Itô's excursion theory and random trees

Jean-François Le Gall

January 3, 2010

Abstract

We explain how Itô's excursion theory can be used to understand the asymptotic behavior of large random trees. We provide precise statements showing that the rescaled contour of a large Galton-Watson tree is asymptotically distributed according to the Itô excursion measure. As an application, we provide a simple derivation of Aldous' theorem stating that the rescaled contour function of a Galton-Watson tree conditioned to have a fixed large progeny converges to a normalized Brownian excursion. We also establish a similar result for a Galton-Watson tree conditioned to have a fixed large height.

Keywords and phrases. Itô excursion theory, Brownian excursion, random tree, Galton-Watson tree.
Mathematics Subject Classification 2010. 60J65, 60J80, 60C05.

1 Introduction

In his celebrated papers [14, 15], Kiyosi Itô introduced the Poisson point process of excursions of a Markov process from a regular point. The idea of considering excursions away from a point had appeared earlier, in particular in the case of linear Brownian motion (see e.g. Chapter VI in Lévy [20]), but Itô's major breakthrough was to observe that the full collection of excursions could be represented as a single Poisson point process on an appropriate space. In the discrete setting, it is very easy to verify that the successive excursions of a Markov chain away from a recurrent point are independent and identically distributed, and this property plays a major role in the analysis of discrete-time Markov chains. In the continuous setting, it is no longer possible to enumerate the successive excursions from a regular point in chronological order (even though, as we shall see in Section 2 below, excursions can be labelled by the value of the local time at their initial time). Therefore, one cannot immediately make sense of the assertion that the different excursions are independent. The correct point of view, which turns out to be extremely powerful, is to consider Itô's point process of excursions.

Itô's excursion theory has many important applications. It can be used to derive lots of explicit distributions for Brownian motion, and often it sheds light on formulas that were initially proved by different methods. See in particular [24] for a recent account of these applications. More generally, Itô's excursion theory is a fundamental tool in the analysis of Lévy processes on the line: See the monograph [5]. Itô's initial motivation was to understand the sample path behaviour of linear diffusions satisfying general Feller's boundary conditions. This approach can in fact be extended to multidimensional diffusion processes in a domain: See [29]. Excursions of multidimensional Brownian motion in a domain are discussed in the monograph [7] in particular.

In the present work, we discuss a perhaps more unexpected application of Itô's excursion theory to the asymptotic properties of large random trees. Connections between branching processes or trees and random walks have been studied for a long time: See in particular Pitman's monograph [23] for extensive references about these connections. Around 1990, both from Aldous' construction of the CRT as the tree coded by a normalized Brownian excursion [1, 2] and from the Brownian snake approach to superprocesses [17], it became clear that Brownian excursions could be used to define and study the continuous random trees that arise as scaling limits of large Galton-Watson trees.

Our main purpose here is to demonstrate that the Itô excursion measure, which corresponds to the intensity of the Poisson point process of excursions, is indeed the fundamental object if one wants to understand asymptotics of (critical, finite variance) Galton-Watson trees conditioned to be large in some sense. We deal only with the case where the offspring distribution is critical with finite variance, although our techniques can be adapted to more general cases (see Chapter 2 of the monograph [9]). Note that there are many different ways of conditioning a Galton-Watson tree to be large, and that the resulting scaling limits will typically be different. Still Theorem 5.1 below strongly suggests that these scaling limits can always be described by a suitable conditioning of the Itô excursion measure.

As a key tool for establishing Theorem 5.1 and the subsequent results, we start by studying scaling limits for a sequence of independent Galton-Watson trees. This sequence is conveniently coded by a contour, or depth-first search process (see Fig.1 in Section 3 below). Under Brownian rescaling, the contour process of the sequence converges in distribution towards reflected Brownian motion (Theorem 4.1). One can then observe that a Galton-Watson tree conditioned to be large (for instance to have height, or total progeny, greater than some large integer p) can be obtained as the first tree in our sequence that satisfies the desired constraint. In terms of the contour process, or rather of the variant of the contour process called the height process, this tree will correspond to the first excursion away from zero that satisfies an appropriate property (for instance having height, or length, greater than p). Under rescaling, this excursion will converge towards the first excursion of reflected Brownian motion satisfying a corresponding property, and Itô's theory tells us that the law of the latter excursion is a certain conditioning of the Itô measure. The preceding arguments are made rigorous in the proof of Theorem 5.1. Note that, although we formulate our limit theorems in terms of contour processes, these results immediately imply analogous limit theorems for the trees viewed as compact metric spaces, in the sense of the Gromov-Hausdorff distance: The point is that a (discrete or continuous) tree can be viewed as a function of its contour process, with a function that is continuous, and even Lipschitz, for the Gromov-Hausdorff distance (see Section 2 of [19], and the discussion at the end of Section 5 of the present work).

Theorem 5.1 can be applied to handle various conditionings of Galton-Watson trees, but it cannot immediately be used to understand "degenerate conditionings", for which the condition imposed on the contour process excursion will lead to a set that is negligible for the Itô excursion measure. In Sections 6 and 7, we show that such cases can still be treated by a more careful analysis. Section 6 is devoted to a new proof of Aldous' theorem [2] concerning the case of Galton-Watson trees conditioned to have a fixed large progeny (then the limit is a normalized Brownian excursion). Section 7 deals with the case of Galton-Watson trees conditioned to have a fixed large height. In that case the scaling limit of the contour process is a Brownian excursion conditioned to have a given height (Theorem 7.1). In both cases, the idea is to study the absolute continuity properties of the law of the tree under the given degenerate conditioning, with respect to its law under a suitable non-degenerate conditioning, and then to use Theorem 5.1 to handle the latter.

Let us briefly discuss the connections of the present article with earlier work. Theorem 4.1 can be viewed as a special case of the more general results presented in Chapter 2 of [9], with certain simplifications in particular in the proof of tightness. Theorem 5.1 is new, even though similar ideas have appeared earlier in [9] and [19]. As mentioned above, Theorem 6.1 is due to Aldous [2]. Our proof is close in spirit to the one provided by Marckert and Mokkadem [21] under the assumption that the offspring distribution has exponential moments (see also Section 1.5 in [19]). To the best of our knowledge, Theorem 7.1 is new, but certain specific aspects of Galton-Watson trees conditioned to have a fixed large height can be found in [13] and [16]. As a general remark, Marc Yor has pointed out that our strategy of understanding the scaling limits of random trees under various conditionings in terms of a single σ-finite measure (namely the Itô measure) is closely related to certain aspects of the study of penalizations of Brownian motion: See in particular the monograph [26].

To conclude this introduction, let us mention that there are other significant applications of Itô's excursion theory in the setting of spatial branching processes. In particular, the excursion measure of the Brownian snake plays a crucial role in the study of this path-valued Markov process and of its connections with semilinear partial differential equations (see [18] and the references therein). In a related direction, the so-called Integrated Super-Brownian Excursion or ISE [3], which has found striking applications in asymptotics of models of statistical physics, is constructed by combining the tree structure of a Brownian excursion with Brownian displacements in space.

The paper is organized as follows. In Section 2, we recall some key results from Itô's excursion theory in the special case of reflected Brownian motion, which is relevant for our applications. In Section 3, we discuss random trees and their coding functions. In particular, we introduce the height function of a tree, which is often more tractable than the contour function thanks to its simple connection with the so-called Lukasiewicz path (formula (3) below). Section 4 is devoted to the proof of Theorem 4.1 describing the scaling limit of the contour and height processes of a sequence of independent Galton-Watson trees. In Section 5 we establish our main result (Theorem 5.1) connecting the asymptotic behavior of large Galton-Watson trees with conditionings of the Itô excursion measure. Section 6 is devoted to the proof of Aldous' theorem, and Section 7 treats Galton-Watson trees conditioned to have a fixed height.

Notation. If I is a closed subinterval of $\mathbb{R}_+$, $C(I, \mathbb{R}_+)$ stands for the space of all continuous functions from I into $\mathbb{R}_+$, which is equipped with the topology of uniform convergence on every compact subset of I. Similarly, $\mathbb{D}(I, \mathbb{R}_+)$ denotes the space of all càdlàg functions from I into $\mathbb{R}_+$, which is equipped with the Skorokhod topology.

Acknowledgements. I thank David Aldous and Jean-François Marckert for useful comments and references. I am also indebted to Marc Yor, Maria Eulalia Vares and Igor Kortchemski for their remarks on the first version of this work.

2 Excursion theory of Brownian motion

In this section, we briefly recall the key facts of Itô's excursion theory in the particular case of reflected Brownian motion. We especially collect those facts that will be needed in our applications to random trees. A much more detailed account of the theory can be found in the paper [29]. The results that are recalled below can also be found in Chapter XII of the book [25].

We consider a standard linear Brownian motion $B = (B_t)_{t \geq 0}$ starting from the origin. The process $\beta_t = |B_t|$ is called reflected Brownian motion. We denote by $(L^0_t)_{t \geq 0}$ the local time process of B (or of β) at level 0, which can be defined by the approximation

$$L^0_t = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_0^t ds\, \mathbf{1}_{[-\varepsilon,\varepsilon]}(B_s) = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_0^t ds\, \mathbf{1}_{[0,\varepsilon]}(\beta_s),$$

for every $t \geq 0$, a.s.

Then $(L^0_t)_{t\geq 0}$ is a continuous increasing process, and the set of increase points of the function $t \to L^0_t$ coincides with the set

$$\mathcal{Z} = \{t \geq 0 : \beta_t = 0\}$$

of all zeros of β. Consequently, if we introduce the right-continuous inverse of the local time process,

$$\tau_\ell := \inf\{t \geq 0 : L^0_t > \ell\} \ , \quad \text{for every } \ell \geq 0,$$

we have

$$\mathcal{Z} = \{\tau_\ell : \ell \geq 0\} \cup \{\tau_{\ell-} : \ell \in D\}$$

where D denotes the countable set of all discontinuity times of the mapping $\ell \to \tau_\ell$.

We call *excursion interval* of β (away from 0) any connected component of the open set $\mathbb{R}_+\backslash\mathcal{Z}$. Excursion intervals for more general continuous random functions will be defined in a similar way. The preceding discussion shows that, with probability one, the excursion intervals of β away from 0 are exactly the intervals $]\tau_{\ell-}, \tau_\ell[$ for $\ell \in D$. Then, for every $\ell \in D$, we define the excursion $e_\ell = (e_\ell(t))_{t\geq 0}$ associated with the interval $]\tau_{\ell-}, \tau_\ell[$ by setting

$$e_\ell(t) = \begin{cases} \beta_{\tau_{\ell-}+t} & \text{if } 0 \leq t \leq \tau_\ell - \tau_{\ell-}\ , \\ 0 & \text{if } t > \tau_\ell - \tau_{\ell-}\ . \end{cases}$$

We view e_ℓ as an element of the excursion space E, which is defined by

$$E = \{e \in C(\mathbb{R}_+, \mathbb{R}_+) : e(0) = 0 \text{ and } \zeta(e) := \sup\{s > 0 : e(s) > 0\} \in]0, \infty[\},$$

where $\sup \varnothing = 0$ by convention. Note that we require $\zeta(e) > 0$, so that the zero function does not belong to E. For technical reasons, we do not require that $e(t) > 0$ for $t \in]0, \zeta(e)[$, although of course the measure $n(de)$ is suppported on the functions that satisfy this property. The space E is equipped with the metric d defined by

$$d(e, e') = \sup_{t\geq 0} |e(t) - e'(t)| + |\zeta(e) - \zeta(e')|$$

and with the associated Borel σ-field. Notice that $\zeta(e_\ell) = \tau_\ell - \tau_{\ell-}$ for every $\ell \in D$.

We can now state the basic theorem of excursion theory in our particular setting. This is a special case of Itô's results in [14, 15] (see also [22]).

Theorem 2.1 *The point measure*

$$\sum_{\ell\in D} \delta_{(\ell, e_\ell)}(ds\, de)$$

is a Poisson measure on $\mathbb{R}_+ \times E$, with intensity

$$ds \otimes \mathbf{n}(de)$$

where $\mathbf{n}(de)$ is a σ-finite measure on E.

The measure $\mathbf{n}(de)$ is called the Itô measure of positive excursions of linear Brownian motion, or simply the Itô excursion measure. Notice that we restrict our attention to positive excursions: The measure $\mathbf{n}$ is denoted by n_+ in Chapter XII of [25].

The next corollary, which follows from standard properties of Poisson measures, will be one of our main tools.

Corollary 2.2 *Let A be a measurable subset of E such that $0 < \mathbf{n}(A) < \infty$, and let $T_A = \inf\{\ell \in D : e_\ell \in A\}$. Then, T_A is exponentially distributed with parameter $\mathbf{n}(A)$, and the distribution of e_{T_A} is the conditional measure*

$$\mathbf{n}(\cdot \,|\, A) = \frac{\mathbf{n}(\cdot \cap A)}{\mathbf{n}(A)}.$$

Moreover, T_A and e_{T_A} are independent.

