Asymptotic Genealogy of a Critical Branching Process

Lea Popovic
Department of Statistics
367 Evans Hall
U.C. Berkeley CA 94720
lea@stat.berkeley.edu

October,2002

Abstract
Let $\mathcal{T}_{t,n}$ be a continuous-time critical branching process conditioned to have population n at time t. Consider $\mathcal{T}_{t,n}$ as a random rooted tree with edge-lengths. We define the genealogy $\mathcal{G}(\mathcal{T}_{t,n})$ of the population at time t to be the smallest subtree of $\mathcal{T}_{t,n}$ containing all the edges at a distance t from the root. We also consider a Bernoulli(p) sampling process on the leaves of $\mathcal{T}_{t,n}$, and define the p-sampled history $\mathcal{H}_p(\mathcal{T}_{t,n})$ to be the smallest subtree of $\mathcal{T}_{t,n}$ containing all the sampled leaves at a distance less than t from the root. We first give a representation of $\mathcal{G}(\mathcal{T}_{t,n})$ and $\mathcal{H}_p(\mathcal{T}_{t,n})$ in terms of point-processes, and then provide their asymptotic behavior as $n \to \infty$, $\frac{t}{n} \to t_0$, and $np \to p_0$. The resulting asymptotic processes are related to a Brownian excursion conditioned to have local time at t_0 equal to 1, sampled at times of a Poisson($\frac{p_0}{2}$) process.

Keywords Galton-Watson process, random tree, point process, Brownian excursion, genealogy

AMS subject classification xxx

1 Introduction

In this paper we study the genealogical structure of a continuous-time critical branching process conditioned on its population size at a given time t. By genealogical structure we mean a particular subtree of the family tree of the branching process up to time t. Within the family tree we consider all the extant individuals at time t, and a subset of the extinct individuals independently sampled with a given chance p. The subtree we consider is the smallest one containing all the common ancestors of the extant individuals and all the sampled extinct individuals. We introduce a point-process representation of this genealogical subtree and derive its law. Our main result is the asymptotic behavior of such point-processes, and their close relationship to a conditioned Brownian excursion.

The relationship between random trees and Brownian excursions has been much explored in the literature. We here note only a small selection that is most relevant to the work in this paper. The appearance of continuous-time critical branching processes embedded in the structure of a Brownian excursion was noted by Neveu-Pitman and Le Gall ([Ne-Pi,89b],[Ne-Pi,89a], and [LG,89]). The construction of an infinite tree within a Brownian excursion, which is in some sense a limit of the trees from the work of Neveu-Pitman, has been considered by Abraham ([Ab,92],[Ab-Ma,92]) and Le Gall ([LG,91]). The convergence of critical branching processes conditioned on total population size to a canonical tree within a Brownian excursion (the *Continuum Random Tree*) was introduced by Aldous ([Al,93]). We present a connection of the asymptotic results in this paper with the above mentioned results.

Some aspects of the genealogy of critical Galton-Watson trees conditioned on non-extinction have been studied in ([Du,78]), without the use of family trees. It has also been studied within the context of super-processes (e.g. [LG,91]).

The paper is structured in two parts. In Section 2 we give a precise definition of the genealogical point-process representing the common ancestry of the extant individuals, and provide its exact law and asymptotic behavior (Theorem 5). Then in Section 3 we give the definition of the corresponding genealogical point-process that includes the sampled extinct individuals as well, provide its exact law and asymptotic behavior (Theorem 9). Section **??** describes the connections of our asymptotic results, with random forests embedded in Brownian motion.

2 Genealogy of extant individuals

Consider a continuous-time critical branching process $\mathcal{T}$, with initial population size 1. In such a process each individual has an Exponential(rate 1) lifetime, in the course of which it gives birth to new individuals according to a Poisson(rate 1) process, with all the individuals living and reproducing independently of each other. Let $\mathcal{T}_{t,n}$ be the process $\mathcal{T}$ conditioned to have population size n at time t. We use the same notation ($\mathcal{T}$ and $\mathcal{T}_{t,n}$) for the random trees with edge-lengths that are family trees of these processes. We picture these family trees as rooted planar trees with the following conventions: each individual is represented with edges whose length is equal to the individuals lifetime, each new individual is attached on the right of the edge of its parent at the point corresponding to its time of birth, with the parent continuing on the left. Such trees are identified by their shape and by the collection of individuals birth times and lifetimes. We shall label the vertices in the tree in a depth-first search manner. An example of a realization of $\mathcal{T}_{t,n}$ is shown in Figure 1

Remark. We point out that $\mathcal{T}$ is closely related to the continuous-time critical binary-branching Galton-Watson process. The difference between the two is in the identities of individuals. If at each branching event with two offspring we imposed the identification of one of the offspring with its parent we would obtain the same random tree as the family tree of $\mathcal{T}$.

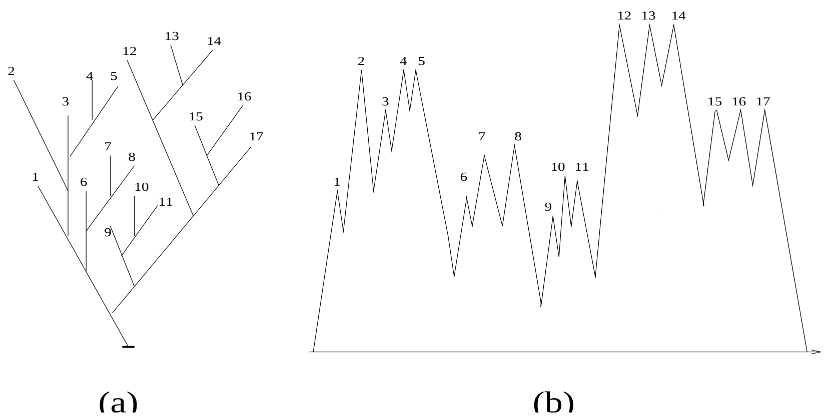

Figure 1: (a) A realization of the tree $\mathcal{T}_{t,n}$ whose population at time t is $N(t) = 5$. The leaves are labeled in depth-first search manner; (b) The contour $\mathcal{C}_{\mathcal{T}_{t,n}}$ process of the tree $\mathcal{T}_{t,n}$. Each local maximum of $\mathcal{C}_{\mathcal{T}_{t,n}}$ corresponds to the height of a leaf of $\mathcal{T}_{t,n}$.

Consider also the contour process $\mathcal{C}_{\mathcal{T}}$ induced by the random tree $\mathcal{T}$. The contour process of a rooted planar tree is a continuous function giving the distance from the root of a unit-speed depth-first search of the tree. Such a process starts at the root of the tree, traverses each edge of the tree once upwards and once downwards following the depth-first search order of the vertices, and ends back at the root of the tree. The contour process consists of line segments of slope $+1$ (the *rises*), and line segments of slope -1 (the *falls*). The unit speed of the traversal insures that the height levels in the process are equivalent to distances from the root in the tree, i.e. times in the branching process. A contour process $\mathcal{C}_{\mathcal{T}_{t,n}}$ induced by a realization of $\mathcal{T}_{t,n}$ is shown in Figure 1 (b). For a formal definition of planar trees with edge lengths, contour processes and their many useful properties see [Pi,02] §6.1 .

Let the *genealogy* of extant individuals at time t be defined as the smallest subtree of the family tree containing all the edges representing these individuals. The genealogy $\mathcal{G}(\mathcal{T}_{t,n})$ at t is thus an n-leaf tree. We introduce a point-process representation of this genealogical tree. The point of doing so is that we get an object that is much simpler to analyze and gives much clearer asymptotic results, than we could have made in the original space of trees with edge-lengths. Informally, think of forming this point-process by taking the height of the branching points of the genealogical tree $\mathcal{G}(\mathcal{T}_{t,n})$ in the order they have as vertices in the tree. For convenience reasons (which will be clear in the asymptotic consideration) we keep track of the heights of the branching points as distances from level t, and for graphical convenience we offset the index of the points by a $\frac{1}{2}$. The genealogical subtree of the tree from Figure 1, and its point-process representation are shown in Figure 2. Formally let $A_i, 1 \le i \le n-1$ be the times of branch points in the tree $\mathcal{G}(\mathcal{T}_{t,n})$, in the order induced from the depth-first search ordering of the vertices in the tree, and $\tau_i = t - A_i$. Set $l_i = i + \frac{1}{2}$, then

Definition. Let the *Genealogical point-process* $\Pi_{t,n}$ be the random finite set

$$\Pi_{t,n} = \{(l_i, \tau_i) : 1 \le i \le n-1,\, 0 < \tau_i < t\} \tag{1}$$

We can equally obtain $A_i, 1 \le i \le n-1$ from the contour process $\mathcal{C}_{\mathcal{T}_{t,n}}$. The ith individual extant at t corresponds to an up-crossing U_i and the subsequent down-crossing D_i of level t. The branch-points $A_i, 1 \le i \le n-1$ of $\mathcal{G}(\mathcal{T}_{t,n})$ correspond to the levels of lowest local minima of the part excursions of $\mathcal{C}_{\mathcal{T}_{t,n}}$ below level t, i.e. $A_i = \inf\{\mathcal{C}_{\mathcal{T}_{t,n}}(u) : D_i < u < U_{i+1}\}$. The usefulness of this observation is to allow us to use some nice properties of the law of $\mathcal{C}_{\mathcal{T}_{t,n}}$ to derive the law of $\Pi_{t,n}$.

