The Annals of Probability
2007, Vol. 35, No. 6, 2091–2121
DOI: 10.1214/009117907000000187

REGENERATIVE REAL TREES

BY MATHILDE WEILL

École Normale Supérieure

In this work, we give a description of all σ-finite measures on the space of rooted compact $\mathbb{R}$-trees which satisfy a certain regenerative property. We show that any infinite measure which satisfies the regenerative property is the "law" of a Lévy tree, that is, the "law" of a tree-valued random variable that describes the genealogy of a population evolving according to a continuous-state branching process. On the other hand, we prove that a probability measure with the regenerative property must be the law of the genealogical tree associated with a continuous-time discrete-state branching process.

1. Introduction. Galton–Watson trees are well known to be characterized among all random discrete trees by a regenerative property. More precisely, if γ is a probability measure on $\mathbb{Z}_+$, the law Π_γ of the Galton–Watson tree with offspring distribution γ is uniquely determined by the following two conditions: Under the probability measure Π_γ:

(i) the ancestor has p children with probability $\gamma(p)$,

(ii) if $\gamma(p) > 0$, then conditionally on the event that the ancestor has p children, the p subtrees which describe the genealogy of the descendants of these children, are independent and distributed according to the probability measure Π_γ.

The aim of this work is to study σ-finite measures satisfying an analogue of this property on the space of equivalence classes of rooted compact $\mathbb{R}$-trees. It would be interesting to study the case of locally compact $\mathbb{R}$-trees. However, we will only be concerned with compact $\mathbb{R}$-trees in this paper.

An $\mathbb{R}$-tree is a metric space $(\mathcal{T}, d)$ such that for any two points σ_1 and σ_2 in $\mathcal{T}$, there is a unique arc with endpoints σ_1 and σ_2 (which is denoted by $[\![\sigma_1, \sigma_2]\!]$), and furthermore this arc is isometric to a compact interval of the real line. In this work, all $\mathbb{R}$-trees are supposed to be compact. A rooted $\mathbb{R}$-tree is an $\mathbb{R}$-tree with a distinguished vertex called the root. Say that two rooted $\mathbb{R}$-trees are equivalent if there is a root-preserving isometry that maps one onto the other. It was noted in [7] that the set $\mathbb{T}$ of all equivalence classes of rooted compact $\mathbb{R}$-trees equipped with the pointed Gromov–Hausdorff distance d_{GH}, is a Polish space. Hence it is legitimate to consider random variables with values in $\mathbb{T}$, that is, random $\mathbb{R}$-trees. A particularly important example is the CRT, which was introduced by Aldous

Received November 2005; revised December 2006.
AMS 2000 subject classification. 60J80.
Key words and phrases. Galton–Watson trees, Lévy trees, branching processes.

[1, 2] with a different formalism. Striking applications of the concept of random $\mathbb{R}$-trees can be found in the recent papers [7] and [8].

Let $\mathcal{T}$ be an $\mathbb{R}$-tree. We write $\mathcal{H}(\mathcal{T})$ for the height of the $\mathbb{R}$-tree $\mathcal{T}$, that is the maximal distance from the root to a vertex of $\mathcal{T}$. For every $t \geq 0$, we denote by $\mathcal{T}_{\leq t}$ the set of all vertices of $\mathcal{T}$ which are at distance at most t from the root, and by $\mathcal{T}_{>t}$ the set of all vertices which are at distance greater than t from the root. To each connected component of $\mathcal{T}_{>t}$ there corresponds a "subtree" of $\mathcal{T}$ above level t (see Section 2.2.3 for a more precise definition). For every $h > 0$, we define $Z(t, t+h)(\mathcal{T})$ as the number of subtrees of $\mathcal{T}$ above level t with height greater than h.

Let Θ be a σ-finite measure on $\mathbb{T}$, such that $0 < \Theta(\mathcal{H}(\mathcal{T}) > t) < \infty$ for every $t > 0$ and $\Theta(\mathcal{H}(\mathcal{T}) = 0) = 0$. For every $t > 0$ we denote by Θ^t the probability measure $\Theta(\cdot \mid \mathcal{H}(\mathcal{T}) > t)$. We say that Θ satisfies the regenerative property (R) if the following holds:

(R) For every $t, h > 0$ and $p \in \mathbb{N}$, under the probability measure Θ^t and conditionally on the event $\{Z(t, t+h) = p\}$, the p subtrees of $\mathcal{T}$ above level t with height greater than h are independent and distributed according to the probability measure Θ^h.

This is a natural analogue of the regenerative property stated above for Galton–Watson trees. Beware that, unlike the discrete case, there is no natural order on the subtrees above a given level. So, the preceding property should be understood in the sense that the unordered collection of the p subtrees in consideration is distributed as the unordered collection of p independent copies of Θ^h.