This corollary can be used to calculate various distributions under the Itô excursion measure. We will use the distribution of the height and the length of the excursion, which are given as follows: For every $\varepsilon > 0$,

$$\mathbf{n}\Big(\max_{t\geq 0} e(t) > \varepsilon\Big) = \frac{1}{2\varepsilon}$$

and

$$\mathbf{n}(\zeta(e) > \varepsilon) = \frac{1}{\sqrt{2\pi\varepsilon}}.$$

The Itô excursion measure enjoys the following scaling property. For every $\lambda > 0$, define a mapping $\Phi_\lambda : E \longrightarrow E$ by setting $\Phi_\lambda(e)(t) = \sqrt{\lambda}\, e(t/\lambda)$, for every $e \in E$ and $t \geq 0$. Then we have $\Phi_\lambda(\mathbf{n}) = \sqrt{\lambda}\, \mathbf{n}$.

This scaling property makes it possible to define conditional versions of the Itô excursion measure. We discuss the conditioning of $\mathbf{n}(de)$ with respect to the length $\zeta(e)$. There exists a unique collection of probability measures $(\mathbf{n}_{(s)}, s > 0)$ on E such that the following properties hold:

(i) For every $s > 0$, $\mathbf{n}_{(s)}(\zeta = s) = 1$.

(ii) For every $\lambda > 0$ and $s > 0$, we have $\Phi_\lambda(\mathbf{n}_{(s)}) = \mathbf{n}_{(\lambda s)}$.

(iii) For every measurable subset A of E,

$$\mathbf{n}(A) = \int_0^\infty \mathbf{n}_{(s)}(A)\, \frac{ds}{2\sqrt{2\pi s^3}}.$$

We may and will write $\mathbf{n}_{(s)} = \mathbf{n}(\cdot \,|\, \zeta = s)$. The measure $\mathbf{n}_{(1)} = \mathbf{n}(\cdot \,|\, \zeta = 1)$ is called the law of the normalized Brownian excursion. Similarly, one can define conditionings with respect to the height of the excursion, and make sense of $\mathbf{n}(de \,|\, \max(e) = x)$ for every $x > 0$.

There are many different descriptions of the Itô excursion measure: See in particular [25, Chapter XII]. We state the following proposition, which emphasizes the Markovian properties of $\mathbf{n}$. For every $t > 0$ and $x > 0$, we set

$$q_t(x) = \frac{x}{\sqrt{2\pi t^3}} \exp(-\frac{x^2}{2t}).$$

Note that the function $t \to q_t(x)$ is the density of the first hitting time of x by B.

Proposition 2.3 *The Itô excursion measure $\mathbf{n}$ is the only σ-finite measure on E that satisfies the following two properties.*

(i) *For every $t > 0$, and every $f \in C(\mathbb{R}_+, \mathbb{R}_+)$,*

$$\mathbf{n}(f(e(t))\, \mathbf{1}_{\{\zeta > t\}}) = \int_0^\infty f(x)\, q_t(x)\, dx.$$

(ii) *Let $t>0$. Under the conditional probability measure $\mathbf{n}(\cdot\,|\,\zeta>t)$, the process $(e(t+r))_{r\geq 0}$ is Markov with the transition kernels of Brownian motion stopped upon hitting* 0.

This proposition can be used to establish absolute continuity properties of the conditional measures $\mathbf{n}_{(s)}$ with respect to $\mathbf{n}$. We will need the following fact. For every $t\geq 0$, let $\mathcal{F}_t$ denote the σ-field on E generated by the mappings $r\longrightarrow e(r)$, for $0\leq r\leq t$. Then, if $0<t<1$, the measure $\mathbf{n}_{(1)}$ is absolutely continuous with respect to $\mathbf{n}$ on the σ-field $\mathcal{F}_t$, with Radon-Nikodym density

$$\frac{d\mathbf{n}_{(1)}}{d\mathbf{n}}\Big|_{\mathcal{F}_t}(e)=2\sqrt{2\pi}\,q_{1-t}(e(t)). \tag{1}$$

This formula is indeed similar to the classical formula relating the law up to time $t<1$ of a "bridge" of duration 1 (associated with a Markov process) to the law of the unconditioned process.

In the same way, we can compute the Radon-Nikodym derivative of the conditional measure $\mathbf{n}(\cdot\,|\,\zeta>1)$ with respect to $\mathbf{n}$ on the same σ-field:

$$\frac{d\mathbf{n}(\cdot\,|\,\zeta>1)}{d\mathbf{n}}\Big|_{\mathcal{F}_t}(e)=\sqrt{2\pi}\int_{1-t}^{\infty}q_r(e(t))\,dr. \tag{2}$$

These formulas will be useful in Section 6 below.

3 Random trees

3.1 Rooted ordered trees

We will primarily be interested in rooted ordered trees, also called plane trees, although our results have applications to other classes of trees. Let us start with some definitions.

We first introduce the set of labels

$$\mathcal{U}=\bigcup_{n=0}^{\infty}\mathbb{N}^n$$

where $\mathbb{N}=\{1,2,\ldots\}$ and by convention $\mathbb{N}^0=\{\varnothing\}$. An element of $\mathcal{U}$ is thus a finite sequence $u=(u^1,\ldots,u^n)$ of positive integers, and we set $|u|=n$. We say that $|u|$ is the *generation* of u. Thus $\mathbb{N}^n$ consists of all elements at the nth generation.

A (finite) rooted ordered tree $\mathbf{t}$ is a finite subset of $\mathcal{U}$ such that:

(i) $\varnothing\in\mathbf{t}$.

(ii) If $u=(u^1,\ldots,u^n)\in\mathbf{t}\backslash\{\varnothing\}$ then $(u^1,\ldots,u^{n-1})\in\mathbf{t}$.

(iii) For every $u=(u^1,\ldots,u^n)\in\mathbf{t}$, there exists an integer $k_u(\mathbf{t})\geq 0$ such that, for every $j\in\mathbb{N}$, $(u^1,\ldots,u^n,j)\in\mathbf{t}$ if and only if $1\leq j\leq k_u(\mathbf{t})$

The number $k_u(\mathbf{t})$ is interpreted as the "number of children" of u in $\mathbf{t}$.

We denote the set of all rooted ordered trees by $\mathbb{T}$. In what follows, we see each element of the tree $\mathbf{t}$ as an individual of a population whose $\mathbf{t}$ is the family tree. By definition, the cardinality $\#\mathbf{t}$ of $\mathbf{t}$ is the total progeny of $\#\mathbf{t}$.

We now explain how trees can be coded by discrete functions (see Fig.1 below). We start by introducing the *contour function*, or Dyck path in the terminology of [28].

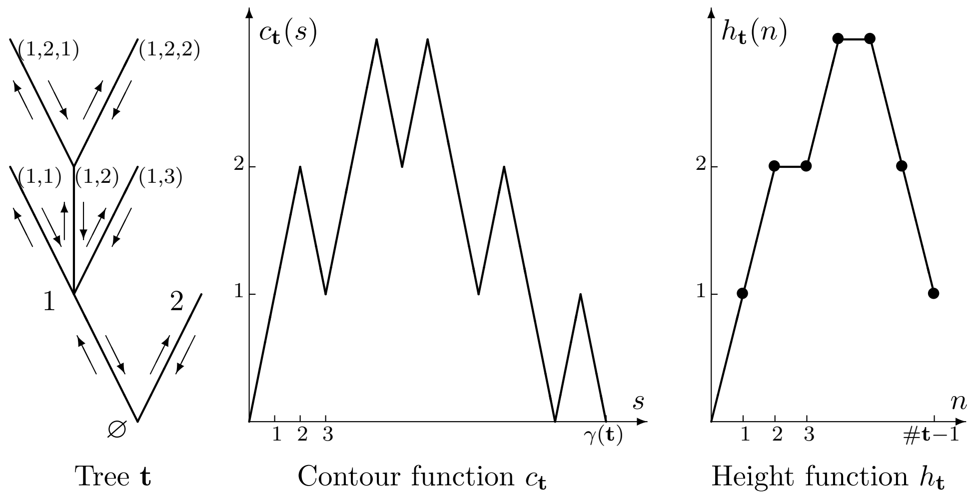

Figure 1

Suppose that the tree $\mathbf{t}$ is embedded in the half-plane as shown on Fig.1, in such a way that edges have length one. Informally, we imagine the motion of a particle that starts at time $s = 0$ from the root of the tree and then explores the tree from the left to the right, moving continuously along the edges at unit speed (in the way explained by the arrows of Fig.1), until all edges have been explored and the particle has come back to the root. Since it is clear that each edge will be crossed twice in this evolution, the total time needed to explore the tree is $\gamma(\mathbf{t}) := 2(\#\mathbf{t} - 1)$. The contour function of $\mathbf{t}$ is the function $(c_{\mathbf{t}}(s))_{0\leq s\leq\gamma(\mathbf{t})}$ whose value at time $s \in [0, \gamma(\mathbf{t})]$ is the distance (on the tree) between the position of the particle at time s and the root. Fig.1 explains the construction of the contour function better than a formal definition. Clearly, $\mathbf{t}$ is determined by its contour function. It will often be convenient to define $c_{\mathbf{t}}(s)$ for every $s \geq 0$, by setting $c_{\mathbf{t}}(s) = 0$ for $s > \gamma(\mathbf{t})$. Then $c_{\mathbf{t}}$ becomes an element of E and $\zeta(c_{\mathbf{t}}) = \gamma(\mathbf{t})$.

The *height function* gives another way of coding the tree $\mathbf{t}$. We denote the elements of $\mathbf{t}$ listed in lexicographical order by $u_0 = \varnothing, u_1, u_2, \ldots, u_{\#\mathbf{t}-1}$. The height function $(h_{\mathbf{t}}(n); 0 \leq n < \#\mathbf{t})$ is defined by

$$h_{\mathbf{t}}(n) = |u_n|, \quad 0 \leq n < \#\mathbf{t}.$$

The height function is thus the sequence of the generations of the individuals of $\mathbf{t}$, when these individuals are listed in lexicographical order (see Fig.1 for an example). It is again easy to check that $h_{\mathbf{t}}$ characterizes the tree $\mathbf{t}$.

We finally introduce the *Lukasiewicz path* of the tree $\mathbf{t}$. This is the finite sequence of integers $(x_0, x_1, \ldots, x_{\#\mathbf{t}})$ determined by the relations $x_0 = 0$ and $x_{j+1} - x_j = k_{u_j}(\mathbf{t}) - 1$ for every $j = 0, 1, \ldots, \#\mathbf{t} - 1$. An easy combinatorial argument shows that $x_j \geq 0$ for $0 \leq j < \#\mathbf{t}$ and $x_{\#\mathbf{t}} = -1$. The height function $h_{\mathbf{t}}$ of the tree $\mathbf{t}$ is related to its Lukasiewicz path by the formula

$$h_{\mathbf{t}}(n) = \#\{j \in \{0, 1, \ldots, n-1\} : x_j = \min_{j\leq\ell\leq n} x_\ell\}, \tag{3}$$

for every $n = 0, 1, \ldots, \#\mathbf{t} - 1$. See e.g. [19, Proposition 1.2] for a proof.

3.2 Galton-Watson trees

Let μ be an offspring distribution, that is a probability measure $(\mu(k), k = 0, 1, \ldots)$ on the nonnegative integers. Throughout this work, we make the following two basic assumptions:

(i) (critical branching) $\displaystyle\sum_{k=0}^{\infty} k\,\mu(k) = 1.$

(ii) (finite variance) $\displaystyle\sum_{k=0}^{\infty} k^2\,\mu(k) < \infty.$

The variance of μ is denoted by $\sigma^2 = \sum_{k=0}^{\infty} k^2\,\mu(k) - 1$. We exclude the trivial case where μ is the Dirac measure at 1, and thus $\sigma > 0$. We also introduce the probability measure ν on $\mathbb{Z}$ defined by

$$\nu(k) = \mu(k+1)$$

for every $k = -1, 0, 1, \ldots$, and $\nu(k) = 0$ if $k < -1$. Notice that ν has zero mean.

The μ-Galton-Watson tree is the genealogical tree of a Galton-Watson branching process with offspring distribution μ starting with a single individual called the ancestor. It thus corresponds to the evolution of a population where each individual has, independently of the others, a random number of children distributed according to μ. Under our assumptions on μ, the population becomes extinct after a finite number of generations, and so the genealogical tree is finite a.s.

More rigorously, a μ-Galton-Watson tree is a random variable $\mathcal{T}$ with values in $\mathbb{T}$, whose distribution is given by

$$P[\mathcal{T} = \mathbf{t}] = \prod_{u \in \mathbf{t}} \mu(k_u(\mathbf{t})) \tag{4}$$

for every $\mathbf{t} \in \mathbb{T}$.

It turns out that the Lukasiewicz path of a μ-Galton-Watson has a simple probabilistic structure.

Lemma 3.1 *Let $(S_n, n \geq 0)$ be a random walk on $\mathbb{Z}$ with initial value $S_0 = 0$ and jump distribution ν. Set*

$$T = \inf\{n \geq 0 : S_n = -1\}.$$

Then the Lukasiewicz path of a μ-Galton-Watson tree $\mathcal{T}$ has the same distribution as $(S_0, S_1, \ldots, S_T)$. In particular, $\#\mathcal{T}$ and T have the same distribution.

The statement of the lemma is intuitively "obvious": If $u_0, u_1, \ldots$ are the individuals of $\mathcal{T}$ listed in lexicographical order, just use the fact that the numbers of children $k_{u_0}(\mathcal{T}), k_{u_1}(\mathcal{T}), \ldots$ are independent and distributed according to μ. See e.g. [19, Corollary 1.6] for a detailed proof.