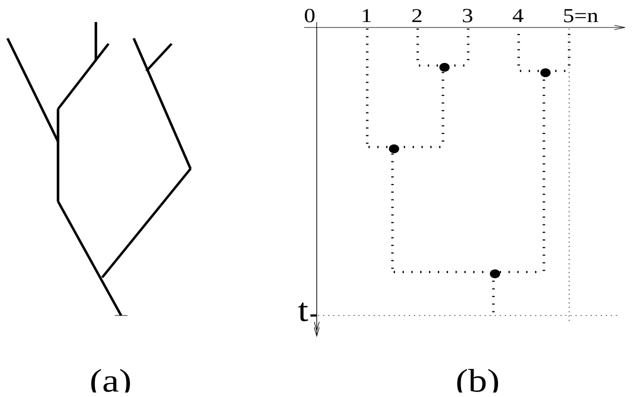

Figure 2: (a) The genealogical tree $\mathcal{G}(\mathcal{T}_{t,n})$ of the extant individuals at time t; (b) The point-process $\Pi_{t,n}$ representation of $\mathcal{G}(\mathcal{T}_{t,n})$. The dotted lines show reconstruction of $\mathcal{G}(\mathcal{T}_{t,n})$ from its point-process.

We now recall the result of Neveu-Pitman-Le Gall, regarding the law of the contour process $\mathcal{C}_{\mathcal{T}}$, referring the reader to either [LG,89], or to [Ne-Pi,89a] for its proof.

Lemma 1. *In the contour process $\mathcal{C}_\mathcal{T}$ of a critical branching process $\mathcal{T}$ the sequence of rises and falls (up to the last fall) has the same distribution as a sequence of independent Exponential(rate 1) variables stopped one step before the sum of successive rises and falls becomes negative (the last fall is then set to equal this sum).*

From the memoryless property of the exponential distribution we immediately get the following corollary.

Corollary 2. *For the contour process $\mathcal{C}_\mathcal{T}$ the process $X_\mathcal{T} = (\mathcal{C}_\mathcal{T}, slope[\mathcal{C}_\mathcal{T}])$ is a time-homogeneous strong Markov process on $\mathbb{R}^+ \times \{+1,-1\}$ stopped when it first reaches $(0,-1)$.*

The law of the Genealogical point-process $\Pi_{t,n}$ can now easily be given with some excursion theory for Markov processes.

Lemma 3. *For any $t > 0$, the random set $\Pi_{t,n}$ is a simple point-process on $\{\frac{3}{2},\dots,n-\frac{1}{2}\} \times (0,t)$ with intensity measure*

$$\nu_{t,n}\big(\{\frac{1}{2}+i\} \times d\tau\big) = \frac{d\tau}{(1+\tau)^2}\frac{1+t}{t} \tag{2}$$

Proof. Consider the Markov process $X_\mathcal{T} = (\mathcal{C}_\mathcal{T}, \mathrm{slope}[\mathcal{C}_\mathcal{T}])$ until the first hitting time $U_{(0,-1)} = \inf\{u \geq 0 : X_\mathcal{T}(u) = (0,-1)\}$, and consider its excursions from the point $(t,+1)$ using the distribution of $\mathcal{C}_\mathcal{T}$ given by Lemma 1. Clearly $\mathbf{P}_{(t,+1)}\big[\inf\{u > 0 : X_\mathcal{T}(u) = (t,+1)\} > 0\big] = 1$, and the visits to $(t,+1)$ at times $U_1 = \inf\{u \geq 0 : X_\mathcal{T}(u) = (t,+1)\}, U_i = \inf\{u > U_{i-1} : X_\mathcal{T}(u) = (t,+1)\}, i > 1$ is discrete. The excursions of $X_\mathcal{T}$ for $i \geq 1$ are

$$\varepsilon_i(u) = X_\mathcal{T}(U_i + u), \text{ for } u \in [0, U_{i+1} - U_i), \text{ and } \varepsilon_i(u) = (0,+1) \text{ else}$$

the number of visits in an interval $(0,u]$ is

$$l(0) = 0,\ l(u) = \sup\{i > 0 : u > U_i\}, u > 0$$

the total number prior to $U_{(0,-1)}$ is $L = \sup\{i \geq 0 : U_{(0,-1)} > U_i\} = l(U_{(0,-1)})$. If $\mathbf{n}$ is the $\mathbf{P}_{(t,+1)}$-law of ε_i, and $\mathcal{E}^{<t}$ is the set of excursions from $(t,+1)$ that return to $(t,+1)$ without reaching $(0,-1)$, and $\mathcal{E}^{>t}$ the set of all others, then (see e.g.[Ro-Wi,87] Vol.2 §VI.50.)

- $\mathbf{P}_{(0,+1)}\big[L \geq i\big] = \big[\mathbf{n}(\mathcal{E}^{<t})\big]^{i-1}$, $i \geq 1$, and $\varepsilon_1, \varepsilon_2, \dots$ are independent
- given that $L \geq i$: the law of $\varepsilon_1, \varepsilon_2, .., \varepsilon_{i-1}$ is $\mathbf{n}(\cdot \cap \mathcal{E}^{<t})/\mathbf{n}(\mathcal{E}^{<t})$
- given that $L = i$: the law of ε_i is $\mathbf{n}(\cdot \cap \mathcal{E}^{>t})/\mathbf{n}(\mathcal{E}^{>t})$

This makes $\{(l(U_i), \varepsilon_i), 1 \leq i \leq L-1\}$ a simple point-process, with the number of points having a Geometric($\mathbf{n}(\mathcal{E}^{>t})$) law, and with each ε_i having the law $\mathbf{n}(\cdot \cap \mathcal{E}^{<t})/\mathbf{n}(\mathcal{E}^{<t})$. This makes it particularly convenient for analyzing $\mathcal{C}_{\mathcal{T}_{t,n}}$,

which is $\mathcal{C}_\mathcal{T}$ conditioned on $L = n$, so then the $n-1$ excursions of $\mathcal{C}_{\mathcal{T}_{t,n}}$ have the law $\mathbf{n}(\cdot \cap \mathcal{E}^{<t})/\mathbf{n}(\mathcal{E}^{<t})$

We have the up-crossing times $U_i, 1 \le i \le n$, then the down-crossing times are $D_i = \inf\{u > U_i : X_\mathcal{T}(u) = (t, -1)\}$. We are interested in the part of the excursions from $(t, +1)$ below level t

$$\varepsilon_i^{<t} = \varepsilon_i(D_i + u), u \in [0, U_{i+1} - D_i), \text{ and } \varepsilon_i^{<t}(u) = (0, +1) \text{ else}$$

We note that the shift and reflection invariance of the transition function of $\mathcal{C}_\mathcal{T}$, as well as its strong Markov property, applied to the law $\mathbf{n}$ for $\varepsilon_i^{<t}$ imply that the law of $\varepsilon_i^+ = t - \varepsilon_i^{<t}$ is the same as the law of $X_\mathcal{T}$. Consequently the law of $t - \inf(\varepsilon_i^{<t}) = \sup(\varepsilon_i^+)$ is the same law as that of $\sup(\mathcal{C}_\mathcal{T})$.

We now recall classical results on the branching process $\mathcal{T}$ ([Fe,68] §XVII.10.11.) which say that for the law of the population size $N(t)$ of $\mathcal{T}$ at time t we have

$$\mathbf{P}\big[N(t) = 0\big] = \frac{t}{1+t}; \;\; \mathbf{P}\big[N(t) = k\big] = \frac{t^{k-1}}{(1+t)^{k+1}}, \text{ for } k \ge 1 \tag{3}$$

hence

$$\mathbf{P}\big[\sup(\mathcal{C}_\mathcal{T}) > t\big] = \mathbf{P}\big[N(t) > 0\big] = \frac{1}{1+t}, \text{ for } t \ge 0 \tag{4}$$

Now for $\mathcal{C}_{\mathcal{T}_{t,n}}$ and for each $1 \le i \le n-1$ we have that $A_i = \inf(\varepsilon_i^{<t})$, and the $\varepsilon_i^{<t}$ are independent with $\varepsilon_i^{<t} \sim \mathbf{n}(\cdot \cap \mathcal{E}^{<t})/\mathbf{n}(\mathcal{E}^{<t})$, hence then each $\tau_i = t - A_i$ has the law

$$\mathbf{P}(\tau_i \in dt) = \frac{1}{(1+\tau)^2}\frac{1+t}{t}, \text{ for } 0 \le \tau \le t \tag{5}$$

□

It would be easy to establish asymptotics for $\Pi_{t,n}$ with a routine calculation, but first we want to give the context in which we see the asymptotic process. We establish a connection with a conditioned Brownian excursion that enables one to derive many more results about the genealogical structure of $\mathcal{T}_{t,n}$ then just $\mathcal{G}(\mathcal{T}_{t,n})$. Asymptotic results for critical Galton-Watson processes conditioned on a "very large" total population size have been established using different techniques. We recall here the result of Aldous ([Al,93] Thm23) which says that its contour process appropriately rescaled converges (as the total population size increases) to a Brownian excursion (doubled in height) conditioned to have length 1. Now note that, if N_{tot} is the total population size of a critical Galton-Watson process, and $N(t)$ its population size at some given time t, then the events $N_{tot} = n$ and $N(t) = n$ are both events of "very small" probability. The first has asymptotic chance $cn^{-\frac{3}{2}}$ as $n \to \infty$, and for $\frac{t}{n} \to t_0$ as $n \to \infty$ the second has asymptotic chance $c(t_0)\, n^{-1}$ (xxx-[Kai,76]). Whereas total population size corresponds to the total length of the contour process, the population size at a particular time t corresponds to the occupation time of the contour process at level t. Hence it is natural to conjecture here that the contour process of a critical Galton-Watson process conditioned on a "very large" population at time t converges to a Brownian excursion conditioned to have local time 1 at level t_0. The asymptotic results we establish in this paper provide some support for its validity.