Property (R) is known to be satisfied by a wide class of infinite measures on $\mathbb{T}$, namely the "laws" of Lévy trees. Lévy trees have been introduced by Duquesne and Le Gall in [6]. Their precise definition is recalled in Section 2.3, but let us immediately give an informal presentation.

Let Y be a critical or subcritical continuous-state branching process. The distribution of Y is characterized by its branching mechanism function ψ. Assume that Y becomes extinct almost surely, which is equivalent to the condition $\int_1^\infty \psi(u)^{-1}\,du < \infty$. The ψ-Lévy tree is a random variable taking values in $(\mathbb{T}, d_{GH})$, which describes the genealogy of a population evolving according to Y and starting with infinitesimally small mass. More precisely, the "law" of the ψ-Lévy tree is defined in [6] as a σ-finite measure Θ_ψ on the space $(\mathbb{T}, d_{GH})$, such that $0 < \Theta_\psi(\mathcal{H}(\mathcal{T}) > t) < \infty$ for every $t > 0$. As a consequence of Theorem 4.2 of [6], the measure Θ_ψ satisfies property (R). In the special case $\psi(u) = u^\alpha$, $1 < \alpha \leq 2$ corresponding to the so-called stable trees, this was used by Miermont [11, 12] to introduce and study certain fragmentation processes.

In the present work we describe all σ-finite measures on $\mathbb{T}$ that satisfy property (R). We show that the only infinite measures satisfying property (R) are the measures Θ_ψ associated with Lévy trees. On the other hand, if Θ is a finite measure satisfying property (R), we can obviously restrict our attention to the case

$\Theta(\mathbb{T}) = 1$ and we obtain that Θ must be the law of the genealogical tree associated with a continuous-time discrete-state branching process.

THEOREM 1.1. *Let Θ be an infinite measure on the space $(\mathbb{T}, d_{\mathrm{GH}})$ such that $\Theta(\mathcal{H}(\mathcal{T}) = 0) = 0$ and $0 < \Theta(\mathcal{H}(\mathcal{T}) > t) < +\infty$ for every $t > 0$. Assume that Θ satisfies property* (R). *Then, there exists a continuous-state branching process, whose branching mechanism is denoted by ψ, which becomes extinct almost surely, such that $\Theta = \Theta_\psi$.*

THEOREM 1.2. *Let Θ be a probability measure on the space $(\mathbb{T}, d_{\mathrm{GH}})$ such that $\Theta(\mathcal{H}(\mathcal{T}) = 0) = 0$ and $0 < \Theta(\mathcal{H}(\mathcal{T}) > t)$ for every $t > 0$. Assume that Θ satisfies property* (R). *Then there exists $a > 0$ and a critical or subcritical probability measure γ on $\mathbb{Z}_+ \setminus \{1\}$ such that Θ is the law of the genealogical tree for a discrete-space continuous-time branching process with offspring distribution γ, where branchings occur at rate a.*

In other words, Θ in Theorem 1.2 can be described in the following way: There exists a real random variable J such that under Θ:

(i) J is distributed according to the exponential distribution with parameter a and there exists $\sigma_J \in \mathcal{T}$ such that $\mathcal{T}_{\le J} = [\![\rho, \sigma_J]\!]$,
(ii) the number of subtrees above level J is distributed according to γ and is independent of J,
(iii) for every $p \ge 2$, conditionally on J and given the event that the number of subtrees above level J is equal to p, these p subtrees are independent and distributed according to Θ.

Theorem 1.1 is proved in Section 3, after some preliminary results have been established in Section 2. A key idea of the proof is to use the regenerative property (R) to embed discrete Galton–Watson trees in our random $\mathbb{R}$-trees (Lemma 3.3). A technical difficulty comes from the fact that $\mathbb{R}$-trees are not ordered whereas Galton–Watson trees are usually defined as random ordered discrete trees (cf. Section 2.2.4 below). To overcome this difficulty, we assign a random ordering to the discrete trees embedded in $\mathbb{R}$-trees. Another major ingredient of the proof of Theorem 1.1 is the construction of a "local time" L_t at every level t of a random $\mathbb{R}$-tree governed by Θ. The local time process is then shown to be a continuous-state branching process with branching mechanism ψ, which makes it possible to identify Θ with Θ_ψ. Theorem 1.2 is proved in Section 4. Several arguments are similar to the proof of Theorem 1.1, so that we have skipped some details.

2. Preliminaries. In this section, we recall some basic facts about branching processes, $\mathbb{R}$-trees and Lévy trees.