We denote the height of the tree $\mathcal{T}$ by $\mathrm{ht}(\mathcal{T}) = \sup\{|u| : u \in \mathcal{T}\}$. Then a classical result of the theory of branching processes (see Theorem 1 in [4, p.19]) states that

$$P[\mathrm{ht}(\mathcal{T}) \geq p] \underset{p\to\infty}{\sim} \frac{2}{\sigma^2 p}, \tag{5}$$

where we use the notation $a_p \underset{p\to\infty}{\sim} b_p$ to mean that the ratio a_p/b_p tends to 1 as $p \to \infty$.

Notice that the function $p \to P[\mathrm{ht}(\mathcal{T}) = p]$ is nonincreasing: To see this, note that for the tree $\mathcal{T}$ to have height $p+1$, it is necessary that the subtree of descendants of one of the children of the ancestor has height p, and use the criticality of the offspring distribution. It follows that we have

$$P[\mathrm{ht}(\mathcal{T}) = p] \underset{p\to\infty}{\sim} \frac{2}{\sigma^2 p^2}. \tag{6}$$

This estimate also follows from Corollary 1 in [4, p.23].

Remark. In the next sections, we discuss asymptotics for Galton-Watson trees conditioned to be large. Up to some point, our results can be extended to offspring distributions that do not satisfy (i) and (ii). Infinite variance offspring distributions are discussed in Chapter 2 of the monograph [9]. In the case of non-critical offspring distributions, if one is interested in the tree conditioned to have a fixed large progeny, it is often possible to apply the results of the critical case, by replacing μ with $\mu_a(k) = C_a\, a^k\, \mu(k)$, where C_a is a normalizing constant and the parameter $a > 0$ is chosen so that μ_a becomes critical.

4 Limit theorems for sequences of independent Galton-Watson trees

Consider a sequence $\mathcal{T}_0, \mathcal{T}_1, \mathcal{T}_2, \ldots$ of independent μ-Galton-Watson trees. The contour process $(C_t)_{t\in\mathbb{R}_+}$ of this sequence is the random process obtained by concatenating the contour functions $(c_{\mathcal{T}_0}(s))_{0\le s\le\gamma(\mathcal{T}_0)}$, $(c_{\mathcal{T}_1}(s))_{0\le s\le\gamma(\mathcal{T}_1)}$,... of the trees $\mathcal{T}_0, \mathcal{T}_1$,... (we may in fact consider only those trees $\mathcal{T}_j$ such that $\#\mathcal{T}_j > 1$, since otherwise the contour function is trivial). Similarly, the height process $(H_n)_{n\in\mathbb{Z}_+}$ of the sequence is obtained by concatenating the (discrete) height functions $(h_{\mathcal{T}_0}(n))_{0\le n\le\#\mathcal{T}_0-1}$, $(h_{\mathcal{T}_1}(n))_{0\le n\le\#\mathcal{T}_1-1}$,... of the trees $\mathcal{T}_0, \mathcal{T}_1$,... Notice that $H_{n+1} - H_n \le 1$ for every $n \ge 0$.

We can now state the main result of this section. If $x \in \mathbb{R}_+$, $[x]$ denotes the integer part of x.

Theorem 4.1 *Let $\mathcal{T}_0, \mathcal{T}_1, \mathcal{T}_2, \ldots$ be a sequence of independent μ-Galton-Watson trees. Let $(H_n)_{n\ge0}$ be the associated height process and let $(C_t)_{t\ge0}$ be the associated contour process. Then*

$$(\frac{1}{\sqrt{p}} H_{[pt]})_{t\ge0} \underset{p\to\infty}{\xrightarrow{\text{(d)}}} (\frac{2}{\sigma}\,\beta_t)_{t\ge0} \tag{7}$$

where β is a reflected Brownian motion started from 0. The convergence holds in the sense of weak convergence of the laws on $\mathbb{D}(\mathbb{R}_+, \mathbb{R}_+)$. Moreover, we have for every $K > 0$,

$$\frac{1}{\sqrt{p}} \sup_{0\le t\le K} |C_{2pt} - H_{[pt]}| \underset{p\to\infty}{\xrightarrow{\text{(P)}}} 0\,, \tag{8}$$

where the notation $\xrightarrow{\text{(P)}}$ refers to convergence in probability. In particular, we have also

$$(\frac{1}{\sqrt{p}} C_{2pt})_{t\ge0} \underset{p\to\infty}{\xrightarrow{\text{(d)}}} (\frac{2}{\sigma}\,\beta_t)_{t\ge0}, \tag{9}$$

in the sense of weak convergence of the laws on $C(\mathbb{R}_+, \mathbb{R}_+)$.

Before we proceed to the proof of Theorem 4.1, we state a couple of lemmas. We first need to introduce an important auxiliary process, which can be viewed as the Lukasiewicz path of the forest $\mathcal{T}_0, \mathcal{T}_1, \ldots$ and which we denote by $(S_n, n \ge 0)$. For every $j = 0, 1, \ldots$, let $(X^j_n, 0 \le n \le \#\mathcal{T}_j - 1)$ be the Lukasiewicz path of the tree $\mathcal{T}_j$, and set $Y^j_n = -j + X^j_n$. Then the random path $(S_n, n \ge 0)$ is just the concatenation of the paths $(Y^j_n, 0 \le n \le \#\mathcal{T}_j - 1)$ for $j = 0, 1, \ldots$

The following simple lemma is a straightforward consequence of the results recalled in the previous section.

Lemma 4.2 *The process* $(S_n)_{n\geq 0}$ *is a random walk with jump distribution* ν *and initial value* $S_0 = 0$. *Moreover, for every* $n \geq 0$,

$$H_n = \#\{k \in \{0,1,\ldots,n-1\} : S_k = \min_{k\leq j\leq n} S_j\}. \tag{10}$$

The fact that S is a random walk with jump distribution ν follows from Lemma 3.1, and formula (10) is an immediate consequence of (3).

As an easy consequence of the representation (10) and the strong Markov property of the random walk S, we get the following lemma, whose proof is left as an exercise for the reader (see [9, Lemma 2.3.5]).

Lemma 4.3 *Let* τ *be a stopping time of the filtration* $(\mathcal{G}_n)$ *generated by the random walk* S. *Then the process*

$$\Big(H_{\tau+n} - \min_{\tau\leq k\leq \tau+n} H_k\Big)_{n\geq 0}$$

is independent of $\mathcal{G}_\tau$ *and has the same distribution as* $(H_n)_{n\geq 0}$.

Proof of Theorem 4.1. The proof of Theorem 4.1 consists of three separate steps. In the first one, we obtain the weak convergence of finite-dimensional marginals of the processes $(\frac{1}{\sqrt{p}}H_{[pt]})_{t\geq 0}$. In the second one, we prove tightness of the laws of these processes, thus establishing the convergence (7). In the last step, we prove the convergence (8).

First step. Let $S = (S_n)_{n\geq 0}$ be as before the Lukasiewicz path of the forest $\mathcal{T}_0, \mathcal{T}_1, \ldots$ Note that the jump distribution ν has mean 0 and finite variance σ^2, and thus the random walk S is recurrent. We also introduce the notation

$$M_n = \max_{0\leq k\leq n} S_k\ , \qquad I_n = \min_{0\leq k\leq n} S_k\ .$$

Donsker's invariance theorem gives

$$(\frac{1}{\sqrt{p}} S_{[pt]})_{t\geq 0} \xrightarrow[p\to\infty]{\text{(d)}} (\sigma\, B_t)_{t\geq 0} \tag{11}$$

where B is as in Section 2 a standard linear Brownian motion started at the origin.

For every fixed $n \geq 0$, introduce the time-reversed random walk $\widehat{S}^n$ defined by $\widehat{S}^n_k = S_n - S_{n-k}$, for $0 \leq k \leq n$. Note that $(\widehat{S}^n_k, 0 \leq k \leq n)$ has the same distribution as $(S_k, 0 \leq k \leq n)$. From formula (10), we have

$$H_n = \#\{k \in \{0,1,\ldots,n-1\} : S_k = \min_{k\leq j\leq n} S_j\} = \Psi_n(\widehat{S}^n_0, \widehat{S}^{(n)}_1, \ldots, \widehat{S}^{(n)}_n),$$

where for any integer sequence $(x_0, x_1, \ldots, x_n)$, we have set

$$\Psi_n(x_0, x_1, \ldots, x_n) = \#\{k \in \{1,\ldots,n\} : x_k = \max_{0\leq j\leq k} x_j\}.$$

We also set $R_n = \Psi_n(S_0, S_1, \ldots, S_n) = \#\{k \in \{1,\ldots,n\} : S_k = M_k\}$.

The following lemma is standard (see e.g. Lemma 1.9 in [19, p.255] for a simple derivation of the law of S_{T_1}).

Lemma 4.4 *Define a sequence of stopping times T_j, $j = 0, 1, \dots$ inductively by setting $T_0 = 0$ and for every $j \geq 1$,*

$$T_j = \inf\{n > T_{j-1} : S_n = M_n\}.$$

Then the random variables $S_{T_j} - S_{T_{j-1}}$, $j = 1, 2, \dots$ are independent and identically distributed, with distribution $P[S_{T_1} = k] = \nu([k, \infty[)$ for every $k \geq 0$.

Note that the distribution of S_{T_1} has a finite first moment:

$$E[S_{T_1}] = \sum_{k=0}^{\infty} k\, \nu([k, \infty[) = \sum_{j=0}^{\infty} \frac{j(j+1)}{2} \nu(j) = \frac{\sigma^2}{2}.$$

The next lemma is the key to the first part of the proof.

Lemma 4.5 *We have*

$$\frac{H_n}{S_n - I_n} \xrightarrow[n\to\infty]{\text{(P)}} \frac{2}{\sigma^2}.$$

Proof. From our definitions, we have

$$M_n = \sum_{T_k \leq n} (S_{T_k} - S_{T_{k-1}}) = \sum_{k=1}^{R_n} (S_{T_k} - S_{T_{k-1}}).$$

Using Lemma 4.4 and the law of large numbers, we get

$$\frac{M_n}{R_n} \xrightarrow[n\to\infty]{\text{(a.s.)}} E[S_{T_1}] = \frac{\sigma^2}{2}.$$

By replacing S with the time-reversed walk $\widehat{S}^n$ we see that for every n, the pair (M_n, R_n) has the same distribution as $(S_n - I_n, H_n)$. Hence the previous convergence entails

$$\frac{S_n - I_n}{H_n} \xrightarrow[n\to\infty]{\text{(P)}} \frac{\sigma^2}{2},$$

and the lemma follows. $\square$

From (11), we have for every choice of $0 \leq t_1 \leq t_2 \leq \cdots \leq t_m$,

$$\frac{1}{\sqrt{p}}\Big(S_{[pt_1]} - I_{[pt_1]}, \dots, S_{[pt_m]} - I_{[pt_m]}\Big) \xrightarrow[p\to\infty]{\text{(d)}} \sigma\Big(B_{t_1} - \inf_{0\leq s\leq t_1} B_s, \dots, B_{t_m} - \inf_{0\leq s\leq t_m} B_s\Big).$$

Therefore it follows from Lemma 4.5 that

$$\frac{1}{\sqrt{p}}\Big(H_{[pt_1]}, \dots, H_{[pt_m]}\Big) \xrightarrow[p\to\infty]{\text{(d)}} \frac{2}{\sigma}\Big(B_{t_1} - \inf_{0\leq s\leq t_1} B_s, \dots, B_{t_m} - \inf_{0\leq s\leq t_m} B_s\Big).$$

A famous theorem of Lévy states that the process $\underline{B}_t := B_t - \inf_{0\leq s\leq t} B_s$ is a reflected Brownian motion. This completes the proof of the convergence of finite-dimensional marginals in (7).

Second step. To simplify notation, set

$$H_t^{(p)} = \frac{1}{\sqrt{p}} H_{[pt]}.$$

We need to prove the tightness of the laws of the processes $H^{(p)}$ in the set of all probability measures on the Skorokhod space $\mathbb{D}(\mathbb{R}_+, \mathbb{R}_+)$. By standard results (see e.g. Corollary 3.7.4 in [10]), it is enough to verify that, for every fixed $T > 0$ and $\delta > 0$,

$$\lim_{n\to\infty} \Big(\limsup_{p\to\infty} \; P\Big[\sup_{1\le i\le 2^n} \; \sup_{t\in[(i-1)2^{-n}T, i2^{-n}T]} |H^{(p)}_t - H^{(p)}_{(i-1)2^{-n}T}| > \delta \Big]\Big) = 0. \tag{12}$$

We fix $\delta > 0$ and $T > 0$ and first observe that

$$P\Big[\sup_{1\le i\le 2^n} \; \sup_{t\in[(i-1)2^{-n}T, i2^{-n}T]} |H^{(p)}_t - H^{(p)}_{(i-1)2^{-n}T}| > \delta \Big] \le A_1(n,p) + A_2(n,p) + A_3(n,p)\,, \tag{13}$$

where

$$\begin{aligned}
A_1(n,p) &= P\Big[\sup_{1\le i\le 2^n} |H^{(p)}_{i2^{-n}T} - H^{(p)}_{(i-1)2^{-n}T}| > \frac{\delta}{5} \Big]\,,\\
A_2(n,p) &= P\Big[\sup_{t\in[(i-1)2^{-n}T, i2^{-n}T]} H^{(p)}_t > H^{(p)}_{(i-1)2^{-n}T} + \frac{4\delta}{5}\,,\ \text{for some } 1\le i\le 2^n \Big]\,,\\
A_3(n,p) &= P\Big[\inf_{t\in[(i-1)2^{-n}T, i2^{-n}T]} H^{(p)}_t < H^{(p)}_{i2^{-n}T} - \frac{4\delta}{5}\,,\ \text{for some } 1\le i\le 2^n \Big]\,.
\end{aligned}$$

The term A_1 is easy to bound. By the convergence of finite-dimensional marginals, we have

$$\limsup_{p\to\infty} A_1(n,p) \le P\Big[\sup_{1\le i\le 2^n} \frac{2}{\sigma} |\beta_{i2^{-n}T} - \beta_{(i-1)2^{-n}T}| \ge \frac{\delta}{5} \Big]$$

and the path continuity of the process β ensures that the right-hand side tends to 0 as $n \to \infty$.