Let us introduce for the state-space for the asymptotic process the notion of a *nice point-process*. A nice point-process on $[0,1]\times(0,\infty)$ (see [Al,93])§2.8.) is a countably infinite set of points such that

- for any $\delta>0$: $[0,1]\times[\delta,\infty)$ contains only finitely many points
- for any $0\le x<y\le 1, \delta>0$: $[x,y]\times(0,\delta)$ contains at least one point.

We construct a point-process from a Brownian excursion conditioned to have local time 1 at level t_0, in the same manner in which $\Pi_{t,n}$ was constructed from the contour process $\mathcal{C}_{\mathcal{T}_{t,n}}$. Consider a Brownian excursion $\mathcal{B}(u), u\ge 0$. For a fixed $t_0>0$, let $\ell_{t_0}(u), u\ge 0$ be its local time at level t_0 up to time u, with standard normalization of local time as occupation density relative to Lebesgue measure. Let $i_{t_0}(\ell), \ell\ge 0$ be the inverse process of $\ell_{t_0},\ i_{t_0}(\ell)=\inf\{u>0:\ell_{t_0}(u)>\ell\}$. Let $\mathcal{B}_{t_0,1}(u), u\ge 0$ then be the excursion $\mathcal{B}$ conditioned to have total local time ℓ_{t_0} equal to 1 (by which we mean the total local time $\ell_{t_0}(\infty)$). Consider excursions ϵ_ℓ of $\mathcal{B}_{t_0,1}$ below level t_0 indexed by the amount of local time ℓ at t_0 at the time $i_{t_0}(\ell^-)$ of their beginning. For each such excursion let a_ℓ be its infimum, and $t_\ell=t-a_\ell$. Ito's excursion theory insures then that the process $\{(\ell,t_\ell): i_{t_0}(\ell^-)\ne i_{t_0}(\ell)\}$ is well defined, then

Definition. Let the *Continuum Genealogical point-process* $\pi_{t_0,1}$ be the random countably infinite set

$$\pi_{t_0,1}=\{(\ell,t_\ell): i_{t_0}(\ell^-)\ne i_{t_0}(\ell)\} \tag{6}$$

The name of the process will be obvious from the theorem on asymptotic behavior of Genealogical point-processes (see Theorem 5). We show that the state-space for $\pi_{t_0,1}$ is the set of nice point-processes, and establish the law of this process using standard results of excursion theory of Levy-Ito-Williams.

Lemma 4. *The random set $\pi_{t_0,1}$ is a Poisson point-process on $[0,1]\times(0,t_0)$ with intensity measure*

$$\nu(\ell\times\tau)=d\ell\,\frac{d\tau}{\tau^2} \tag{7}$$

The random set $\pi_{t_0,1}$ is a.s. a nice point-processes.

Proof. Consider the path of an (unconditioned) Brownian excursion $\mathcal{B}$ after the first hitting time of t_0, $U_{t_0}=\inf\{u\ge 0:\mathcal{B}(u)=t_0\}$, shifted and reflected about the u-axis

$$\beta(u)=t_0-\mathcal{B}(U_{t_0}+u),\ \text{for } u\ge 0 \tag{8}$$

Let $\ell_0^\beta(u), u\ge 0$ be the local time of β at level 0 up to time u, and let $i_0^\beta(\ell), \ell\ge 0$ be the inverse process of ℓ_0^β, $i_0^\beta(\ell)=\inf\{u>0:\ell_0^\beta(u)>\ell\}$. Then the process

$\beta(u), u \geq 0$ is a standard Brownian motion stopped at the first hitting time of t_0, $U^\beta_{t_0} = \inf\{u \geq 0 : \beta(u) = t_0\}$. The excursions of β from 0 (with a change of sign) are the excursions of $\mathcal{B}$ from t_0, and its local time process ℓ^β_0 is equivalent to the local time process ℓ_{t_0} of $\mathcal{B}$. We are only interested in the excursions of $\mathcal{B}$ below t_0, which are the positive excursions of β defined for $i^\beta_0(\ell^-) \neq i^\beta_0(\ell)$ and $\beta(i^\beta_0(\ell)^+) > 0$ as

$$\epsilon^+_\ell = \beta(i^\beta_0(\ell^-) + u), u \in [0, i^\beta_0(\ell) - i^\beta_0(\ell^-)), \text{ and } \epsilon^+_\ell(u) = 0 \text{ else}$$

Note that the infimum of an excursion of $\mathcal{B}$ below t_0 is equivalent to $t_0 - \sup(\epsilon^+_\ell)$. Standard results of Ito's excursion theory ([Ro-Wi,87] Vol.2 §VI.47.) imply that for a standard Brownian motion β the random set $\{(\ell, \sup(\epsilon^+_\ell)) : i^\beta_0(\ell^-) \neq i^\beta_0(\ell)\}$ is a Poisson point-process on $\mathbb{R}^+ \times \mathbb{R}^+$ with intensity measure $d\ell \, \frac{d\tau}{\tau^2}$.
Now let $L = \inf\{\ell \geq 0 : \sup(\epsilon^+_\ell) \geq t_0)\}$. Then stopping ℓ^β_0 at the hitting time L is equivalent to stopping β at its hitting time $U^\beta_{t_0}$. Define a random set π_{t_0} from the unconditioned Brownian excursion $\mathcal{B}$ in the same manner in which we defined $\pi_{t_0,1}$ from a conditioned Brownian excursion $\mathcal{B}_{t_0,1}$. Recalling the relationship (8) of $\mathcal{B}$ and β, we obtain that π_{t_0} is equivalent to a restriction of $\{(\ell, \sup(\epsilon^+_\ell)) : i^\beta_0(\ell^-) \neq i^\beta_0(\ell)\}$ on the random set $[0, L] \times (0, t_0)$. In other words π_{t_0} is a Poisson point-process on $\mathbb{R}^+ \times \mathbb{R}^+$ with intensity measure $d\ell \, \frac{d\tau}{\tau^2}$ restricted to the random set $[0, L] \times (0, t_0)$.
Next note that the condition $\{\ell_{t_0} = 1\}$ for $\mathcal{B}$ is equivalent to the condition $\{\ell^\beta_0(U^\beta_{t_0}) = 1\}$ for β, which is further equivalent to the condition $\{L = 1\}$ for π_{t_0}. This then establishes that $\pi_{t_0,1} \overset{d}{=} \pi_{t_0} | \{L = 1\}$. Further, the condition $\{L = 1\}$ on π_{t_0} is equivalent to the condition that π_{t_0} has no points in $[0, 1) \times [t, \infty)$ and has a point in $\{1\} \times [t, \infty)$. But since π_{t_0} is Poisson, independence of Poisson random measures on disjoint sets implies that conditioning π_{t_0} on $\{L = 1\}$ will not alter its law on the set $[0, 1] \times (0, t)$. However, since $\pi_{t_0,1}$ is supported precisely on $[0, 1] \times (0, t)$, the above results together establish that $\pi_{t_0,1}$ is a Poisson point-process on $[0, 1] \times (0, t)$ with intensity measure $d\ell \, \frac{d\tau}{d\tau^2}$.
It is now easy to see from the intensity measure of $\pi_{t_0,1}$ that its realizations are a.s. nice point-processes, namely

- for any $\delta > 0$: $d\ell\big([0, 1]\big) \times \frac{d\tau}{d\tau^2}\big([\delta, \infty)\big) = \frac{1}{\delta} < \infty$
- fro any $0 \leq x < y \leq 1$,$\delta > 0$: $d\ell\big([x, y]\big) \times \frac{d\tau}{d\tau^2}\big((0, \delta)\big) = (y - x) \cdot \infty$.

And since $\pi_{t_0,1}$ is Poisson, finiteness of its intensity measure on $[0, 1] \times [\delta, \infty)$ implies that it has a.s. only finitely many points in the set $[0, 1] \times [\delta, \infty)$, while infiniteness of its intensity measure on $[x, y] \times (0, \delta)$ implies that it has a.s. at least one point on the set $[x, y] \times (0, \delta)$. $\square$

The right rescaling for $\mathcal{T}_{t,n}$ is to speed up time by n and to assign mass n^{-1} to each extant individual, which implies rescaling each coordinate of $\Pi_{t,n}$ by n^{-1}. The asymptotic behavior of the rescaled Genealogical point-process

$$n^{-1}\Pi_{t,n} = \{(n^{-1}l_i, n^{-1}\tau_i) : (l_i, \tau_i) \in \Pi_{t,n}\} \tag{9}$$

is now easily established.

Theorem 5. *For any* $\{t_n > 0\}_{n\geq 1}$ *such that* $\frac{t_n}{n} \underset{n\to\infty}{\to} t_0$ *we have*

$$n^{-1}\Pi_{n,t_n} \underset{n\to\infty}{\overset{d}{\Longrightarrow}} \pi_{t_0,1}$$

Remark. We use the notation $\overset{d}{\Longrightarrow}$ for weak convergence of processes.

