ is a positive constant. Next observe that

$$\mathbf{E}\left[\int_{\partial \mathbb{U}} M(d\mathbf{u}) \, G\big(M(B(\mathbf{u}, 1)) \big) \right] = \mathbf{E}\left[\sum_{i=1}^{k_\emptyset(\mathscr{T})} \mathrm{m}^{-1} W_i \, G(\mathrm{m}^{-1} W_i) \right] = \mathbf{E}\left[W G(\mathrm{m}^{-1} W) \right] .$$

Therefore (27) with $r = 1$ entails

$$\mathbf{E}\left[W \, G(\mathrm{m}^{-1} W) \right] = \mathbf{E}\left[G\left(\mathrm{m}^{-2} X_2 \right) \right] = \mathbf{E}\left[G\left(\mathrm{m}^{-2} X_1 \right) \right] \, .$$

Cauchy-Schwarz inequality and the previous identity both imply the following:

$$\mathbf{P}(X_1 > \mathrm{m}\, x) = \mathbf{E}[W \, \mathbf{1}_{\{W > x\}}] \leq C_1 \exp\left(-F(x)/2 \right) \, , \tag{29}$$

where we have set $C_1 := \sqrt{\mathbf{E}\left[W^2 \right]}$, which is finite if F satisfies (13), as noticed in Remark 1.2.

4 Proof of Theorem 1.1.

First observe that (28) implies that $\mathbf{P}(Y_n > F^{-1}(\log n)) \geq C_0 n^{-1}$. Since the Y_n's are independent, the converse of Borel-Cantelli lemma entails that Y_n is larger than $F^{-1}(\log n)$ for infinitely many n almost surely. Now, since $X_n \geq Y_n$, we easily get

$$\mathbf{P}\text{-a.s.} \quad \limsup_{n \to \infty} \frac{\mathrm{m}^{-n} X_n}{g(e^{-n})} \geq 1 \, , \tag{30}$$

where we recall that g is given by (6). Next (29) implies that $\mathbf{P}(X_n > \mathrm{m} F^{-1}(3 \log n)) \leq C_1 \cdot n^{-3/2}$. Then, Borel-Cantelli combined with (12) gives

$$\mathbf{P}\text{-a.s.} \quad \limsup_{n \to \infty} \frac{\mathrm{m}^{-n} X_n}{g(e^{-n})} \leq 6^a \mathrm{m} \, . \tag{31}$$

Kolmogorov 0-1 law combined with (30) and (31) imply that $\limsup_{r\to 0} g(r)^{-1}\mathrm{m}^{-n(r)-1}X_{n(r)+1}$ is a.s. equal to a constant $\kappa = c_\xi^{-1}$ that is *positive* and *finite*, which entails (10) by (27).

Let us prove (11): by Lemma 2.2, we only need to prove that there exists $\kappa_0 > 0$ such that

$$\mathscr{H}_g\left(\left\{\mathbf{u}\in\partial\mathscr{T}\ :\quad \limsup_{r\to 0} g(r)^{-1}M(B(\mathbf{u},r)) < \kappa_0\right\}\right) = 0\,. \tag{32}$$

Recall that $M(B(\mathbf{u},r)) = \mathrm{m}^{-n(r)}W_{\mathbf{u}_{|n(r)}}$. Thus, if we set

$$E_{n_0} = \left\{\mathbf{u}\in\partial\mathscr{T}\ :\ \forall n\geq n_0\ ,\ W_{\mathbf{u}_{|n}} < \mathrm{m}^{-1}\,F^{-1}(\tfrac{1}{2}\log n)\right\}$$

then, an easy argument using (12) entails that Claim (32) is true with $\kappa_0 = 4^{-a}\mathrm{m}^{-1}$ as soon as almost surely the following holds:

$$\forall n_0\geq 2\,,\quad \mathscr{H}_g\left(E_{n_0}\right) = 0\,. \tag{33}$$

Let us prove (33): to that end, we set for any $N > n_0$:

$$J_{n_0,N} = \left\{v\in\mathscr{T}\ :\ |v| = N\ ;\ \forall n\in\{n_0,\dots,N\}\ ,\ W_{v_{|n}} < \mathrm{m}^{-1}\,F^{-1}(\tfrac{1}{2}\log n)\ \right\}.$$

Thus, $\{B_v\ ;\ v\in J_{n_0,N}\}$ is a (e^{-N})-cover of E_{n_0} and $\mathscr{H}_g^{(e^{-N})}\left(E_{n_0}\right)\leq g(e^{-N})\#J_{n_0,N}$. We now estimate $g(e^{-N})\#J_{n_0,N}$. First observe that

$$\mathbf{E}\left[g(e^{-N})\#J_{n_0,N}\right] = g(e^{-N})\mathbf{E}\left[\sum_{u\in\mathscr{T}:|u|=N}\mathbf{1}_{\{\forall n\in\{n_0,\dots,N\}\ ,\ W_{u_{|n}}<\mathrm{m}^{-1}\,F^{-1}(\frac{1}{2}\log n)\}}\right]. \tag{34}$$

Let us now compute the right hand side of the last display thanks to (24): (18) entails that a.s. for any $u\in\mathscr{T}$ such that $|u| = N$, and for any $n\in\{n_0,\dots,N\}$, one has $W_{u_{|n}} = G_n(\theta_u\mathscr{T}) + K_n(\mathrm{Cut}_u(\mathscr{T}))$, where have set $G_n(\theta_u\mathscr{T}) = \mathrm{m}^{-(N-n)}W_u$ and

$$K_n(\mathrm{Cut}_u(\mathscr{T})) = \sum_{p=n+1}^{N}\mathrm{m}^{-(p-n)}\sum_{\substack{v\in\mathscr{T}\,:\,|v|=p\\ v\neq u_{|p}\,,\,v_{|p-1}=u_{|p-1}}} W_v\,.$$