To bound the terms A_2 and A_3, we fix $p \ge 1$ and we introduce the stopping times $\tau^{(p)}_k$, $k \ge 0$ defined by induction by $\tau^{(p)}_0 = 0$ and

$$\tau^{(p)}_{k+1} = \inf\{t \ge \tau^{(p)}_k : H^{(p)}_t > \inf_{\tau^{(p)}_k \le r\le t} H^{(p)}_r + \frac{\delta}{5}\}.$$

Let $i \in \{1, \ldots, 2^n\}$ be such that

$$\sup_{t\in[(i-1)2^{-n}T, i2^{-n}T]} H^{(p)}_t > H^{(p)}_{(i-1)2^{-n}T} + \frac{4\delta}{5}. \tag{14}$$

Then at least one of the random times $\tau^{(p)}_k$, $k \ge 0$ must lie in the interval $[(i-1)2^{-n}T, i2^{-n}T]$. Let $\tau^{(p)}_j$ be the first such time. By construction we have

$$\sup_{t\in[(i-1)2^{-n}T, \tau^{(p)}_j[} H^{(p)}_t \le H^{(p)}_{(i-1)2^{-n}T} + \frac{\delta}{5},$$

and since the positive jumps of $H^{(p)}$ are of size $\frac{1}{\sqrt{p}}$, we get also

$$H^{(p)}_{\tau^{(p)}_j} \le H^{(p)}_{(i-1)2^{-n}T} + \frac{\delta}{5} + \frac{1}{\sqrt{p}} < H^{(p)}_{(i-1)2^{-n}T} + \frac{2\delta}{5}$$

provided that $p > (5/\delta)^2$, which we assume from now on. From (14), we have then

$$\sup_{t\in[\tau_j^{(p)},i2^{-n}T]} H_t^{(p)} > H^{(p)}_{\tau_j^{(p)}} + \frac{\delta}{5},$$

which implies that $\tau_{j+1}^{(p)} \leq i2^{-n}T$. Summarizing, we get

$$A_2(n,p) \leq P\Big[\tau_k^{(p)} < T \text{ and } \tau_{k+1}^{(p)} - \tau_k^{(p)} < 2^{-n}T\,,\ \text{for some } k \geq 0\Big]. \tag{15}$$

A similar argument gives exactly the same bound for the quantity $A_3(n,p)$.

In order to bound the right-hand side of (15), we will use the next lemma.

Lemma 4.6 *The random variables $\tau_{k+1}^{(p)} - \tau_k^{(p)}$ are independent and identically distributed. Furthermore, for every $x > 0$,*

$$\lim_{x\downarrow 0}\Big(\limsup_{p\to\infty} P[\tau_1^{(p)} < x]\Big) = 0.$$

Proof. The first assertion is a straightforward consequence of Lemma 4.3. Let us turn to the second assertion. To simplify notation, we write $\delta' = \delta/5$. For every $\eta > 0$, set

$$T_\eta^{(p)} = \inf\{t \geq 0 : \frac{1}{\sqrt{p}} S_{[pt]} < -\eta\}.$$

Then, from the definition of $\tau_1^{(p)}$, we get

$$P[\tau_1^{(p)} < x] = P\Big[\sup_{s<x} H_s^{(p)} > \delta'\Big] \leq P\Big[\sup_{s\leq T_\eta^{(p)}} H_s^{(p)} > \delta'\Big] + P[T_\eta^{(p)} < x].$$

On one hand, by (11)

$$\limsup_{p\to\infty} P[T_\eta^{(p)} < x] \leq \limsup_{p\to\infty} P\Big[\inf_{t\leq x} \frac{1}{\sqrt{p}} S_{[pt]} \leq -\eta\Big] \leq P\Big[\inf_{t\leq x} B_t \leq -\eta/\sigma\Big],$$

and the right-hand side goes to zero as $x \downarrow 0$, for any choice of $\eta > 0$.

On the other hand, we first note that $H_n = 0$ if and only if there exists an integer $j \geq 0$ such that $n = \gamma_j := \inf\{k \geq 0 : S_k = -j\}$ (this is immediate from (10)). It follows that the time interval $[\gamma_j, \gamma_{j+1}[$ exactly corresponds in the time scale of the height process to the visits of the individuals of the tree $\mathcal{T}_j$. Consequently,

$$\sqrt{p} \sup_{s\leq T_\eta^{(p)}} H_s^{(p)} = \max(\mathrm{ht}(\mathcal{T}_0), \mathrm{ht}(\mathcal{T}_1), \ldots, \mathrm{ht}(\mathcal{T}_{[\eta\sqrt{p}]})),$$

and therefore,

$$P\Big[\sup_{s\leq T_\eta^{(p)}} H_s^{(p)} > \delta'\Big] = 1 - (1 - P[\mathrm{ht}(\mathcal{T}_0) > \delta'\sqrt{p}])^{[\eta\sqrt{p}]}.$$

From this identity and (5), we get

$$\lim_{\eta\to 0}\Big(\limsup_{p\to\infty} P\Big[\sup_{s\leq T_\eta^{(p)}} H_s^{(p)} > \delta'\Big]\Big) = 0.$$

The second assertion of the lemma now follows. $\square$

We can now complete the proof of tightness. From (15) and the first assertion of Lemma 4.6, we have, for every integer $L \geq 1$,

$$\begin{aligned} A_2(n,p) &\leq \sum_{k=0}^{L-1} P[\tau^{(p)}_{k+1} - \tau^{(p)}_k < 2^{-n}T] + P[\tau^{(p)}_L < T] \\ &\leq L\,P[\tau^{(p)}_1 < 2^{-n}T] + e^T\,E[e^{-\tau^{(p)}_L}] \\ &= L\,P[\tau^{(p)}_1 < 2^{-n}T] + e^T\,(E[e^{-\tau^{(p)}_1}])^L. \end{aligned}$$

The second assertion of the lemma implies that

$$\limsup_{p\to\infty} E[e^{-\tau^{(p)}_1}] = a < 1.$$

It follows that, for every integer $L \geq 1$,

$$\limsup_{p\to\infty} A_2(n,p) \leq L\,\limsup_{p\to\infty} P[\tau^{(p)}_1 < 2^{-n}T] + e^T\,a^L.$$

By choosing L large and then letting $n \to \infty$, we deduce from this bound and the second assertion of Lemma 4.6 that

$$\lim_{n\to\infty}\Big(\limsup_{p\to\infty} A_2(n,p)\Big) = 0.$$

The same result holds for $A_3(n,p)$. This completes the proof of (12) and of the convergence (7).

Third step. It remains to establish the convergence (8). We let $v_0, v_1, \ldots$ denote the vertices of the trees $\mathcal{T}_0, \mathcal{T}_1, \ldots$ enumerated one tree after another and in lexicographical order for each tree. Note that we have $H_n = |v_n|$, by the definition of the height function of a tree. By preceding observations, the interval corresponding to the tree $\mathcal{T}_j$ in the enumeration $v_0, v_1, \ldots$ is $[\gamma_j, \gamma_{j+1}[$, where $\gamma_j = \inf\{k \geq 0 : S_k = -j\}$. Thus the individual v_n belongs to $\mathcal{T}_j$ if $j = -I_n$. We then set, for every $n \geq 0$,

$$K_n = 2n - H_n + 2I_n.$$

From the convergences (7) and (11), it is immediate that

$$\frac{K_n}{n} \xrightarrow[n\to\infty]{\text{(P)}} 2.$$

As in Section 3, we can define a contour exploration of the sequence of trees $\mathcal{T}_0, \mathcal{T}_1, \ldots$, in such a way that, for every integer $n \geq 0$, C_n is the generation of the individual visited at time n in this exploration (when completing the visit of a tree, the contour exploration immediately jumps to the root of the next tree, and so there may be several individuals visited at the same time n, but of course $C_n = 0$ in that case). It is easily checked by induction on n that K_n is the time at which the contour exploration first visits the individual v_n, and in particular $C_{K_n} = H_n$. From this observation, we get

$$\sup_{t\in[K_n,K_{n+1}]} |C_t - H_n| \leq |H_{n+1} - H_n| + 1.$$

Define a random function $\varphi : \mathbb{R}_+ \longrightarrow \mathbb{N}$ by setting $\varphi(t) = n$ iff $t \in [K_n, K_{n+1}[$. Fix an integer $m \geq 1$ in the remaining part of the proof. From the previous bound, we get

$$\sup_{t\in[0,K_m]} |C_t - H_{\varphi(t)}| \leq 1 + \sup_{n\leq m} |H_{n+1} - H_n|. \tag{16}$$

Similarly, it follows from the definition of K_n that

$$\sup_{t\in[0,K_m]} |\varphi(t) - \frac{t}{2}| \leq \frac{1}{2} \sup_{n\leq m} H_n - I_m + 1. \tag{17}$$

For every $p \geq 1$, set $\varphi_p(t) = p^{-1}\varphi(pt)$. By (16), we have

$$\sup_{t\leq p^{-1}K_{mp}} \Big|\frac{1}{\sqrt{p}} C_{pt} - \frac{1}{\sqrt{p}} H_{p\varphi_p(t)}\Big| \leq \frac{1}{\sqrt{p}} + \frac{1}{\sqrt{p}} \sup_{n\leq mp} |H_{n+1} - H_n| \xrightarrow[p\to\infty]{\text{(P)}} 0 \tag{18}$$

by the convergence (7).

On the other hand, we get from (17) that

$$\sup_{t\leq p^{-1}K_{mp}} |\varphi_p(t) - \frac{t}{2}| \leq \frac{1}{2p} \sup_{k\leq mp} H_k - \frac{1}{p} I_{mp} + \frac{1}{p} \xrightarrow[p\to\infty]{\text{(P)}} 0 \tag{19}$$

by (7) and (11).

The convergence (8) now follows from (18) and (19), using also the fact that $p^{-1}K_{mp}$ converges in probability to $2m$ as $p \to \infty$. □

Remark. There is one special case where the convergence (9) is easy. This is the case where μ is the geometric distribution $\mu(k) = 2^{-k-1}$, which satisfies our assumptions with $\sigma^2 = 2$. In that case, it is not hard to see that the restriction to integers of the contour process is distributed as a simple random walk reflected at the origin. Thus the convergence (9) follows from Donsker's invariance theorem.