Proof. Using Lemma 3 and the rescaling (9) we have that $n^{-1}\Pi_{n,t_n}$ is a simple point-process on $\{\frac{3}{2n}, \ldots, 1 - \frac{1}{2n}\} \times (0, \frac{t_n}{n})$ with intensity measure

$$\frac{1}{n}\sum_{i=1}^{n-1} \delta_{\{\frac{1}{2n}+\frac{i}{n}\}}(l)\, \frac{nd\tau}{(1+n\tau)^2}\frac{1+t_n}{t_n} \tag{10}$$

If $\{t_n\}_{n\geq 1}$ is such that $\frac{t_n}{n} \underset{n\to\infty}{\to} t_0$ then clearly the support set of the process $n^{-1}\Pi_{n,t_n}$ converges to $[0,1] \times (0, t_0)$ and its intensity measure converges to $d\ell\, \frac{d\tau}{\tau^2}$. For simple point-processes this is sufficient to insure weak convergence of the process (see e.g.[Bi,99] §12.3.) to a Poisson point-process on $[0,1] \times (0, t_0)$ with intensity measure $d\ell\, \frac{d\tau}{d\tau^2}$. Using Lemma 4, we then established that $n^{-1}\Pi_{n,t_n} \underset{n\to\infty}{\overset{d}{\Longrightarrow}} \pi_{t_0,1}$. □

3 Genealogy of sampled extinct individuals

We now want to extend this genealogical structure to include a proportion of extinct individuals as well. Suppose that each individuals independently has some given chance p of appearing in the genealogical history of the extant individuals. We indicate this by putting a star mark on the leaf of $\mathcal{T}_{t,n}$ corresponding to that individual. An example of a realization of such p-sampling is shown in Figure 3.

Then let the *p-sampled history* of the extant individuals as time t be defined as the smallest subtree of the family tree containing all the edges of the extant individuals and all the leaves of the p-sampled extinct individuals. The p-sampled history $\mathcal{H}_p(\mathcal{T}_{t,n})$ contains the genealogy $\mathcal{G}(\mathcal{T}_{t,n})$. In fact we think of $\mathcal{G}(\mathcal{T}_{t,n})$ as the "main genealogical tree" with the rest of $\mathcal{H}_p(\mathcal{T}_{t,n})$ as the "p-sampled subtrees" attached to this main tree. Also, we construct a point-process representation of $\mathcal{H}_p(\mathcal{T}_{t,n})$ so that it contains $\Pi_{t,n}$ as the "main points" of the process. Informally, think of forming the rest of this point-process by taking the height of the branching points in $\mathcal{H}_p(\mathcal{T}_{t,n})$ at which the p-sampled subtrees get attached to the edges of the main tree $\mathcal{G}(\mathcal{T}_{t,n})$ (again we keep track of these heights as distances from level t, and in the order in between the points $\Pi_{t,n}$ that they have as vertices in a depth-search first ordered tree). The p-sampled

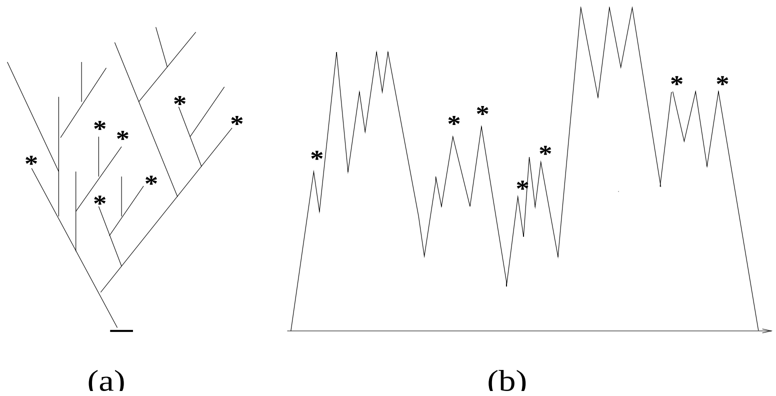

Figure 3: (a) The tree $\mathcal{T}_{t,n}$ with p-sampling on its individuals. The sampled individuals are represented by $*$ marks; (b) Its contour process with the sampling on the corresponding local maxima.

history of the tree from Figure 3, and its point-process representation are shown in Figure 4.

Formally, we define the point-process of $\mathcal{H}_p(\mathcal{T}_{t,n})$ from the contour process $\mathcal{C}_{\mathcal{T}_{t,n}}$. The p-sampling can be equally defined as sampling of the local maxima of $\mathcal{C}_{\mathcal{T}_{t,n}}$. The heights of the branch-points of $\mathcal{G}(\mathcal{T}_{t,n})$ as defined earlier are $A_i = \inf\{\mathcal{C}_{\mathcal{T}_{t,n}}(u) : D_i < u < U_{i+1}\}$, occurring in the contour process $\mathcal{C}_{\mathcal{T}_{t,n}}$ at times $U_{A_i} = \{u \in (D_i, U_{i+1}) : \mathcal{C}_{\mathcal{T}_{t,n}}(u) = A_i\}$. The p-subtree that attaches to $\mathcal{G}(\mathcal{T}_{t,n})$ on the left (right) of the branch-point A_i is defined from the part of the excursion of $\mathcal{C}_{\mathcal{T}_{t,n}}$ below t before (respectively after) time U_{A_i}. In addition, the p-subtrees on the left of the first branching point are defined by the part of $\mathcal{C}_{\mathcal{T}_{t,n}}$ prior to the first up-crossing time U_1, and analogously, the p-subtrees on the right of the last branching point are defined by the part of $\mathcal{C}_{\mathcal{T}_{t,n}}$ after the last down-crossing time D_n (see Figure 3). The part of an excursion of $X_{\mathcal{T}_{t,n}} = (\mathcal{C}_{\mathcal{T}_{t,n}}, \text{slope}[\mathcal{C}_{\mathcal{T}_{t,n}}])$ below t as defined earlier is $\varepsilon_i^{<t} = X_{\mathcal{T}_{t,n}}(D_i + u), u \in [0, U_{i+1} - D_i)$. Let

$$\varepsilon_{i,L}^{<t}(u) = X_{\mathcal{T}_{t,n}}(D_i + u), \;\; \varsigma_{i,L}(u) = \inf_{0 \leq v \leq u} \varepsilon_{i,L}^{<t}(v), \;\; u \in [0, U_{A_i} - D_i)$$

Figure 5 shows $\varepsilon_{i,L}^{<t}$ with its infimum process $\varsigma_{i,L}^{<t}$ (all the following definitions for the right part $\varepsilon_{i,R}^{<t}$ of $\varepsilon_i^{<t}$ are identical up to symmetry about the vertical axis through A_i).

Let $a_{i,L}(j), j \geq 0$ be successive levels of constancy of $\varsigma_{i,L}$, $t_{i,L}(j) = t - a_{i,L}(j)$. For each level of constancy, let $\varepsilon_{i,L}^{<t}(j)$ be the excursion of $\varepsilon_{i,L}^{<t} - \varsigma_{i,L}$ that lies above the level $a_{i,L}(j)$, of $\varsigma_{i,L}$ (see Figure 5). Let $h_{i,L}(j)$ be the height of this excursion, $h_{i,L}(j) = \sup(\varepsilon_{i,L}^{<t}(j))$. Also let $\Upsilon_{i,L}(j)$ be the tree whose contour process is the excursion $\varepsilon_{i,L}^{<t}(j)$. At this point note that all the star marks due to p-sampling are contained in the excursions $\varepsilon_{i,L}^{<t}(j)$, hence are contained in the

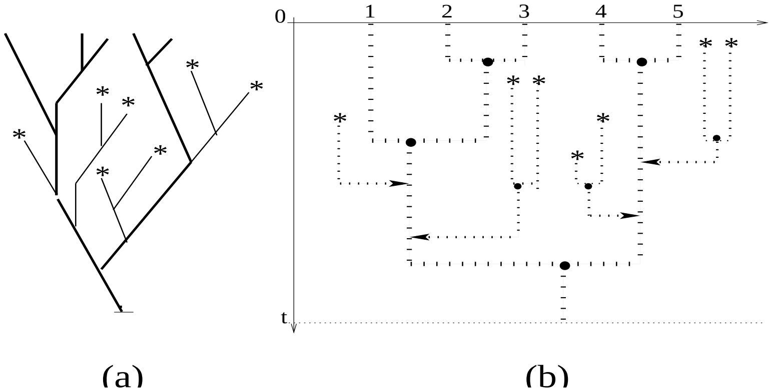

Figure 4: (a) The p-sampled tree $\mathcal{H}_p(\mathcal{T}_{t,n})$, and (b) its point-process representation $\mathcal{P}_p(\Xi_{t,n})$ Each point of $\Pi_{t,n}$ has a left set and a right set of points attached to it, representing the p-sampled subtrees attaching to the left and right of that branch-point.

subtrees $\Upsilon_{i,L}(j)$. We attach to each point (l_i, t_i) of $\Pi_{t,n}$ one "left" set and one (analogously defined) "right" set

$$\mathcal{L}_i = \{(t_{i,L}(j), h_{i,L}(j), \Upsilon_{i,L}(j))\}_{j\geq 0},\ \mathcal{R}_i = \{(t_{i,R}(j), h_{i,R}(j), \Upsilon_{i,R}(j))\}_{j\geq 0} \quad (11)$$

In addition, we define one set $\mathcal{R}_0 = \{(t_{0,R}(j), h_{0,R}(j), \Upsilon_{0,R}(j))\}_{j\geq 0}$, from the first part of $\mathcal{C}_{\mathcal{T}_{t,n}}$: $\varepsilon^{<t}_{0,R}(u) = X_{\mathcal{T}}(u), u \in [0, U_1)$; and also from the last part of $\mathcal{C}_{\mathcal{T}_{t,n}}$: $\varepsilon^{<t}_{n,L}(u) = X_{\mathcal{T}}(D_n + u)$, $u \in [0, U_{(0,-1)} - D_n)$ we likewise define a set $\mathcal{L}_n = \{(t_{n,L}(j), h_{n,L}(j), \Upsilon_{n,L}(j))\}_{j\geq 0}$. To make notation easier we set $\mathcal{L}_0 = \emptyset$, $\mathcal{R}_n = \emptyset$, $(l_0, t_0) = (1, t)$, $(l_n, t_n) = (n, t)$, then

Definition. Let the *Historical point-process* with parameter p be the random set

$$\mathcal{P}_p(\Xi_{t,n}) = \{(l_i, t_i, \mathcal{L}_i, \mathcal{R}_i) : (l_i, t_i) \in \Pi_{t,n},\ 0 \leq i \leq n\} \quad (12)$$

Remark. We have in fact implicitly defined a point-process representation $\Xi_{t,n}$ of an unsampled (i.e.$p = 1$) Historical point-process, the only difference between $\Xi_{t,n}$ and $\mathcal{P}_p(\Xi_{t,n})$ is in the star marks on the leaves in the latter. It will however be clear that for nice asymptotic behavior we need to consider $\mathcal{P}_p(\Xi_{t,n})$ with $p < 1$ (i.e. we can only keep track of a proportion of the extinct individuals).