Then, (34) and (24) entail

$$\mathbf{E}\left[g(e^{-N})\#J_{n_0,N}\right] = F^{-1}(\log N)\,\mathbf{P}\left(\forall n\in\{n_0,\dots,N\}\ ,\ X_n^* < \mathrm{m}^{-1}\,F^{-1}(\tfrac{1}{2}\log n)\right)\,,$$

where $X_n^* = \mathbf{m}^{-(N-n)}W' + \sum_{p=n+1}^{N} m^{-(p-n)}Y_p$ and where W' is distributed as W and is independent of the Y_n's. Recall that the Y_n's are i.i.d. and observe that $X_n^*\geq \mathrm{m}^{-1}Y_{n+1}$ for any $n_0\leq n < N$. Therefore, (28) implies

$$\begin{aligned}
\mathbf{P}\left(\forall n\in\{n_0,\dots,N\}\ ,\ X_n^* < \mathrm{m}^{-1}F^{-1}(\tfrac{1}{2}\log n)\right) &\leq \prod_{n=n_0}^{N-1}\mathbf{P}\left(Y_{n+1} < F^{-1}(\tfrac{1}{2}\log n)\right)\\
&\leq \prod_{n=n_0}^{N-1}\left(1 - C_0 n^{-1/2}\right)\\
&\leq \exp\left(-\tfrac{1}{2}C_0\left(\sqrt{N}-\sqrt{n_0}\right)\right)\,.
\end{aligned}$$

Since (12) implies $F^{-1}(\log N)\leq F^{-1}(1)(2\log N)^a$, we easily get $\lim_{N\to\infty} g(e^{-N})\#J_{n_0,N} = 0$, which completes the proof of (33) and thus of (11). ■

References

[1] Aldous, D., and Pitman, J. Tree-valued Markov chains derived from Galton-Watson processes. *Ann. Inst. H. Poincaré. 34* (1998), 637–686. MR1641670

[2] Aldous, D. J. Asymptotic fringe distributions for general families of random trees. *Ann. Appl. Probab. 1*, 2 (1991), 228–266. MR1102319

[3] Athreya, K., and Ney, P. *Branching process*. No. 196 in Grundlehren der Mathematischen Wissenschaften. Springer, 1972. MR0373040

[4] Bingham, N. H., and Doney, R. A. Asymptotic properties of super-critical branching processes. I: The Galton-Watson process. *Adv. Appl. Prob. 6* (1974), 711–731. MR0362525

[5] Duquesne, T. Continuum random trees and branching processes with immigration. *Stoch. Proc and Appl. 119*, 1 (2009), 99–129. MR2485021

[6] Edgar, G. Centered densities and fractal measures. *New York J. Math. 13* (2007), 33–87. MR2288081

[7] Grimmett, G. R. Random labelled trees and their branching networks. *J. Austral. Math. Soc. (Ser. A) 30* (1980), 229–237. MR0607933

[8] Hawkes, J. Trees generated by a simple branching process. *J. London. Math. Soc. 24* (1981), 373–384. MR0631950

[9] Holmes, R. A. Local asymptotic law and the exact Hausdorff measure for a simple branching process. *Proc. London Math. Soc. 26-(3)* (1973), 577–604. MR0326853

[10] Kesten, H., and Stigum, B. A limit theorem for multidimensional Galton-Watson processes. *Ann. Math. Statist. 37* (1966), 1211–1223. MR0198552

[11] Liu, Q. The exact Hausdorff dimension of a branching set. *Prob. Th. Rel.Fields 104* (1996), 515–538. MR1384044

[12] Liu, Q. Exact packing measure of a Galton-Watson tree. *Stoch. Proc. Appl. 85* (2000), 19–28. MR1730621

[13] Liu, Q. Local dimension of the branching measure on a Galton-Watson tree. *Ann. Inst. H. Poincaré, Statist. 37* (2001), 195–222. MR1819123

[14] Lyons, R. Random walks and percolation on trees. *Ann. Probab. 18* (1990), 931–958. MR1062053

[15] Lyons, R., Pemantle, R., and Peres, Y. Conceptual proof of $L\log L$ criteria for mean behavior of branching processes. *Ann. Probab. 23* (1995), 1125–1138. MR1349164

[16] Morters, P., and Shieh, N. R. Thin and thick points for the branching measure on a Galton-Watson tree. *Statist. Probab. Lett. 58* (2002), 13–22. MR1900336

[17] Morters, P., and Shieh, N. R. On the multifractal spectrum of the branching measure of a Galton-Watson tree. *J. Appl. Probab. 41* (2004), 1223–1229. MR2122818

[18] Neveu, J. Arbres et processus de Galton-Watson. *Ann. Inst. H. Poincaré 26* (1986), 199–207. MR0850756

[19] Rogers, C., and Taylor, S. Functions continuous and singular with respect to Hausdorff measures. *Mathematika 8* (1961), 1–31. MR0130336

[20] Watanabe, T. Exact packing measure on the boundary of a Galton-Watson tree. *J. London Math. Soc. 69*, 2 (2004), 801–816. MR2050047

[21] Watanabe, T. Exact Hausdorff measure on the boundary of a Galton-Watson tree. *Ann. Probab. 35*, 3 (2007), 1007–1038. MR2319714