To conclude this section, we note that the convergence (7) can be reinforced in the following way. Set $J_n = -I_n$, so that J_n corresponds to the index j such that $\mathcal{T}_j$ contains the individual v_n. Then we have

$$\Big(\frac{\sigma}{2\sqrt{p}} H_{[pt]}, \frac{1}{\sigma\sqrt{p}} J_{[pt]}\Big)_{t\geq 0} \xrightarrow[p\to\infty]{\text{(d)}} (\beta_t, L_t^0)_{t\geq 0}, \tag{20}$$

where L_t^0 is the local time at 0 at time t of the reflected Brownian motion β (defined as in Section 2). Indeed, it readily follows from the proof of Theorem 4.1 and (11) that we have the joint convergence

$$\Big(\frac{\sigma}{2\sqrt{p}} H_{[pt]}, \frac{1}{\sigma\sqrt{p}} J_{[pt]}\Big)_{t\geq 0} \xrightarrow[p\to\infty]{\text{(d)}} \Big(B_t - \inf_{0\leq s\leq t} B_s, -\inf_{0\leq s\leq t} B_s\Big)_{t\geq 0}.$$

However, by an already mentioned famous theorem of Lévy,

$$\Big(B_t - \inf_{0\leq s\leq t} B_s, -\inf_{0\leq s\leq t} B_s\Big)_{t\geq 0} \overset{\text{(d)}}{=} (\beta_t, L_t^0)_{t\geq 0}.$$

5 Convergence towards the Itô measure

Recall our notation E for the set of excursions, and d for the distance on E. If A is a nonempty subset of E and $e \in E$, we set

$$d(e, A) = \inf_{e'\in A} d(e, e').$$

A nonempty open subset A of E is called regular if $\mathbf{n}(A) < \infty$ and if

$$\mathbf{n}(\{e \in E : d(e, A) < \varepsilon\}) \xrightarrow[\varepsilon\to 0]{} \mathbf{n}(A).$$

Let $(h_n)_{0\leq n\leq \#\mathcal{T}-1}$ be the height function of a μ-Galton-Watson tree $\mathcal{T}$. By convention we set $h_n = 0$ for $n \geq \#\mathcal{T}$. The function $n \to h_n$ is extended to $\mathbb{R}_+$ by linear interpolation over each interval $[i-1, i]$, $i \in \mathbb{N}$. If the tree $\mathcal{T}$ is nontrivial, meaning that $\#\mathcal{T} > 1$, $(h_t)_{t\geq 0}$ is a random element of E (of course if $\#\mathcal{T} = 1$, $(h_t)_{t\geq 0}$ is just the zero function). We rescale the random function $(h_t)_{t\geq 0}$ by setting, for every integer $p \geq 1$,

$$h_t^{(p)} = \frac{\sigma}{2\sqrt{p}}\, h_{pt}\ , \qquad t \geq 0\,.$$

Similarly, we let $(c_t)_{t\geq 0}$ be the contour function of $\mathcal{T}$ and we rescale this function by setting, for every integer $p \geq 1$,

$$c_t^{(p)} = \frac{\sigma}{2\sqrt{p}}\, c_{2pt}\ , \qquad t \geq 0\,.$$

Theorem 5.1 *Let A be a regular open subset of E. Then,*

$$P[h^{(p)} \in A] \underset{p\to\infty}{\sim} \frac{\mathbf{n}(A)}{\sigma\sqrt{p}}. \tag{21}$$

Furthermore, the law of $h^{(p)}$ under the conditional probability measure $P[\cdot\,|\, h^{(p)} \in A]$ converges to $\mathbf{n}(\cdot\,|\,A)$ as $p \to \infty$. Similarly, the law of $c^{(p)}$ under $P[\cdot\,|\, h^{(p)} \in A]$ converges to $\mathbf{n}(\cdot\,|\,A)$ as $p \to \infty$.

Proof. We start by proving the second assertion of the theorem. As in the preceding section, we consider the height process $(H_n)_{n\geq 0}$ associated with a sequence $\mathcal{T}_0, \mathcal{T}_1, \ldots$ of independent μ-Galton-Watson trees. We extend the function $n \to H_n$ to $\mathbb{R}_+$ using linear interpolation as above. By Theorem 4.1, we have

$$\Big(\frac{\sigma}{2\sqrt{p}} H_{pt}\Big)_{t\geq 0} \underset{p\to\infty}{\overset{\text{(d)}}{\longrightarrow}} (\beta_t)_{t\geq 0} \tag{22}$$

where $(\beta_t)_{t\geq 0}$ is a reflected Brownian motion. We now observe that the successive excursions of $(H_t)_{t\geq 0}$ from 0 correspond to the height processes associated with the trees $\mathcal{T}_0, \mathcal{T}_1, \ldots$ (in fact only with those nontrivial trees in the sequence). It follows that the law of $h^{(p)}$ under the conditional probability measure $P[\cdot \mid h^{(p)} \in A]$ coincides with the law of the first excursion of the process $(\frac{\sigma}{2\sqrt{p}} H_{pt})_{t\geq 0}$ away from 0 that falls in the set A. We write $h^{(p),A}$ for this excursion.

On the other hand, let $e^A = (e_t^A)_{t\geq 0}$ be the first excursion of β away from 0 that falls in A. By Corollary 2.2, the law of e^A is $\mathbf{n}(\cdot\,|\,A)$. Thus the proof of the second assertion of the theorem reduces to checking that

$$h^{(p),A} \underset{p\to\infty}{\overset{\text{(d)}}{\longrightarrow}} e^A\,. \tag{23}$$

To this end, it will be convenient to use the Skorokhod representation theorem in order to replace the convergence in distribution in (22) by an almost sure convergence. Let the process $(J_n)_{n\geq 0}$ be defined as in the previous section, and recall the convergence (20). From the Skorokhod representation theorem, we can then find for every integer $p \geq 1$ a pair $(\widetilde{H}_t^{(p)}, \widetilde{J}_t^{(p)})_{t\geq 0}$ such that

$$(\widetilde{H}_t^{(p)}, \widetilde{J}_t^{(p)})_{t\geq 0} \overset{\text{(d)}}{=} \Big(\frac{\sigma}{2\sqrt{p}} H_{pt}, \frac{1}{\sigma\sqrt{p}} J_{[pt]}\Big)_{t\geq 0}$$

and

$$(\widetilde{H}_t^{(p)}, \widetilde{J}_t^{(p)})_{t\geq 0} \underset{p\to\infty}{\longrightarrow} (\beta_t, L_t^0)_{t\geq 0}, \tag{24}$$

uniformly on every compact subset of $\mathbb{R}_+$, almost surely. We write $\widetilde{h}^{(p),A}$ for the first excursion of $\widetilde{H}^{(p)}$ away from 0 that falls in A. Then $\widetilde{h}^{(p),A} \overset{\text{(d)}}{=} h^{(p),A}$, and thus (23) will follow if we can verify that

$$\widetilde{h}^{(p),A} \xrightarrow[p\to\infty]{\text{(a.s.)}} e^A\,. \tag{25}$$

To this end, write $]g^A, d^A[$ for the time interval associated with the excursion e^A, and $m^A = \frac{1}{2}(g^A + d^A)$. By (24), we have $\widetilde{H}^{(p)}_{m^A} > 0$ for p large enough, and thus we can consider the excursion interval of $\widetilde{H}^{(p)}$ that straddles m^A. Denote this interval by $]g^{(p),A}, d^{(p),A}[$. It is now easy to see that we have the almost sure convergences

$$\lim_{p\to\infty} d^{(p),A} = d^A\,, \tag{26}$$

and

$$\lim_{p\to\infty} g^{(p),A} = g^A\,. \tag{27}$$

Let us verify (26) for instance. From the convergence of $\widetilde{H}^{(p)}_t$ towards β_t in (24), it immediately follows that

$$\liminf_{p\to\infty} d^{(p),A} \geq d^A.$$

On the other hand, by the results recalled in Section 2, we know that, for every fixed $\varepsilon > 0$, $L^0_{d^A+\varepsilon} > L^0_{d^A}$ a.s., and thus the convergence of $\widetilde{J}^{(p)}_t$ towards L^0_t in (24) ensures that $\widetilde{J}^{(p)}_{d^A+\varepsilon} > \widetilde{J}^{(p)}_{d^A}$ for all p large enough, a.s. Now note that J_n only increases when H_n vanishes, and thus a similar property holds for $\widetilde{J}^{(p)}$ and $\widetilde{H}^{(p)}$. It follows that $H^{(p)}$ vanishes somewhere between d^A and $d^A + \varepsilon$, and therefore $d^{(p),A} \leq d^A + \varepsilon$, for all p large enough, a.s. This completes the proof of (26), and a similar argument applies to (27).

From (26), (27) and the convergence (24), we deduce that the excursion of $\widetilde{H}^{(p),A}$ that straddles m_A, which we denote by $\overline{h}^{(p),A}$, converges a.s. to e^A as $p \to \infty$. Since $e^A \in A$ and A is open, we have $\overline{h}^{(p),A} \in A$ for all p large enough, a.s. In order to establish (25), and thus to complete the proof of the second assertion of the theorem, it remains to establish that $\widetilde{h}^{(p),A} = \overline{h}^{(p),A}$ for all p large enough, a.s. Equivalently, we need to check that no excursion of $\widetilde{H}^{(p)}$ before time $g^{(p),A}$ belongs to A.

To this end, we set $A_\varepsilon = \{e \in E : d(e, A) < \varepsilon\}$ and we use the fact that $\mathbf{n}(A_\varepsilon)$ tends to $\mathbf{n}(A)$ as $\varepsilon \to 0$. It follows that almost surely there exists a (random) value $\varepsilon_0 > 0$ such that the first excursion of β that belongs to A is also the first excursion of β that belongs to A_{ε_0}. On the other hand, the convergence (24) implies that for every $t \geq 0$, the zero set of the random function $\widetilde{H}^{(p)}$ before time t converges a.s. as $p \to \infty$ to the zero set of β before time t, in the sense of the Hausdorff distance between compact subsets of $\mathbb{R}_+$ (use the fact that the points of increase of the local time process L^0 are exactly the zeros of β). Using (24) once more, it follows that if p is large enough, any excursion of $\widetilde{H}^{(p)}$ before time d^A lies within distance less than ε_0 from an excursion of β before time d^A. Since no excursion of β before g^A belongs to A_{ε_0}, it follows that for p large enough no excursion of $\widetilde{H}^{(p)}$ before time $g^{(p),A}$ belongs to A. This completes the proof of the second assertion of the theorem.

The proof of the first assertion is then easy. For every $p \geq 1$, set

$$\widetilde{N}^{(p)}_A = \sigma\sqrt{p}\,\widetilde{J}^{(p)}_{g^{A,(p)}}.$$

From (24) and (27), we have

$$\frac{1}{\sigma\sqrt{p}}\,\widetilde{N}_A^{(p)} \xrightarrow[p\to\infty]{\text{a.s.}} L^0_{g^A}.$$

Notice that $L^0_{g^A}$ is exponentially distributed with parameter $\mathbf{n}(A)$, by Corollary 2.2. On the other hand, $\widetilde{N}_A^{(p)}$ has the same distribution as the index of the first tree in the sequence $\mathcal{T}_0, \mathcal{T}_1, \ldots$ whose rescaled height process belongs to A. Hence $\widetilde{N}_A^{(p)}$ has a geometric distribution,

$$P[\widetilde{N}_A^{(p)} = k] = \alpha_A^{(p)}\,(1-\alpha_A^{(p)})^k$$

with $\alpha_A^{(p)} = P[h^{(p)} \in A]$. Consequently, for every $x > 0$,

$$(1-\alpha_A^{(p)})^{[x\sigma\sqrt{p}]} = P[\widetilde{N}_A^{(p)} \geq x\sigma\sqrt{p}] \xrightarrow[p\to\infty]{} e^{-x\mathbf{n}(A)}.$$

The estimate (21) follows.

Finally, the last assertion is obtained by verifying that, for every $\varepsilon > 0$,

$$\lim_{p\to\infty} P[d(h^{(p)}, c^{(p)}) > \varepsilon \,|\, h^{(p)} \in A] = 0,$$

where d is the distance on E. This can be derived from the second assertion of the theorem, by arguments very similar to the third step of the proof of Theorem 4.1. We omit details. $\square$

In the next section, we will need a minor strengthening of the second assertion of Theorem 4.1. Denote the Lukasiewicz path (up to time $\#\mathcal{T}-1$) of the tree $\mathcal{T}$ by $(\Sigma_k)_{0\leq k<\#\mathcal{T}}$, and set $\Sigma_k = 0$ for $k \geq \#\mathcal{T}$. Use linear interpolation to define Σ_t for every real $t \geq 0$, and set

$$\Sigma_t^{(p)} = \frac{1}{\sigma\sqrt{p}}\,\Sigma_{pt}$$

for every $t \geq 0$ and $p \geq 1$. Then the law of the pair $(h^{(p)}, \Sigma^{(p)})$ under $P[\cdot\,|\,h^{(p)} \in A]$ converges as $p \to \infty$ to the law of (e, e) under $\mathbf{n}(de\,|\,A)$. To see this, first note that the limiting result (7) can be combined with (11) to give the joint convergence

$$\Big(\frac{\sigma}{2\sqrt{p}}H_{pt}, \frac{1}{\sigma\sqrt{p}}(S_{pt}-I_{pt})\Big)_{t\geq 0} \xrightarrow[p\to\infty]{\text{(d)}} \Big(\beta_t, \beta_t\Big)_{t\geq 0} \tag{28}$$

where $(S_t)_{t\geq 0}$ denotes the Lukasiewicz path of the forest $\mathcal{T}_0, \mathcal{T}_1, \ldots$ (extended to real values of t by linear interpolation) and $I_t := \inf\{S_r : 0 \leq r \leq t\}$. This joint convergence is indeed immediate from the method we used to prove Theorem 4.1. Then, if $h^{(p),A}$ is as in the previous proof, the interval associated with the excursion $h^{(p),A}$ is also an excursion interval of the process $\frac{1}{\sigma\sqrt{p}}(S_{pt}-I_{pt})$ (this is so because the integer times at which H vanishes are precisely those at which $S-I$ vanishes). Write $\Sigma^{(p),A}$ for the excursion of the process $\frac{1}{\sigma\sqrt{p}}(S_{pt}-I_{pt})$ corresponding to the interval associated with $h^{(p),A}$. Then the law of $(h^{(p),A}, \Sigma^{(p),A})$ coincides with the law of $(h^{(p)}, \Sigma^{(p)})$ under $P[\cdot\,|\,h^{(p)} \in A]$. The remaining part of the argument is exactly the same as in the above proof, using (28) instead of (7).

Examples. Let us discuss a few examples of application of Theorem 5.1, which will be useful in the next sections.

(a) Let $A = \{e \in E : \zeta(e) > 1\}$. Then it is easy to verify that A is a regular open subset of E. Theorem 5.1 implies that the rescaled height function (or contour function) of a μ-Galton-Watson tree conditioned to have total progeny greater than p converges to a Brownian excursion conditioned to have duration greater than 1.