Let $\mathbf{T}$ denote the space of rooted planar trees with edge-lengths with finitely many leaves, and $\mathbf{\Lambda}^p$ the law on $\mathbf{T}$ induced by the p-sampled trees $\mathcal{T}$. For any $h > 0$, let $\mathbf{\Lambda}^p_h$ be the law induced by restricting $\mathbf{\Lambda}^p$ to the trees $\mathcal{T}$ of height h. The following Lemma describes the law of $\mathcal{P}_p(\Xi_{t,n})$.

Lemma 6. *For any $p \in (0, 1)$, the random set $\mathcal{P}_p(\Xi_{t,n})$ is such that:*
- $\{(l_i, t_i) : 1 \leq i \leq n-1\}$ *is the simple point-process $\Pi_{t,n}$ of Lemma 3,*

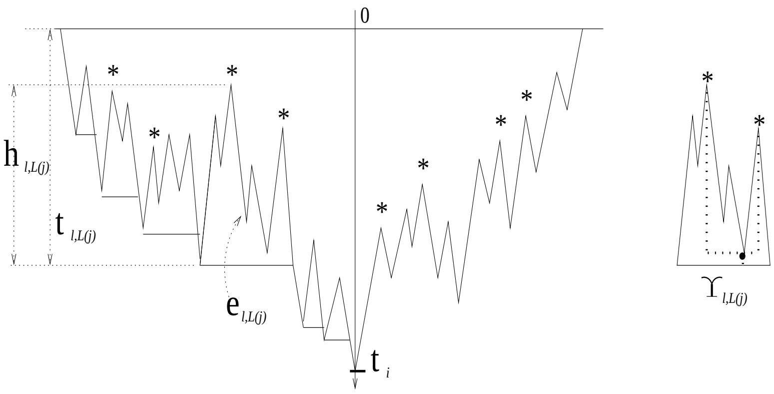

Figure 5: The left half of an excursion of $\mathcal{C}_{\mathcal{T}_{t,n}}$ below t, $\varepsilon^{<t}_{i,L}$, with its supremum process $\varsigma_{i,L}$ whose levels of constancy are $\{t_{i,L}(j)\}_j$, above which lie p-marked subtrees $\{\Upsilon_{i,L}(j)\}_j$ of heights $\{h_{i,L}(j)\}_j$.

• *given* $\{(l_i, t_i), 1 \leq i \leq n-1\}$ *the random sets* $\{\mathcal{L}_i\}_i$ *and* $\{\mathcal{R}_i\}_i$ *are independent over the index i, for each* $0 \leq i \leq n$ $\mathcal{L}_i$ *and* $\mathcal{R}_i$ *are independent Poisson point-processes on* $(0,t) \times (0,t) \times \mathbf{T}$ *with the same intensity measure*

$$dt 1_{\{0<t<t_i\}} \frac{dh}{(1+h)^2} \frac{1+t}{t} 1_{\{0<h<t\}} \mathbf{\Lambda}^p_h \tag{13}$$

Proof. The independence of the sets $\mathcal{L}_i$ over the index i follows from the independence of the excursions $\varepsilon^{<t}_i$ of $X_{\mathcal{T}_{t,n}}$ below level t (same for the sets $\mathcal{R}_i$). The strong Markov property of $X_\mathcal{T}$ extends this claim to the independence of $\mathcal{R}_0$ and $\mathcal{L}_n$ from these sets as well. The conditional independence and the equality in law of $\mathcal{L}_i$ and $\mathcal{R}_i$ given t_i, follows from the time reversibility and the strong Markov property of $X_\mathcal{T}$. Consider the left half $\varepsilon^{<t}_{i,L}$ of an excursion below level t. By Lemma 2, the conditional law of $t - \varepsilon^{<t}_i$ given (l_i, t_i) is that of $X_\mathcal{T} | \{\sup(\mathcal{C}_\mathcal{T}) = t_i\}$, hence $t - \varepsilon^{<t}_{i,L}$ has the law of $X_\mathcal{T} | \{\tau_{t_i} < \tau_0\}$ where τ_{t_i}, τ_0 are the first hitting times of $(t_i, +1), (0, -1)$ respectively by $X_\mathcal{T}$. We then consider the levels $\{t_{i,L}(j)\}_j$ of constancy of $t - \varsigma_{i,L} = \sup(t - \varepsilon^{<t})$. The fact that $\mathcal{C}_\mathcal{T}$ is an alternating sum of exponential variables implies that $\{t_{i,L}(j)\}_j$ form a Poisson process of rate 1 on the set $(0, t_i)$. It also implies that the excursions $\{\varepsilon^{<t}_{i,L}(j)\}_j$ of $\varepsilon^{<t}_{i,L} - \varsigma_{i,L}$ above these levels have the laws of $X_\mathcal{T} | \{\sup(X_\mathcal{T}) < t_{i,L}(j)\}$. Hence given $t_{i,L}(j)$ the law of $h_{i,L}(j) = \sup(\varepsilon^{<t}_{i,L}(j))$ by (4) has the density $\frac{dh}{(1+h)^2} \frac{1+t_{i,L}(j)}{t_{i,L}(j)}$ on the set $(0, t_{i,L}(j))$. Then given $h_{i,L}(j)$ $\varepsilon^{<t}_{i,L}(j)$ has the law of $X_\mathcal{T} | \{\sup(X_\mathcal{T}) = h_{i,L}(j)\}$, and the tree whose contour process is $\varepsilon^{<t}_{i,L}(j)$ has the law of of $\mathcal{T}_{\Delta = h_{i,L}(j)}$. Now, the strong Markov property implies that the p-sampling on the local maxima of $\mathcal{C}_{\mathcal{T}_{t,n}}$ is for each $\varepsilon^{<t}_{i,L(j)}$ again a Bernoulli p-sampling on its local maxima. Thus the law of the p-sampled tree $\Upsilon_{i,L}(j)$ is $\mathbf{\Lambda}^p_{h_{i,L}(j)}$. Putting all the above results together we have that the set $\{(t_{i,L}(j), h_{i,L}(j), \Upsilon_{i,L}(j))\}_j$ is a Poisson point-process with intensity measure

$dt 1_{\{0<t<t_i\}} \frac{dh}{(1+h)^2} 1_{\{0<h<t\}} \mathbf{\Lambda}^p_h$. $\square$

In the context of the contour process the p-sampled individuals form a random set of marks along its time coordinate. The fact that $\mathcal{C}_{\mathcal{T}}$ is an alternating sum of independent Exponential(rate 1) random variables implies that the local maxima of $\mathcal{C}_{\mathcal{T}}$ form a Poisson process of rate $\frac{1}{2}$ along its time axis, and in fact the same is true for the local maxima of each part of an excursion of $\mathcal{C}_{\mathcal{T}_{t,n}}$ below t. The p-sampled local maxima thus form a Poisson process of rate $\frac{p}{2}$ along the time axis. The appropriate rescaling (as considered in the asymptotics of Section 2) speeds up the time axis of $\mathcal{C}_{\mathcal{T}_{t,n}}$ by n. Hence for $np_n \to p_0$ as $n \to \infty$ the p_n-sampling converges along the time axis to a Poisson process of rate $\frac{p_0}{2}$. With this in mind we turn to a conditioned Brownian excursion $\mathcal{B}_{t_0,1}$ with a Poisson(rate $\frac{p_0}{2}$) process of marks along its time axis.

Let us introduce a process derived from a conditioned Brownian excursion $\mathcal{B}_{t_0,1}$ in the same manner that $\mathcal{P}_p(\Xi_{t,n})$ was derived from the contour process of the conditioned branching process $\mathcal{C}_{\mathcal{T}_{t,n}}$. Define p_0-sampling on $\mathcal{B}_{t_0,1}$ to be a Poisson(rate $\frac{p_0}{2}$) process along the time axis of $\mathcal{B}_{t_0,1}$. We indicate this by putting a star mark on the graph of $\mathcal{B}_{t_0,1}$ corresponding to the times of this process. Let $\varepsilon^{<t_0}$ be the excursions of $\mathcal{B}_{t_0,1}$ below level t_0

$$\varepsilon^{<t_0}_{\ell,L}(u) = \mathcal{B}_{t_0,1}(i_{t_0}(\ell^-) + u),\ u \in [0, i_{t_0}(\ell) - i_{t_0}(\ell^-))$$

The points of their infima as defined earlier are $a_\ell = t_0 - t_\ell$, occurring at times $U_{a_\ell} = \{u \in (i_{t_0}(\ell^-), i_{t_0}(\ell)) : \mathcal{B}_{t_0,1}(u) = a_\ell\}$. For each $\varepsilon^{<t_0}_\ell$ we define its left part (left relative to the point of its infimum) to be

$$\varepsilon^{<t_0}_{\ell,L}(u) = \mathcal{B}_{t_0,1}(i_{t_0}(\ell^-) + u),\ \ \varsigma_{\ell,L}(u) = \inf_{0\le v\le u} \varepsilon^{<t_0}_{\ell,L}(v),\ \ u \in [0, U_{a_\ell} - i_{t_0}(\ell^-))$$

Figure 6 shows $\varepsilon^{<t_0}_{\ell,L}$ with its infimum process $\varsigma_{\ell,L}$ (all the following definitions that we make for $\varepsilon^{<t_0}_{\ell,L}$ apply equivalently to the part $\varepsilon^{<t_0}_{i,R}$ of $\varepsilon^{<t_0}$ on the right of its infimum up to symmetry about the vertical axis through a_ℓ).

Let $a_{\ell,L}(j), j \ge 0$ be the successive levels of constancy of $\varsigma_{\ell,L}$, and let $t_{\ell,L}(j) = t_0 - a_{\ell,L}(j)$. For each level $a_{\ell,L}(j)$ let $\varepsilon^{<t_0}_{\ell,L}(j)$ be the excursion of $\varepsilon^{<t_0}_{\ell,L} - \varsigma_{\ell,L}$ that lies above this level, and $h_{\ell,L}(j) = \sup(\varepsilon^{<t_0}_{\ell,L}(j))$ be its height (see Figure 6). Note that a.s. all the p_0-sampled points on $\mathcal{B}_{t_0,1}$ lie on these excursions $\varepsilon^{<t_0}_{\ell,L}(j)$. We define a tree induced by a p_0-sampled excursion $\varepsilon^{<t_0}_{\ell,L}(j)$ as the tree whose contour process is the linear interpolation of the sequence of the values of $\varepsilon^{<t_0}_{\ell,L}(j)$ at the p_0-sampling times, alternating with the sequence of the minima of $\varepsilon^{<t_0}_{\ell,L}(j)$ between the p_0-sampling times. Denote this tree by $\gamma_{\ell,L}(j)$.

Remark. This definition of a tree from an excursion path sampled at given times has been explored for different sampling distributions in the literature (see [Pi,02] §6. for examples). Since for each $\varepsilon^{<t_0}_\ell$ there are a.s. only finitely

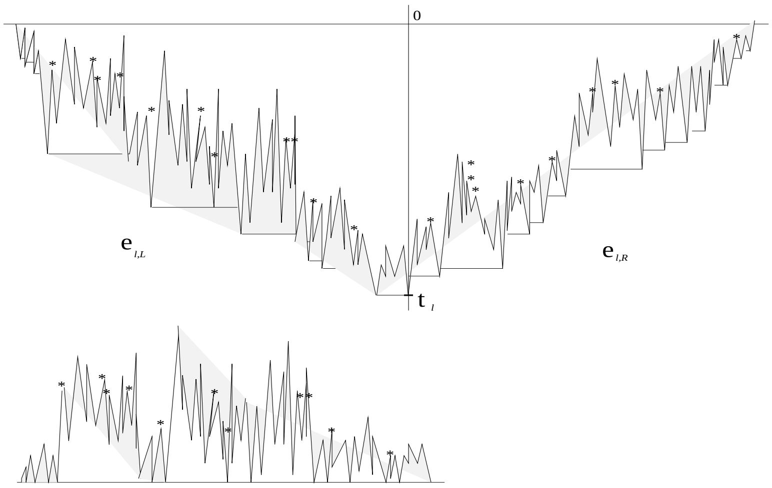

Figure 6: A half of an excursion of $\mathcal{B}_{t_0,1}$ below t_0, $\varepsilon^{<t_0}_{\ell,L}$, with its infimum process $\varsigma_{\ell,L}$ whose levels of constancy are $\{t_{\ell,L}(j)\}_j$, above which lie p-marked subtrees $\{\Upsilon_{\ell,L}(j)\}_j$ of heights $\{h_{\ell,L}(j)\}_j$. The corresponding process $\varepsilon^{<t_0}_{\ell,L} - \varsigma_{\ell,L}$ is shown bellow it.

many p_0-sampled points the trees $\{\gamma_{\ell,L}(j)\}_j, \{\gamma_{\ell,R}(j)\}_j$ are a.s. in the space $\mathbf{T}$ of rooted planar trees with edge-lengths and finitely many leaves.

Now to each point (ℓ, t_ℓ) of $\pi_{t_0,1}$ we attach a "left" set and an (analogously defined) "right" set

$$\mathcal{L}_\ell = \{(t_{\ell,L}(j), h_{\ell,L}(j), \gamma_{\ell,L}(j))\}_{j\geq 0},\ \mathcal{R}_\ell = \{(t_{\ell,R}(j), h_{\ell,R}(j), \gamma_{\ell,R}(j))\}_{j\geq 0} \quad (14)$$

We also define the first "right" set $\mathcal{R}_0$ and the last "left" set $\mathcal{L}_1$ from paths $\varepsilon^{<t_0}_{0,R}$ of $\mathcal{B}_{t_0,1}$ before the first hitting time of t_0, and $\varepsilon^{<t_0}_{1,L}$ of $\mathcal{B}_{t_0,1}$ after the last hitting time of t_0 (let $\mathcal{L}_0 = \mathcal{R}_1 = \emptyset$, $t_0 = t_1 = t_0$). Then

Definition. Let the *Continuum Historical point-process* with parameter p_0 be the random set

$$\mathcal{P}_{p_0}(\xi_{t_0,1}) = \{(\ell, t_\ell, \mathcal{L}_\ell, \mathcal{R}_\ell) : (\ell, t_\ell) \in \pi_{t_0,1},\ i_{t_0}(\ell^-) \neq i_{t_0}(\ell)\} \quad (15)$$

For $p_0 \in (0,1)$ let λ^{p_0} denote the law on $\mathbf{T}$ induced by the p_0-sampled unconditioned Brownian excursion $\mathcal{B}$ (the definition is the same as for $\gamma_{\ell,L}(j)$ from the p_0-sampled $\varepsilon^{<t_0}_{\ell,L}(j)$). Then for any $h > 0$, let $\lambda^{p_0}_h$ be the law induced by restricting λ^{p_0} to the set of Brownian excursions $\mathcal{B}$ of height h. The law of $\mathcal{P}_{p_0}(\xi_{t_0,1})$ is now described by the following Lemma.

Lemma 7. *The random set $\mathcal{P}_{p_0}(\xi_{t_0,1})$ is such that:*
- $\{(\ell, t_\ell) : i_{t_0}(\ell^-) \neq i_{t_0}(\ell)\}$ *is the Poisson point-process $\pi_{t_0,1}$ of Lemma 4,*

• *given* $\{(\ell, t_\ell) : i_{t_0}(\ell^-) \neq i_{t_0}(\ell)\}$ *the random sets* $\{\mathcal{L}_\ell\}_\ell$ *and* $\{\mathcal{R}_\ell\}_\ell$ *are independent over the index* ℓ, *for each* $\ell : i_{t_0}(\ell^-) \neq i_{t_0}(\ell)$ $\mathcal{L}_\ell$ *and* $\mathcal{R}_\ell$ *are independent Poisson point-processes on* $(0, t_0) \times (0, t_0) \times \mathbf{T}$ *with the same intensity measure*

$$dt 1_{\{0<t<t_\ell\}} \frac{dh}{h^2} 1_{\{0<h<t\}} \lambda_h^{p_0} \tag{16}$$

Proof. The independence of the sets $\mathcal{L}_\ell$ over the index ℓ (and the same for the sets $\mathcal{R}_\ell$) follows from the independence of the excursions of $\mathcal{B}_{t_0,1}$ below level t_0. This extends (using the strong Markov property of $\mathcal{B}$) to the sets $\mathcal{R}_0$ and $\mathcal{L}_\infty$ defined from the parts of the path of $\mathcal{B}_{t_0,1}$ of its ascent to level t_0 and its descent from it. For each $\varepsilon_\ell^{<t_0}$ as earlier defined we let $\epsilon_\ell^+ = t_0 - \varepsilon_\ell^{<t_0}$. By Lemma 4, the conditional law of ϵ_ℓ^+ given (ℓ, t_ℓ) is that of a Brownian excursion $\mathcal{B}$ conditioned on the value of its supremum $\mathcal{B}|\{\sup(\mathcal{B}) = t_\ell\}$. Let $\tau_{t_\ell} = \inf\{u > 0 : \epsilon_\ell^+(u) = t_\ell\}$, then by Williams' decomposition of a Brownian excursion $\mathcal{B}$ ([Ro-Wi,87] Vol.1 §III.49.), the law of $\epsilon_{\ell,L}^+ = t_0 - \varepsilon_{\ell,L}^{<t_0}$ is that of a 3-dimensional Bessel (Bess(3)) process ρ stopped at its first hitting time $\tau_{t_\ell}^\rho = \inf\{u > 0 : \rho(u) = t_\ell\}$ of t_ℓ. By time reversibility of $\mathcal{B}$ the process

$$r_{\ell,L}(u) = t_\ell - \epsilon_{\ell,L}^+(\tau_{t_\ell} - u), 0 \in (0, \tau_{t_\ell})$$

also has the law of the stopped Bess(3) process $\rho(u), u \in (0, \tau_{t_\ell}^\rho)$. Let

$$j_{\ell,L}(u) = \inf_{u \le v \le \tau_{t_\ell}} r_{\ell,L}, u \in (0, \tau_{t_\ell})$$