(b) Let $A = \{e \in E : \max(e) > \sigma/2\}$. Again, A is regular. Theorem 5.1 implies that the rescaled height function (or contour function) of a μ-Galton-Watson tree conditioned to have height greater than $\sqrt{p}$ converges in distribution to a Brownian excursion conditioned to have height greater than $\sigma/2$. For $\varepsilon > 0$, we can also take $A = \{e \in E : \sigma/2 < \max(e) < (1+\varepsilon)\sigma/2\}$ and obtain the scaling limit of the height or contour process of a μ-Galton-Watson tree conditioned to have height between $\sqrt{p}$ and $(1+\varepsilon)\sqrt{p}$. This last case will be relevant in Section 7 below.

Convergence of rescaled trees in the Gromov-Hausdorff topology.

Theorem 5.1 immediately implies a result of convergence of (conditioned) rescaled random trees in the sense of the Gromov-Hausdorff topology. To state this precisely, we need to define the real tree coded by an excursion. So let $e \in E$, and for every $s, t \in [0, \zeta(e)]$ set

$$d_e(s,t) = e(s) + e(t) - 2 \inf_{s \wedge t \le r \le s \vee t} e(r).$$

Then d_e is a pseudo-distance on $[0, \zeta(e)]$, and we can consider the associated equivalence relation

$$s \sim_e t \quad \text{iff} \quad d_e(s,t) = 0.$$

The quotient space $\mathbf{T}_e := [0, \zeta(e)]/ \sim_e$ equipped with the distance d_e is a compact metric space, and is in fact a real tree (see e.g. Section 2 of [19], and [11] for a general overview of real trees in probability theory). Furthermore the mapping $e \longrightarrow \mathbf{T}_e$ is continuous in the following sense. Let D_{GH} denote the Gromov-Hausdorff distance between (isometry classes of) compact metric spaces. Then, if $e, e' \in E$,

$$D_{GH}(\mathbf{T}_e, \mathbf{T}_{e'}) \le 2\, d(e, e') \tag{29}$$

(see Lemma 2.4 in [19]).

Write d_{gr} for the usual graph distance on the tree $\mathcal{T}$, so that $(\mathcal{T}, d_{gr})$ is a random compact metric space. Then Theorem 5.1 implies that the law of the rescaled space $(\mathcal{T}, \frac{\sigma}{2\sqrt{p}} d_{gr})$ under $P[\cdot \,|\, h^{(p)} \in A]$ converges as $p \to \infty$ to the law of $\mathbf{T}_e$ under $\mathbf{n}(de \,|\, A)$, in the sense of the Gromov-Hausdorff topology.

To see this, note that the Gromov-Hausdorff distance between $(\mathcal{T}, d_{gr})$ and the space $(\mathbf{T}_c, d_c)$ is easily bounded above by 1 (here c is the contour function of $\mathcal{T}$). It follows that

$$D_{GH}\Big((\mathcal{T}, \frac{\sigma}{2\sqrt{p}} d_{gr}), (\mathbf{T}_{c^{(p)}}, d_{c^{(p)}})\Big) \le \frac{\sigma}{2\sqrt{p}}.$$

However, Theorem 5.1 and (29) immediately imply that the law of $\mathbf{T}_{c^{(p)}}$ under $P[\cdot \mid h^{(p)} \in A]$ converges as $p \to \infty$ to the law of $\mathbf{T}_e$ under $\mathbf{n}(de \,|\, A)$.

6 Aldous' theorem

In this section, we show how our methods can be used to recover a famous theorem of Aldous [2, Theorem 23] about the scaling limit of the contour function of a Galton-Watson tree conditioned to have a fixed total progeny.

We consider an offspring distribution μ satisfying the assumptions of the previous sections, and $\nu(k) = \mu(k+1)$ for every $k = -1, 0, 1, \ldots$, as previously. Let G be the smallest subgroup of $\mathbb{Z}$ that contains the support of μ. Clearly, the total progeny of a μ-Galton-Watson tree belongs to $1+G$, and conversely, for every sufficiently large integer $p \in G$, the probability that the total progeny is equal to $1+p$ is nonzero (this can be seen by combining the last assertion of Lemma 3.1 with Kemperman's formula recalled below and a suitable local limit theorem).

Theorem 6.1 *For every sufficiently large integer $p \in G$, let $C^{\langle p\rangle} = (C_t^{\langle p\rangle})_{0\le t\le 2p}$ be the contour function of a μ-Galton-Watson tree conditioned to have exactly $p+1$ vertices. Then,*

$$\Big(\frac{1}{\sqrt{p}}\,C^{\langle p\rangle}_{2pt}\Big)_{0\le t\le 1} \underset{p\to\infty,\, p\in G}{\xrightarrow{\text{(d)}}} \Big(\frac{2}{\sigma}\,\mathbf{e}_t\Big)_{0\le t\le 1}$$

where $(\mathbf{e}_t)_{0\le t\le 1}$ is a normalized Brownian excursion (i.e., $\mathbf{e}$ is distributed according to $\mathbf{n}(\cdot\,|\,\zeta=1)$) and the convergence is in the sense of weak convergence of the laws on $C([0,1],\mathbb{R}_+)$.

To avoid technicalities, we will concentrate on the aperiodic case where $G=\mathbb{Z}$. We let $S=(S_n)_{n\ge 0}$ be a random walk with jump distribution ν started from 0 under P. For $k\neq 0$, we also use the notation P_k for the probability measure under which the random walk S starts from k, but unless otherwise indicated we argue under the probability measure P.

We will make use of a classical local limit theorem (see e.g. [27], pp.77-79). We have

$$\lim_{\ell\to\infty}\sup_{x\in\mathbb{Z}}\,(1\vee\frac{x^2}{\ell})\,\Big|\sqrt{\ell}\,P[S_\ell=x]-g_{\sigma^2}(\frac{x}{\sqrt{\ell}})\Big| = 0, \tag{30}$$

where

$$g_t(y)=\frac{1}{\sqrt{2\pi t}}\,\exp(-\frac{y^2}{2t})$$

for $t>0$ and $y\in\mathbb{R}$, is the standard Brownian transition density.

We also recall Kemperman's formula (see e.g. [23], p.122). Set $T=\inf\{n\ge 0: S_n=-1\}$. Then, for every integers $k\ge 0$ and $n\ge 1$,

$$P_k[T=n]=\frac{k+1}{n}P_k[S_n=-1]=\frac{k+1}{n}P[S_n=-k-1].$$

Proof of Theorem 6.1. The random walk S is (under the probability measure P) the Lukasiewicz path of a forest $\mathcal{T}_0,\mathcal{T}_1,\dots$ of independent μ-Galton-Watson trees. We denote by $(H_n)_{n\ge 0}$ the height process of this forest and by $(C_t)_{t\ge 0}$ the corresponding contour process. Then the process $(C_t)_{0\le t\le 2(T-1)}$ is the contour function of $\mathcal{T}_0$, and $\#\mathcal{T}_0=T$. Consequently, the law of $(C_t^{(p)})_{0\le t\le 2p}$ coincides with the law of $(C_t)_{0\le t\le 2p}$ under $P[\cdot\,|\,T=p+1]$.

Also note that, by Kemperman's formula, we have for every integer $p\ge 0$

$$P[T=p+1]=\frac{1}{p+1}P[S_{p+1}=-1] \tag{31}$$

which is positive for p large enough by (30).

For every (sufficiently large) integer p, let us consider the conditional probability measures

$$\mathbf{Q}_p=P[\cdot\,|\,T=p+1]\quad,\quad \mathbf{Q}^*_p=P[\cdot\,|\,T>p]\,.$$

Theorem 5.1 yields information about the scaling limit of the process $(H_n)_{0\le n\le T-1}$ under $\mathbf{Q}^*_p$ when $p\to\infty$. Precisely, if we set

$$h_n=\begin{cases} H_n & \text{if } 0\le n\le T-1\\ 0 & \text{if } n\ge T,\end{cases}$$

Theorem 5.1 (cf Example (a) in Section 5) implies that the law under $\mathbf{Q}_p^*$ of the rescaled process

$$h_t^{(p)} := \frac{\sigma}{2\sqrt{p}}\, h_{[pt]}, \qquad t \geq 0$$

converges as $p \to \infty$ towards $\mathbf{n}(\cdot\,|\,\zeta > 1)$.

The next step is to compare $\mathbf{Q}_p^*$ and $\mathbf{Q}_p$ when p is large. We fix $a \in]0,1[$. For every $p \geq 1$, let f_p be a nonnegative function on $\mathbb{Z}^{[ap]+1}$. An easy application of the Markov property of the random walk S shows that

$$\mathbf{Q}_p[f_p((S_k)_{0\leq k\leq [ap]})] = E\Big[f_p((S_k)_{0\leq k\leq [ap]})\,\mathbf{1}_{\{T>[ap]\}}\,\frac{\phi_p(S_{[ap]})}{P[T=p+1]}\Big] \tag{32}$$

where, for every integer $k \geq 0$,

$$\phi_p(k) = P_k[T = p+1-[ap]\,].$$

Similarly,

$$\mathbf{Q}_p^*[f_p((S_k)_{0\leq k\leq [ap]})] = E\Big[f_p((S_k)_{0\leq k\leq [ap]})\,\mathbf{1}_{\{T>[ap]\}}\,\frac{\phi_p^*(S_{[ap]})}{P[T>p]}\Big] \tag{33}$$

where

$$\phi_p^*(k) = P_k[T > p-[ap]\,].$$

As a simple consequence of (31) and (30), we have

$$\lim_{p\to\infty} p^{3/2}\,P[T=p+1] = \frac{1}{\sigma\sqrt{2\pi}} \tag{34}$$

and

$$\lim_{p\to\infty} p^{1/2}\,P[T>p] = \frac{2}{\sigma\sqrt{2\pi}}. \tag{35}$$

Let us then consider $\phi_p(k)$ and $\phi_p^*(k)$. By Kemperman's formula again, we have

$$\phi_p(k) = P_k[T=p+1-[ap]\,] = \frac{k+1}{p+1-[ap]}\,P[S_{p+1-[ap]} = -k-1].$$

It then follows from (30) that

$$\lim_{p\to\infty}\Big(\sup_{k\geq 0}\Big|p\phi_p(k) - (1-a)^{-3/2}\,\frac{k+1}{\sqrt{p}}\,g_{\sigma^2}\Big(\frac{k+1}{\sqrt{p+1-[ap]}}\Big)\Big|\Big) = 0.$$

Furthermore,

$$\phi_p^*(k) = \sum_{\ell>p-[ap]} P_k[T=\ell] = \sum_{\ell>p-[ap]} \frac{k+1}{\ell}\,P[S_\ell = -k-1]$$

and from (30) again we have

$$\lim_{p\to\infty}\Big(\sup_{k\geq 0}\Big|\phi_p^*(k) - \sum_{\ell>p-[ap]}\frac{k+1}{\ell^{3/2}}\,g_{\sigma^2}\Big(\frac{k+1}{\sqrt{\ell}}\Big)\Big|\Big) = 0$$

(note that we use the precise form of the estimate (30), and in particular the presence of the factor $1 \vee \frac{x^2}{k}$ in this estimate). Recall the notation $q_t(x)$ from Section 2, and note that $q_t(x) = \frac{x}{t}\, g_t(x)$ for every $t > 0$ and $x > 0$. Elementary analysis shows that

$$\lim_{p\to\infty} \Big(\sup_{k\geq 0} \Big| \sum_{\ell > p-[ap]} \frac{k+1}{\ell^{3/2}}\, g_{\sigma^2}\Big(\frac{k+1}{\sqrt{\ell}}\Big) - \int_{1-a}^{\infty} ds\, q_s\Big(\frac{k+1}{\sigma\sqrt{p}}\Big)\Big|\Big) = 0.$$

Summarizing, we have

$$\lim_{p\to\infty} \Big(\sup_{k\geq 0} \Big| p\phi_p(k) - q_{1-a}(\frac{k+1}{\sigma\sqrt{p}})\Big|\Big) = 0$$

and

$$\lim_{p\to\infty} \Big(\sup_{k\geq 0} \Big| \phi^*_p(k) - \int_{1-a}^{\infty} ds\, q_s\Big(\frac{k+1}{\sigma\sqrt{p}}\Big)\Big|\Big) = 0.$$

From (32) and (33) we have

$$\mathbf{Q}_p[f_p((S_k)_{0\leq k\leq [ap]})] = \mathbf{Q}^*_p\Big[f_p((S_k)_{0\leq k\leq [ap]})\, \frac{\phi_p(S_{[ap]})}{\phi^*_p(S_{[ap]})}\, \frac{P[T>p]}{P[T=p+1]}\Big].$$

From the preceding estimates, we have for every $c > 0$,

$$\lim_{p\to\infty} \Big(\sup_{k\geq c\sqrt{p}} \Big| \frac{\phi_p(k)}{\phi^*_p(k)}\, \frac{P[T>p]}{P[T=p+1]} - \Gamma_a(\frac{k}{\sigma\sqrt{p}})\Big|\Big) = 0,$$

where, for every $x > 0$,

$$\Gamma_a(x) = \frac{2q_{1-a}(x)}{\int_{1-a}^{\infty} ds\, q_s(x)}.$$

On the other hand, using (30) once again, it is easy to verify that,

$$\lim_{p\to\infty} \mathbf{Q}_p[S_{[ap]} \leq c\sqrt{p}\,] = 0 \quad , \quad \lim_{p\to\infty} \mathbf{Q}^*_p[S_{[ap]} \leq c\sqrt{p}\,] = 0.$$