Then $\{t_\ell - t_{\ell,L}(j)\}_j$ are (in reversed index order) the successive levels of constancy of the process $j_{\ell,L}(u), u \in (0, \tau_{t_\ell})$, $\{h_{\ell,L}(j)\}_j$ (in reversed index order) are the heights of the successive excursions from 0 of the process $r_{\ell,L}(u) - j_{\ell,L}(u), u \in (0, \tau_{t_\ell})$, and $\{\gamma_{\ell,L}(j)\}_j$ (in reversed index order) are the trees induced by the p_0-sampled points on these excursions. To obtain the law of $j_{\ell,L}$ and $r_{\ell,L} - j_{\ell,L}$ consider the Bess(3) process $\rho(u), u \ge 0$ and its future infimum process $\jmath(u) = \inf_{v \ge u} \rho(v), u \ge 0$. We note that the law of $j_{\ell,L}(u), u \in (0, \tau_{t_\ell})$ is equivalent to that of $\jmath(u), u \in (0, \tau_{t_\ell}^\rho)$ if $\jmath(\tau_{t_\ell}^\rho) = t_\ell$, in other words, if $\rho(u), u \ge 0$ after it first reaches t_ℓ never returns to that height again. So,

$$(j_{\ell,L}, r_{\ell,L} - j_{\ell,L}) \stackrel{d}{=} (\jmath, \rho - \jmath)|\{\jmath(\tau_{t_\ell}^\rho) = t_\ell\} \text{ for } u \in (0, \tau_{t_\ell})$$

By Pitman's theorem, then by Levy's theorem ([Re-Yo,91] VI.§3.)

$$(\jmath, \rho - \jmath) \stackrel{d}{=} (\zeta, \zeta - \beta) \stackrel{d}{=} (\widetilde{\ell}, |\widetilde{\beta}|)$$

where β is a standard Brownian motion, ζ its supremum process; $|\widetilde{\beta}|$ is a reflected Brownian motion, $\widetilde{\ell}$ its local time at 0 (with the occupation time normalization). Thus, for $\widetilde{\tau}_{t_\ell} := \inf\{u \ge 0 : |\widetilde{\beta}|_u + \widetilde{\ell}_u = t_\ell\}$,

$$(j_{\ell,L}, r_{\ell,L} - j_{\ell,L}) \stackrel{d}{=} (\widetilde{\ell}, |\widetilde{\beta}|)\, |\{\widetilde{\ell}_{\widetilde{\tau}_{t_\ell}} = t_\ell\} \quad \text{for } u \in (0, \tau_{t_\ell})$$

The condition $\{\widetilde{\ell}_{\widetilde{\tau}_{t_\ell}} = t_\ell\}$ is equivalent to the condition $\{\widetilde{\ell}_{\widetilde{\tau}_{t_\ell}} = t_\ell, |\widetilde{\beta}|_{\widetilde{\tau}_{t_\ell}} = 0\}$ and $\{\widetilde{\ell}_u < t_\ell, |\widetilde{\beta}|_u < t_\ell - \widetilde{\ell}_u,\ \text{for } u \in (0, \widetilde{\tau}_{t_\ell})\}$. Hence,

$$
\begin{aligned}
&(j_{\ell,L}, r_{\ell,L} - j_{\ell,L}) \stackrel{d}{=} \\
&\quad (\widetilde{\ell}, |\widetilde{\beta}|)\,|\{\widetilde{\ell}_u < t_\ell, |\widetilde{\beta}|_u < t_\ell - \widetilde{\ell}_u : u \in (0, \widetilde{\tau}_{t_\ell}); \widetilde{\ell}_{\widetilde{\tau}_{t_\ell}} = t_\ell, |\widetilde{\beta}|_{\widetilde{\tau}_{t_\ell}} = 0\}
\end{aligned}
\tag{17}
$$

Since $\big(\widetilde{\ell}, \sup(|\widetilde{\beta}|)\big)$ is a Poisson point-process with intensity measure $\widetilde{d\ell} \times \frac{\widetilde{dh}}{\widetilde{dh^2}}$, then using the independence property of a Poisson random measure on disjoint sets in (17), we obtain for $t = t_\ell - \widetilde{\ell}$ that $\big(t_\ell - j_{\ell,L}, \sup(r_{\ell,L} - j_{\ell,L})\big)$ is a Poisson point-process with intensity measure

$$
dt 1_{(0<t<t_\ell)} \frac{dh}{h^2} 1_{(0<h<t)}
$$

Recall the relationship of the values $\{t_{\ell,L}(j), h_{\ell,L}(j), \gamma_{\ell,L}(j)\}_j$ of $\mathcal{L}_\ell$ with the processes $j_{\ell,L}$ and $r_{\ell,L} - j_{\ell,L}$. The above result thus implies that $\mathcal{L}_\ell$ is a Poisson point-process with intensity measure

$$
dt 1_{(0<t<t_\ell)} \frac{dh}{h^2} 1_{(0<h<t)}\, \lambda_h^{p_0}
$$

where the last factor comes from the fact that $\gamma_{\ell,L}(j)$ is just the tree induced by the p_0-sampled excursion of $|\widetilde{\beta}|$ of height $h_{\ell,L}(j)$. $\square$

We now consider more closely the trees $\Upsilon_{i,L}(j)$ and $\gamma_{\ell,L}(j)$ induced by the sampled excursions appearing in the historical point-processes above. In both cases we have an excursion type function $\mathcal{C}_\mathcal{T}$ and $\mathcal{B}$ of some given height with marks on it produced by a sampling process. The laws of trees induced by sampled excursions of unrestricted height are known in the literature (see [Ho,00] for the Brownian excursion case). For the trees from excursions of a given height that we need to consider here, we use a recursive description for both $\mathcal{C}_\mathcal{T}$ and $\mathcal{B}$ (see [Ab-Ma,92] for a similar recursive description of an infinite tree induced by an unsampled Brownian excursion). Consider a "spine" of the tree extending from the root of the tree to the point of maximal height in the excursion. An equivalent representation of the tree is one in which subtrees of the trees on the left and on the right of the axis through the spine are attached to this spine (see Figure 7). We obtain the levels at which these subtrees are attached as well as the description of the subtrees as follows.

Let us denote the excursion function defining this tree, either $\mathcal{C}_\mathcal{T}$ or $\mathcal{B}$, by $\varepsilon(u), u \geq 0$. Let $h = \sup(\varepsilon)$ be its given height, and $U_h = \{u \geq 0 : \varepsilon(u) = h\}$ the time at which it is achieved. Then let $\varepsilon_L(u), u \in [0, U_h]$ be the left part of the excursion, and let $\varsigma_L(u) = \inf_{0\leq v\leq u} \varepsilon(v), u \in [0, U_h]$ be its past infimum process (note that all the following definitions are to be equivalently made for

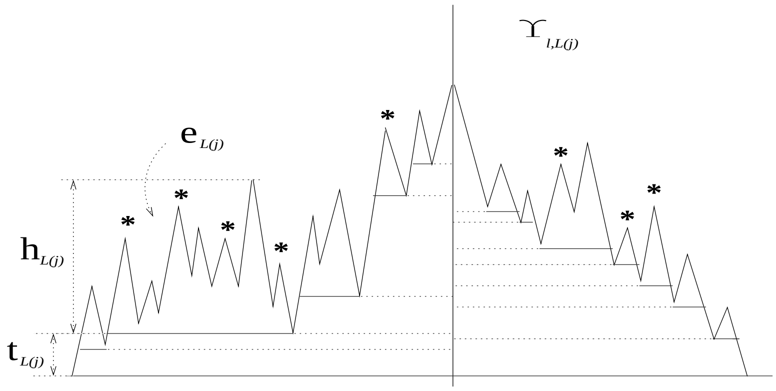

Figure 7: A recursive derivation of the tree induced by an excursion ε of height h. The "first" set of the tree description is shown: branch levels $\{t_L(j)\}_j$ at which subtrees induced by sampled excursions $\{\varepsilon_L(j)\}_j$ of $\varepsilon_L - \varsigma_L$ of heights $\{h_L(j)\}_j$

ε_R up to symmetry about the vertical axis through h). Then the subtrees attaching on the left of the spine are defined by those excursions $\varepsilon_L(j)$ of the process $\varepsilon_L - \varsigma_L$ that contain at least one sampled mark in it. They are the trees induced by the sampled excursions $\varepsilon_L(j)$ of heights $h_L(j)$. The levels at which they are attached to the spine are the levels of constancy $t_L(j)$ of ς_L at which the excursions of $\varepsilon_L - \varsigma_L$ occur. Then the set $\{(t_L(j), h_L(j))\}_j$ is the "first" set of points defining our tress. The "second" set is derived with the same procedure from the sampled excursions $\{\varepsilon_L(j)\}_j$, and so on recursively.

Remark. This recursive procedure is very similar to our definition of the left and right sets, $\mathcal{L}_i, \mathcal{R}_i$ for $t - \varepsilon^{<t}$ and $\mathcal{L}_\ell, \mathcal{R}_\ell$ for $t_0 - \varepsilon^{<t_0}$, as defined earlier. The difference lies in the fact that the subtrees here are defined from excursions above the levels of constancy of the infimum process for ε, whereas earlier they were defined from excursions below the levels of constancy of the supremum process for $t - \varepsilon^{<t}$ and $t_0 - \varepsilon^{<t_0}$. However, the time inversion and reflection invariance of the transition function of ε will allow us to easily derive the laws of the "first" set of points here from the results of Lemma 6 and Lemma 7.