Assume that the functions f_p are uniformly bounded. From the previous discussion, we now get that

$$\lim_{p\to\infty} \Big| \mathbf{Q}_p[f_p((S_k)_{0\leq k\leq [ap]})] - \mathbf{Q}^*_p\Big[f_p((S_k)_{0\leq k\leq [ap]})\, \Gamma_a(\frac{S_{[ap]}}{\sigma\sqrt{p}})\Big]\Big| = 0.$$

We apply this convergence to

$$f_p((S_k)_{0\leq k\leq [ap]}) = F((h^{(p)}_t)_{0\leq t\leq a}),$$

where F is a bounded continuous function on $\mathbb{D}([0,a],\mathbb{R})$. By the remarks following the proof of Theorem 5.1, we know that the law under $\mathbf{Q}^*_p$ of

$$\Big(h^{(p)}_t, \frac{S_{[pt]}}{\sigma\sqrt{p}}\Big)_{0\leq t\leq 1}$$

converges to the law of $(e_t, e_t)_{0\leq t\leq 1}$ under $\mathbf{n}(\cdot\,|\,\zeta > 1)$, and thus we have

$$\lim_{p\to\infty} \mathbf{Q}^*_p\Big[F((h^{(p)}_t)_{0\leq t\leq a})\, \Gamma_a(\frac{S_{[ap]}}{\sigma\sqrt{p}})\Big] = \mathbf{n}\Big(F((e_t)_{0\leq t\leq a})\, \Gamma_a(e_a)\,\Big|\, \zeta > 1\Big).$$

It follows that we have also

$$\lim_{p\to\infty} \mathbf{Q}_p\Big[F((h_t^{(p)})_{0\le t\le a})\Big] = \mathbf{n}\Big(F((e_t)_{0\le t\le a})\,\Gamma_a(e_a)\,\Big|\,\zeta>1\Big) = \mathbf{n}\Big(F((e_t)_{0\le t\le a}\,\Big|\,\zeta=1\Big),$$

by (1) and (2). In other words, the law of $(h_t^{(p)})_{0\le t\le a}$ under $\mathbf{Q}_p$ converges to the law of $(e_t)_{0\le t\le a}$ under $\mathbf{n}(\cdot\,|\,\zeta=1)$.

Then, by the very same arguments as in the third step of the proof of Theorem 4.1, we can verify that, for every $\varepsilon>0$,

$$\lim_{p\to\infty} \mathbf{Q}_p\Big[\sup_{0\le t\le a}\frac{1}{\sqrt{p}}|H_{[pt]}-C_{2pt}|>\varepsilon\Big]=0.$$

Hence, we also get that the law under $\mathbf{Q}_p$ of

$$\Big(\frac{\sigma C_{2pt}}{2\sqrt{p}}\Big)_{0\le t\le a}$$

converges to the law of $(e_t)_{0\le t\le a}$ under $\mathbf{n}(\cdot\,|\,\zeta=1)$. Since this holds for every $a\in]0,1[$, and since $\mathbf{Q}_p[C_{2p}=0]=1$, the finite-dimensional marginals of the processes

$$\Big(\frac{\sigma C_{2pt}}{2\sqrt{p}}\Big)_{0\le t\le 1}$$

under $\mathbf{Q}_p$ converge to the finite-dimensional marginals of $(e_t)_{0\le t\le 1}$ under $\mathbf{n}(\cdot\mid\zeta=1)$. On the other hand, tightness of this sequence of processes is immediate from the convergence over the time interval $[0,a]$ and the fact that the processes $(C_{2pt})_{0\le t\le 1}$ and $(C_{2p(1-t)})_{0\le t\le 1}$ have the same distribution under $\mathbf{Q}_p$. This completes the proof of the theorem. □

Remark. Here again, Theorem 6.1 can be formulated as a convergence of rescaled random trees in the Gromov-Hausdorff topology. For every $p\ge1$, let $\mathcal{T}^{\langle p\rangle}$ be a μ-Galton-Watson tree conditioned to have exactly p vertices, and let d_{gr} be the usual graph distance on $\mathcal{T}^{\langle p\rangle}$. Then Theorem 6.1 and the arguments discussed at the end of the previous section entail that

$$(\mathcal{T}^{\langle p\rangle},\frac{\sigma}{2\sqrt{p}}d_{gr}) \xrightarrow[p\to\infty]{\rm (d)} (\mathbf{T_e},d_{\mathbf{e}})$$

in the sense of the Gromov-Hausdorff topology. The limiting space $(\mathbf{T_e},d_{\mathbf{e}})$, which is the tree coded by a normalized Brownian excursion $\mathbf{e}$, coincides with Aldous' Continuum Random Tree [1, 2], up to an unimportant scaling factor 2. A similar remark will apply to Theorem 7.1 in the next section.

An application. Let us mention a typical application of Theorem 6.1, which was already discussed by Aldous [1]. Let $\mathcal{T}^{\langle p\rangle}$ be a μ-Galton-Watson tree conditioned to have exactly p vertices. As an immediate application of Theorem 6.1, we have

$$\frac{\sigma}{2\sqrt{p}}\,\mathrm{ht}(\mathcal{T}^{\langle p\rangle}) \xrightarrow[p\to\infty]{\rm (d)} \max_{0\le t\le 1}\mathbf{e}_t\ .$$

Thus, for every $x>0$,

$$\lim_{p\to\infty} P\Big[\mathrm{ht}(\mathcal{T}^{\langle p\rangle})>\frac{2\sqrt{p}}{\sigma}\,x\Big]=P\Big[\max_{0\le t\le 1}\mathbf{e}_t>x\Big].$$

The right-hand side is known (see e.g. Chung [8]) in the form of a series,

$$P\Big[\max_{0\le t\le 1} \mathbf{e}_t > x\Big] = 2\sum_{k=1}^{\infty}(4k^2x^2-1)\,\exp(-2k^2x^2). \tag{36}$$

In the case of the geometric distribution $\mu(k)=2^{-k-1}$, the tree $\mathcal{T}^{\langle p\rangle}$ is uniformly distributed over the set of all rooted ordered trees with p vertices (this follows from (4)). Thus the preceding considerations give the asymptotic behavior of the proportion of those trees with p vertices which have height greater than $y\sqrt{p}$, for any $y>0$. By letting μ be the binary distribution $\mu=\frac{1}{2}(\delta_0+\delta_2)$, respectively the Poisson distribution with parameter 1, we get a similar result for binary trees, resp. rooted Cayley trees on p vertices. Interestingly, the limiting distribution for the height of random trees in the right-hand side of (36) had been derived by Flajolet and Odlyzko [12] using analytic methods, before Aldous' theorem was proved. Obviously the interpretation in terms of Brownian excursions provides a very satisfactory explanation for the appearance of this distribution.

7 Conditioning the tree to have a fixed height

Recall that $\mathrm{ht}(\mathcal{T})$ denotes the height of the tree $\mathcal{T}$. The next result is an analogue of Theorem 6.1.

Theorem 7.1 *For every integer $p\ge 1$, let $C^{\{p\}}=(C^{\{p\}}_t)_{t\ge 0}$ be the contour function of a μ-Galton-Watson tree conditioned to have height equal to p. Then, the law of*

$$\Big(\frac{\sigma}{2p}\,C^{\{p\}}_{2p^2t}\Big)_{t\ge 0}$$

converges to $\mathbf{n}(de\,|\max(e)=\frac{\sigma}{2})$ as $p\to\infty$, in the sense of weak convergence of probability measures on E.

Proof. Let $\mathcal{T}$ be a μ-Galton-Watson tree. We denote the contour function of the tree $\mathcal{T}$ by $(C^{\mathcal{T}}_t)_{t\ge 0}$, and we set

$$C^{(p)}_t = \frac{1}{p}\,C^{\mathcal{T}}_{2p^2t}.$$

For every integer $k\ge 0$, we denote the contour function of $\mathcal{T}$ "truncated at the kth generation" by $(C^k_t)_{t\ge 0}$. This is simply the contour function of the truncated tree $\mathcal{T}^k := \{u\in\mathcal{T} : |u|\le k\}$. We fix $a\in\,]0,1[$ and we set

$$C^{[ap],(p)}_t = \frac{1}{p}\,C^{[ap]}_{2p^2t}.$$

Let $\varepsilon>0$. As a consequence of Theorem 5.1 (cf Example (b) in Section 5), we know that the law of $(C^{(p)}_t)_{t\ge 0}$ under the conditional measure $P[\cdot\,|\,p<\mathrm{ht}(\mathcal{T})<(1+\varepsilon)p]$ converges towards the law of $\frac{2}{\sigma}\,e$ under $\mathbf{n}(de\,|\,\frac{\sigma}{2}<\max(e)<(1+\varepsilon)\frac{\sigma}{2})$ as $p\to\infty$.

If $f\in E$ and $b\ge 0$, define the truncation of f at level b to be the function

$$\mathrm{Tr}_b(f) = f\circ\eta^f_b$$

where for every $t\ge 0$,

$$\eta^f_b(t) = \inf\{s\ge 0 : \int_0^s dr\,\mathbf{1}_{[0,b]}(f(r)) > t\}.$$

Note that $\mathrm{Tr}_b(f) \in E$. Furthermore, if f_n is a sequence in E converging to f, and if b_n is a sequence of nonnegative reals converging to b, then the condition $\int_0^\infty \mathbf{1}_{\{f(t)=b\}} dt = 0$ ensures that $\mathrm{Tr}_{b_n}(f_n)$ converges to $\mathrm{Tr}_b(f)$ in E.

By definition

$$C^{[ap],(p)} = \mathrm{Tr}_{[ap]/p}(C^{(p)}).$$

Thus the preceding considerations imply that the law of $(C_t^{[ap],(p)})_{t\geq 0}$ under the conditional probability measure $P[\cdot \,|\, p < \mathrm{ht}(\mathcal{T}) < (1+\varepsilon)p]$ converges as $p \to \infty$ towards the law of $\mathrm{Tr}_a(\frac{2}{\sigma}\, e)$ under $\mathbf{n}(de \,|\, \frac{\sigma}{2} < \max(e) < (1+\varepsilon)\frac{\sigma}{2})$. Notice that when ε is small, the latter law is close to the law of $\mathrm{Tr}_a(\frac{2}{\sigma}\, e)$ under $\mathbf{n}(de \,|\, \max(e) = \frac{\sigma}{2})$ (this can be seen for instance by using the scaling properties of the Itô measure).

We next need to compare the law of the process $(C_t^{[ap],(p)})_{t\geq 0}$ under $P[\cdot \,|\, p < \mathrm{ht}(\mathcal{T}) < (1+\varepsilon)p]$ with the law of the same process under $P[\cdot \,|\, \mathrm{ht}(\mathcal{T}) = p]$.

Lemma 7.2 *Let F be a bounded measurable function on E. Then,*

$$\lim_{\varepsilon\to 0} \Big(\limsup_{p\to\infty} \Big| E[F((C_t^{[ap],(p)})_{t\geq 0}) \,|\, p < \mathrm{ht}(\mathcal{T}) < (1+\varepsilon)p] - E[F((C_t^{[ap],(p)})_{t\geq 0}) \,|\, \mathrm{ht}(\mathcal{T}) = p] \Big| \Big) = 0.$$

Proof. Recall our notation $\mathcal{T}^{[ap]}$ for the tree $\mathcal{T}$ truncated at generation $[ap]$. We will prove that, for any uniformly bounded sequence $(f_p)_{p\geq 1}$ of functions defined on the set $\mathbb{T}$,

$$\lim_{\varepsilon\to 0} \Big(\limsup_{p\to\infty} \Big| E[f_p(\mathcal{T}^{[ap]}) \,|\, p < \mathrm{ht}(\mathcal{T}) < (1+\varepsilon)p] - E[f_p(\mathcal{T}^{[ap]}) \,|\, \mathrm{ht}(\mathcal{T}) = p] \Big| \Big) = 0. \tag{37}$$

Since $C^{[ap]}$ is a function of the tree $\mathcal{T}^{[ap]}$, the result of the lemma will immediately follow. Without loss of generality, we assume that all functions $|f_p|$ are bounded by 1. We denote by $Z_{[ap]}$ the number of individuals of the tree $\mathcal{T}$ at generation $[ap]$. We first observe that for every $\delta > 0$ we can choose a constant K large enough so that, for every $\varepsilon \in (0,1)$,

$$\limsup_{p\to\infty} P[Z_{[ap]} > Kp \,|\, p < \mathrm{ht}(\mathcal{T}) < (1+\varepsilon)p] \leq \delta$$

and

$$\limsup_{p\to\infty} P[Z_{[ap]} > Kp \,|\, \mathrm{ht}(\mathcal{T}) = p] \leq \delta.$$

These estimates are easily obtained by using the branching property of the tree $\mathcal{T}$ at generation $[ap]$ together with our estimates (5) and (6). We leave details to the reader.