Lemma 8. *The law $\mathbf{\Lambda}_h^{p_n}$ of a tree induced by a p_n-sampled contour process $\mathcal{C}_\mathcal{T}$ of a given height h is such that the first sets of points $\{t_L(j), h_L(j)\}_j$ and $\{t_R(j), h_R(j)\}_j$ are independent Poisson point-processes with intensity measure*

$$\frac{1}{\sqrt{p_n}}\, d\tau 1_{(0<\tau<h)} \, \frac{d\hbar}{(1+\hbar)^2} \frac{1+\tau}{\tau} 1_{(0<\hbar<h-\tau)} \tag{18}$$

The law $\lambda_h^{p_0}$ of tree induced by a p_0-sampled Brownian excursion $\mathcal{B}$ of a given height h is such that the first sets of points $\{t_L(j), h_L(j)\}_j$ and $\{t_R(j), h_R(j)\}_j$

are independent Poisson point-processes with intensity measure

$$\frac{1}{\sqrt{p_0}}\,d\tau 1_{(0<\tau<h)}\,\frac{d\hbar}{\hbar^2}1_{(0<\hbar<h-\tau)} \tag{19}$$

Let $n^{-1}\mathbf{\Lambda}_h^{p_n}$ be the law of the tree induced by a rescaled p_n-sampled contour process $\mathcal{C}_{\mathcal{T}}$ by n^{-1} in the vertical coordinate .Then for any $\{p_n \in (0,1)\}_{n\geq 1}$ such that $np_n \underset{n\to\infty}{\to} p$ we have $n^{-1}\mathbf{\Lambda}_h^{p_n} \underset{n\to\infty}{\Longrightarrow} \lambda_h^{p_0}$.

Proof. The key for this proof is to observe the following. If $\varepsilon(u), u \geq 0$ is the p_n-sampled process $X_{\mathcal{T}}|\{\sup(\mathcal{C}_{\mathcal{T}}) = h\}$ then $\varepsilon_L(u) = \varepsilon(u), u \in [0, U_h]$ has the law of a p_n-sampled $X_{\mathcal{T}}|\{\tau_h < \tau_0\}$ where τ_h, τ_0 are the first hitting times of $(h, +1), (0, -1)$ respectively by $X_{\mathcal{T}}$. Then time reversibility and the reflection invariance of the transition function of $X_{\mathcal{T}}$ imply that $h-\varepsilon_L(U_h-u), u \in [0, U_h]$ has the same law as $\varepsilon_L(u), u \in [0, U_h]$. Now the levels of constancy of ς_L, and the corresponding excursions $\varepsilon_L - \varsigma_L$ above them, are equivalent to the levels of constancy and excursions of a set $\mathcal{L}_i$ considered in Lemma 6, thus giving a Poisson process of intensity measure as in (13). The factor $\frac{1}{\sqrt{p}}$ in the intensity measure (18) comes from the fact that here we only consider the excursions of $\varepsilon_L - \varsigma_L$ that have at least one sampled mark in them. Namely, for the branching process $\mathcal{T}$, if N_{tot} denote the total population size of $\mathcal{T}$, then the generating function of N_{tot} is $\mathbf{E}(x^{N_{tot}}) = 1-\sqrt{1-x}$ ([Fe,68] xxx-a br.pr. book). Hence, the chance of at least one mark in the p_n-sampled point-process of $\mathcal{T}$ is $1 - \mathbb{E}((1-p_n)^{N_{tot}}) = \sqrt{p_n}$.

A similar argument applies when $\varepsilon(u), u \geq 0$ is the process $\mathcal{B}|\{\sup(\mathcal{B}) = h\}$ sampled at Poisson(rate $\frac{p_0}{2}$) times. Time reversibility and reflection invariance of the transition function of $\mathcal{B}$ allow us to identify that the law of the levels of constancy of of ς_L, and the corresponding excursions $\varepsilon_L - \varsigma_L$ above them are the same as those for a set $\mathcal{L}_\ell$ considered in Lemma 7, which we know form a Poisson process with intensity measure as in (16). The factor $\frac{1}{\sqrt{p_0}}$ in the intensity measure of (19) then comes from the rate of excursions with at least one sampled mark. Namely, a Poisson(rate $\frac{p_0}{2}$) process of marks on $\mathcal{B}$ along its time coordinate is in its local time coordinate a Poisson(rate $\sqrt{p_0}$) process of marks ([Ro-Wi,87] Vol.2§VI.50.).

Now the law of the first set of the rescaled process with under $n^{-1}\mathbf{\Lambda}_h^{p_n}$ converges to the law of the first set of the process with the law $\lambda_h^{p_0}$. This follows from the fact that the former is a sequence of Poisson point-processes whose support set and intensity measure converge to those of the latter Poisson point-process. Since for Poisson random measures the convergence of finite dimensional sets is sufficient to insure weak convergence of the whole process our claim follows for the first sets, and by recursion for the whole process. □

We apply the same rescaling as earlier for $\mathit{\Pi}_{t,n}$ now on $\mathcal{P}_p(\Xi_{t,n})$. Namely we rescale both coordinates for $\mathit{\Pi}_{t,n} \subset \mathcal{P}_p(\Xi_{t,n})$ by n^{-1}, so that the vertical coordinate of the sets $\mathcal{L}_i$,$\mathcal{R}_i$ is also rescaled by n^{-1}, and the sampling rate is rescaled by n. Denote this by

$$n^{-1}\mathcal{P}_p(\Xi_{t,n}) = \{(n^{-1}l_i, n^{-1}\tau_i, n^{-1}\mathcal{L}_i, n^{-1}\mathcal{R}_i) : (l_i, \tau_i, \mathcal{L}_i, \mathcal{R}_i) \in \mathcal{P}_p(\Xi_{t,n})\} \quad (20)$$

The asymptotic properties of the rescaled p-sampled historical process are now easily established from our earlier results.

Theorem 9. *For any $\{t_n > 0\}_{n\geq 1}$, and $\{p_n \in (0,1)\}_{n\geq 1}$ such that $\frac{t_n}{n} \underset{n\to\infty}{\to} t_0$, and $np_n \underset{n\to\infty}{\to} p$ we have $n^{-1}\mathcal{P}_{p_n}(\Xi_{t_n,n}) \underset{n\to\infty}{\Longrightarrow} \mathcal{P}_{p_0}(\xi_{t_0,1})$.*

Proof. By Theorem 5 we already have that $n^{-1}\mathit{\Pi}_{n,t_n} \underset{n\to\infty}{\Longrightarrow} \pi_{t_0,1}$. Applying the rescaling to the results of Lemma 6 together with the result of Lemma 8 now implies that the support set and intensity measure of the Poisson point-process of each $\mathcal{L}_i$ after rescaling converges to those of the Poisson point-process $\mathcal{L}_\ell$ as given by Lemma 7. Then the convergence of the support set and intensity measure for the Poisson random measure $\mathcal{P}_{p_n}(\Xi_{t_n,n})$ to those of $\mathcal{P}_{p_0}(\xi_{t_0,1})$ implies the weak convergence of these processes. $\square$

Acknowledgements. The work in this paper has been done under the guidance of D.Aldous. I would also like to thank J.Pitman and S.Evans for valuable comments and suggestions.

References

[Ab,92] Abraham,R. (1992) Un Arbre Aletoire Infini Associe a l'Excursion Brownienne. *Seminaire de Probabilites XXVI*, p.374-397, Springer, Lect. Notes in Math. 1526.

[Ab-Ma,92] Abraham,R. and Mazliak,L. (1998) Branching Properties of Brownian Paths and Trees. *Expositiones Mathematicae*, Vol. 16, p.59-74.

[Al,93] Aldous,D. (1993) The Continuum Random Tree III. *Annals of Probability*, Vol.21, p.248-289.

[Bi,99] Billingsley,P. (1999) *Convergence of Probability Measures.* 3rd ed., John Wiley & Sons.

[Du,78] Durrett,R. (1978) The Genealogy of Critical Branching Processes. *Stochastic Processes and their Applications.* Vol.8, p.101-116.

[Fe,68] Feller,W. (1968) *An Introduction to Probability Theory and its Applications.* Vol.1, 3rd ed., John Wiley & Sons.

[Ho,00] Hobson,D.G. (2000) Marked excursions and random trees. *Seminaire de Probabilites XXXIV*, p.289-301, Springer, 2000.

[Kai,76] Kaigh,W.D. (1976) An invariance principle for random walk conditioned by a late return to zero. *Annals of Probability*, Vol.4, p.115-121.

[LG,89] Le Gall,J.F. (1989) Marches aleatoires, mouvement brownien et processus de branchement. *Seminaire de Probabilites XXIII*, p.258-274, Springer, Lect. Notes in Math. 1372.

[LG,91] Le Gall,J.F. (1991) Brownian Excursions, Trees and Measure-valued Branching Processes. *Annals of Probability*, Vol 19, p.1399-1439.

[Ne-Pi,89a] Neveu,J. and Pitman,J.W. (1989) Renewal Property of the Extrema and the Tree Property of a One-dimensional Brownian Motion. *Seminaire de Probabilites XXIII*, p.239-247, Springer, Lect. Notes in Math. 1372.

[Ne-Pi,89b] Neveu,J. and Pitman,J.W. (1989) The Branching process in a Brownian excursion. *Seminaire de Probabilites XXIII*, p.248-257, Springer, Lect. Notes in Math. 1372.

[Pi,02] Pitman,J.W. (2002) Combinatorial Stochastic Processes. *Lecture Notes for Saint-Flour Course, July 2002*, Technical Report, UCBerkeley, SD-621.

[Re-Yo,91] Revuz,D. and Yor,M. (1991) *Continuous Martingales and Brownian Motion*. 1st ed., Springer-Verlag.

[Ro-Wi,87] Rogers,L.C.G. and Williams,D. (1987) *Diffusions, Markov processes, and Martingales*. Vol.1 &2, 2nd ed., John Wiley & Sons.