Thanks to the preceding considerations, it is enough to prove that, for every fixed $K > 0$,

$$\lim_{\varepsilon\to 0} \Big(\limsup_{p\to\infty} \Big| E[f_p(\mathcal{T}^{[ap]})\, \mathbf{1}_{\{Z_{[ap]}\leq Kp\}} \,|\, p < \mathrm{ht}(\mathcal{T}) < (1+\varepsilon)p] - E[f_p(\mathcal{T}^{[ap]})\, \mathbf{1}_{\{Z_{[ap]}\leq Kp\}} \,|\, \mathrm{ht}(\mathcal{T}) = p] \Big| \Big) = 0. \tag{38}$$

To simplify notation, we set

$$\gamma(\ell) = P[\mathrm{ht}(\mathcal{T}) > \ell]$$

for every integer $\ell \geq 0$. We also write $\mathcal{T}_1, \mathcal{T}_2, \ldots$ for a sequence of independent μ-Galton-Watson trees. By applying the branching property at generation $[ap]$, we have

$$E[f_p(\mathcal{T}^{[ap]})\, \mathbf{1}_{\{\mathrm{ht}(\mathcal{T})=p\}}] = E[f_p(\mathcal{T}^{[ap]})\, \Delta_p(Z_{[ap]})]$$

where, for every integer $k \geq 0$,

$$\Delta_p(k) = P[\max(\mathrm{ht}(\mathcal{T}_1), \ldots, \mathrm{ht}(\mathcal{T}_k)) = p - [ap]\,] = (1 - \gamma(p - [ap]))^k - (1 - \gamma(p - [ap] - 1))^k.$$

Similarly,

$$E[f_p(\mathcal{T}^{[ap]})\, \mathbf{1}_{\{p < \mathrm{ht}(\mathcal{T}) < (1+\varepsilon)p\}}] = E[f_p(\mathcal{T}^{[ap]})\, \Delta_p^\varepsilon(Z_{[ap]})],$$

where

$$\Delta_p^\varepsilon(k) = (1 - \gamma(p - [ap] + [\varepsilon p]))^k - (1 - \gamma(p - [ap]))^k.$$

Hence, we have, for every $p \geq 1$,

$$\begin{aligned}
&E[f_p(\mathcal{T}^{[ap]})\, \mathbf{1}_{\{Z_{[ap]} \leq Kp\}} \,|\, p < \mathrm{ht}(\mathcal{T}) < (1+\varepsilon)p] - E[f_p(\mathcal{T}^{[ap]})\, \mathbf{1}_{\{Z_{[ap]} \leq Kp\}} \,|\, \mathrm{ht}(\mathcal{T}) = p] \\
&= E\Big[f_p(\mathcal{T}^{[ap]})\; \mathbf{1}_{\{0 < Z_{[ap]} \leq Kp\}} \Big(\frac{\Delta_p(Z_{[ap]})}{P[\mathrm{ht}(\mathcal{T}) = p]} - \frac{\Delta_p^\varepsilon(Z_{[ap]})}{P[p < \mathrm{ht}(\mathcal{T}) < (1+\varepsilon)p]} \Big)\Big].
\end{aligned}$$

Now recall our estimates (5) and (6), and also note that

$$P[p < \mathrm{ht}(\mathcal{T}) < (1+\varepsilon)p] \underset{p\to\infty}{\sim} \frac{2}{\sigma^2 p} - \frac{2}{\sigma^2(1+\varepsilon)p} = \frac{2\varepsilon}{\sigma^2(1+\varepsilon)p}.$$

Elementary calculations show that

$$\sup_{1 \leq k \leq Kp} \Big| \frac{\Delta_p(k)}{P[\mathrm{ht}(\mathcal{T}) = p]} - \frac{k}{(1-a)^2} \exp\Big(- \frac{2k}{\sigma^2(1-a)p} \Big) \Big| = o(p)$$

as $p \to \infty$, and similarly,

$$\sup_{1 \leq k \leq Kp} \Big| \frac{\Delta_p^\varepsilon(k)}{P[p < \mathrm{ht}(\mathcal{T}) < (1+\varepsilon)p]} - \frac{\sigma^2(1+\varepsilon)p}{2\varepsilon} \Big(\exp\Big(\frac{2k\varepsilon}{\sigma^2(1-a)^2 p} \Big) - 1 \Big) \exp\Big(- \frac{2k}{\sigma^2(1-a)p} \Big) \Big| = o(p)$$

as $p \to \infty$. Noting that

$$E[f_p(\mathcal{T}^{[ap]})\, \mathbf{1}_{\{Z_{[ap]} > 0\}}] = O(\frac{1}{p})$$

(by (5)), we get that

$$\begin{aligned}
&\limsup_{p\to\infty} \Big| E[f_p(\mathcal{T}^{[ap]})\, \mathbf{1}_{\{Z_{[ap]} \leq Kp\}} \,|\, p < \mathrm{ht}(\mathcal{T}) < (1+\varepsilon)p] - E[f_p(\mathcal{T}^{[ap]})\, \mathbf{1}_{\{Z_{[ap]} \leq Kp\}} \,|\, \mathrm{ht}(\mathcal{T}) = p] \Big| \\
&\leq \limsup_{p\to\infty} \Big(P[Z_{[ap]} > 0] \times \sup_{1 \leq k \leq Kp} \Big| \frac{k}{(1-a)^2} - \frac{\sigma^2(1+\varepsilon)p}{2\varepsilon} \Big(\exp\Big(\frac{2k\varepsilon}{\sigma^2(1-a)^2 p} \Big) - 1 \Big) \Big| \Big).
\end{aligned}$$

Using once again the fact that $P[Z_{[ap]} > 0] = P[\mathrm{ht}(\mathcal{T}) \geq [ap]\,] = O(1/p)$, we see that the right-hand side of the last display can be made arbitrarily small by choosing $\varepsilon > 0$ small. This completes the proof of (38) and of the lemma. $\square$

From Lemma 7.2 and the considerations preceding the statement of this lemma, we get that, for every fixed $a \in (0,1)$, the law of the process $(C_t^{[ap],(p)})_{t \geq 0}$ under $P[\cdot \,|\, \mathrm{ht}(\mathcal{T}) = p]$ converges as $p \to \infty$ towards the law of $\mathrm{Tr}_a(\frac{2}{\sigma}\, e)$ under $\mathbf{n}(de \,|\, \max(e) = \frac{\sigma}{2})$. Clearly, the latter law is close to the law of $\frac{2}{\sigma}\, e$ under $\mathbf{n}(de \,|\, \max(e) = \frac{\sigma}{2})$ when a is close to 1. So in order to complete the proof of Theorem 7.1, we only need to verify that the law of $(C_t^{(p)})_{t \geq 0}$ under $P[\cdot \,|\, \mathrm{ht}(\mathcal{T}) = p]$ is close to the law of $(C_t^{[ap],(p)})_{t\geq 0}$ under the same probability measure, when a is close to 1 and p is large. More precisely, recalling our notation d for the distance on E, it suffices to verify that, for every $\delta > 0$,

$$\lim_{a\uparrow 1}\Big(\limsup_{p\to\infty} P\Big[d((C_t^{[ap],(p)})_{t\geq 0}),(C_t^{(p)})_{t\geq 0}) > \delta \,\Big|\, \mathrm{ht}(\mathcal{T}) = p\Big]\Big) = 0. \tag{39}$$

From the relations between $C^{(p)}$ and the truncated function $C^{[ap],(p)}$, we see that (39) will follow if we can prove that

$$\lim_{a\uparrow 1}\Big(\limsup_{p\to\infty} P\Big[\int_0^\infty dt\,\mathbf{1}_{\{a\leq C_t^{(p)}\leq 1\}} > \delta \,\Big|\, \mathrm{ht}(\mathcal{T}) = p\Big]\Big) = 0. \tag{40}$$

However, by arguments that we already used in the proof of Lemma 7.2, it is easy to verify that, for every fixed $\eta > 0$, we can choose a constant M large enough so that, for every $p \geq 1$,

$$P\Big[\exists k \in \{[ap],[ap]+1,\ldots,p\} : Z_k \geq Mp \,\Big|\, \mathrm{ht}(\mathcal{T}) = p\Big] \leq \eta,$$

where Z_k denotes the number of vertices of $\mathcal{T}$ at generation k. The estimate (40) readily follows. This completes the proof of Theorem 7.1. $\square$

An application. Let $\mathcal{T}^{\{p\}}$ be a μ-Galton-Watson tree conditioned to have height equal to p. As a consequence of Theorem 7.1, we get

$$\frac{1}{p^2}\,\#\mathcal{T}^{\{p\}} \xrightarrow[p\to\infty]{(d)} \zeta(e^{\{\sigma/2\}})$$

where $e^{\{x\}}$ stands for a Brownian excursion conditioned to have height x, and $\zeta(e)$ denotes the duration of e as in Section 2. The Laplace transform of the limiting distribution can be computed from the Williams decomposition of the Brownian excursion with a fixed height (see e.g. [25, Chapter XII]): For every $x > 0$, and $\lambda > 0$,

$$E[\exp(-\lambda\zeta(e^{\{x\}}))] = \Big(\frac{x\sqrt{2\lambda}}{\sinh(x\sqrt{2\lambda})}\Big)^2.$$

The distribution of $\zeta(e^{\{x\}})$ is closely related to that of the maximum of a normalized excursion, which appears in (36). See [6] (in particular formula (3k) in [6, Théorème 3.4]) and the references therein for more information about these distributions.

References

[1] Aldous, D. (1991) The continuum random tree II: An overview. In: *Stochastic Analysis* (M.T. Barlow, N.H. Bingham eds), pp. 23-70. Cambridge University Press, Cambridge.

[2] Aldous, D. (1993) The continuum random tree III. *Ann. Probab.* **21**, 248-289.

[3] Aldous, D. (1993) Tree-based models for random distribution of mass. *J. Stat. Phys.* **73**, 625-641.

[4] Athreya, K.B., Ney, P.E. (1972) *Branching Processes.* Springer, Berlin.

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

[6] Biane, P., Yor, M. (1987) Valeurs principales associées aux temps locaux browniens. *Bull. Sci. Math.* **111**, 23-101.

[7] Burdzy, K. (1987) *Multidimensional Brownian Excursions and Potential Theory.* Longman, London.

[8] Chung, K.L. (1976) Excursions in Brownian motion. *Ark. Mat.* **14**, 155-177.

[9] Duquesne, T., Le Gall, J.F. (2002) *Random Trees, Lévy Processes and Spatial Branching Processes. Astérisque* **281**.

[10] Ethier, S.N., Kurtz, T. (1985) *Markov Processes: Characterization and Convergence.* Wiley, New York.

[11] Evans, S.N. (2007) *Probability and Real Trees. Lecture Notes Math.* **1920**. Springer, Berlin.

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

[13] Geiger, J., Kersting, G. (1999) The Galton-Watson tree conditioned on its height. *Proceedings 7th Vilnius Conference on Probability and Mathematical Statistics*, pp.277-286. VSP, Utrecht.

[14] Itô, K. (1969) Point processes and their applications to Markov processes. Lecture Notes Kyoto University.

[15] Itô, K. (1972) Poisson point processes attached to Markov processes. *Proceedings Sixth Berkeley Symposium on Mathematical Statistics and Probability*, vol. III, pp.225-239. University of California Press, Berkeley.

[16] Kesten, H., Pittel, B. (1996) A local limit theorem for the number of nodes, the height, and the number of final leaves in a critical branching process tree. *Random Structures Algorithms* **8**, 243-299.

[17] Le Gall, J.F. (1991) Brownian excursions, trees and measure-valued processes. *Ann. Probab.* **19**, 1399-1439.

[18] Le Gall, J.F. (1999) *Spatial Branching Processes, Random Snakes and Partial Differential Equations.* Birkhäuser, Boston.

[19] Le Gall, J.F. (2005) Random trees and applications. *Probability Surveys* **2**, 245-311.

[20] Lévy, P. (1948) *Processus Stochastiques et Mouvement Brownien.* Gauthier-Villars, Paris.

[21] Marckert, J.F., Mokkadem, A. (2003) The depth first processes of Galton-Watson trees converge to the same Brownian excursion. *Ann. Probab.* **31**, 1655-1678.

[22] Meyer, P.A. (1971) Processus de Poisson ponctuels, d'après K.Itô. Séminaire de Probabilités V. *Lecture Notes Math.* **191**. Springer, Berlin.

[23] Pitman, J. (2006) *Combinatorial Stochastic Processes. Lecture Notes Math.* **1875**. Springer, Berlin.

[24] Pitman, J., Yor, M. (2007) Itô's excursion theory and its applications. *Japanese J. Math.* **2**, 83-96.

[25] Revuz, D., Yor, M. (1992) *Continuous Martingales and Brownian Motion.* Springer, Berlin.

[26] Roynette, B., Yor, M. (2009) *Penalizing Brownian Paths. Lecture Notes Math.* **1969**. Springer, Berlin.

[27] Spitzer, F. (2001) *Principles of Random Walk.* Springer, Berlin.

[28] Stanley, R.P. (1999) *Enumerative Combinatorics, Vol.2.* Cambridge University Press, Cambridge.

[29] Watanabe, S. (2010) Itô's theory of excursion point processes and its developments. In this volume.

Département de mathématiques, Université Paris-Sud, 91405 ORSAY Cédex, FRANCE
e-mail: *jean-francois.legall@math.u-psud.fr*