2.1. *Branching processes.*

2.1.1. *Continuous-state branching processes.* A (continuous-time) continuous-state branching process (in short a CSBP) is a Markov process $Y = (Y_t, t \geq 0)$ with values in the positive half-line $[0, +\infty)$, with a Feller semigroup $(Q_t, t \geq 0)$ satisfying the following branching property: For every $t \geq 0$ and $x, x' \geq 0$,

$$Q_t(x, \cdot) * Q_t(x', \cdot) = Q_t(x + x', \cdot).$$

Informally, this means that the union of two independent populations started respectively at x and x' will evolve like a single population started at $x + x'$.

We will consider only the critical or subcritical case, meaning that, for every $t \geq 0$ and $x \geq 0$,

$$\int_{[0,+\infty)} y Q_t(x, dy) \leq x.$$

Then, if we exclude the trivial case where $Q_t(x, \cdot) = \delta_0$ for every $t > 0$ and $x \geq 0$, the Laplace functional of the semigroup can be written in the following form: For every $\lambda \geq 0$,

$$\int_{[0,+\infty)} e^{-\lambda y} Q_t(x, dy) = \exp(-x u(t, \lambda)),$$

where the function $(u(t, \lambda), t \geq 0, \lambda \geq 0)$ is determined by the differential equation

$$\frac{\partial u(t, \lambda)}{\partial t} = -\psi(u(t, \lambda)), \qquad u(0, \lambda) = \lambda,$$

and $\psi : \mathbb{R}_+ \to \mathbb{R}_+$ is of the form

$$\psi(u) = \alpha u + \beta u^2 + \int_{(0,+\infty)} (e^{-ur} - 1 + ur)\pi(dr), \tag{1}$$

where π is a σ-finite measure on $(0, +\infty)$ such that $\int_{(0,+\infty)} (r \wedge r^2)\pi(dr) < \infty$ and $\alpha, \beta \geq 0$. The process Y is called the ψ-continuous-state branching process (in short the ψ-CSBP).

Continuous-state branching processes may also be obtained as weak limits of rescaled Galton–Watson processes. We recall that an offspring distribution is a probability measure on $\mathbb{Z}_+$. An offspring distribution μ is said to be critical if $\sum_{i \geq 0} i\mu(i) = 1$ and subcritical if $\sum_{i \geq 0} i\mu(i) < 1$. Let us state a result that can be derived from [9] and [10].

THEOREM 2.1. *Let $(\mu_n)_{n \geq 1}$ be a sequence of offspring distributions. For every $n \geq 1$, denote by X^n a Galton–Watson process with offspring distribution*

μ_n, started at $X_0^n = n$. Let $(m_n)_{n\geq 1}$ be a nondecreasing sequence of positive integers converging to infinity. We define a sequence of processes $(Y^n)_{n\geq 1}$ by setting, for every $t \geq 0$ and $n \geq 1$,

$$Y_t^n = n^{-1} X_{[m_n t]}^n.$$

Assume that, for every $t \geq 0$, the sequence $(Y_t^n)_{n\geq 1}$ converges in distribution to Y_t where $Y = (Y_t, t \geq 0)$ is an almost surely finite process such that $\mathbb{P}(Y_\delta > 0) > 0$ for some $\delta > 0$. Then, Y is a continuous-state branching process and the sequence of processes $(Y^n)_{n\geq 1}$ converges to Y in distribution in the Skorokhod space $\mathbb{D}(\mathbb{R}_+)$.

PROOF. It follows from the proof of Theorem 1 of [10] that Y is a CSBP. Then, thanks to Theorem 2 of [10] and Theorem 3.4 of [9], there exists a sequence of offspring distributions $(\nu_n)_{n\geq 1}$ and a nondecreasing sequence of positive integers $(c_n)_{n\geq 1}$ such that we can construct for every $n \geq 1$ a Galton–Watson process Z^n started at c_n and with offspring distribution ν_n satisfying

$$\left(c_n^{-1} Z_{[nt]}^n, t \geq 0\right) \xrightarrow[n\to\infty]{\text{(d)}} (Y_t, t \geq 0),$$

where the symbol $\xrightarrow{\text{(d)}}$ indicates convergence in distribution in $\mathbb{D}(\mathbb{R}_+)$.

Let $(m_{n_k})_{k\geq 1}$ be a strictly increasing subsequence of $(m_n)_{n\geq 1}$. For $n \geq 1$, we set $B^n = X^{n_k}$ and $b_n = n_k$ if $n = m_{n_k}$ for some $k \geq 1$, and we set $B^n = Z^n$ and $b_n = c_n$ if there is no $k \geq 1$ such that $n = m_{n_k}$. Then, for every $t \geq 0$, $(b_n^{-1} B_{[nt]}^n)_{n\geq 1}$ converges in distribution to Y_t. Applying Theorem 3.4 of [9], we obtain that

$$\left(b_n^{-1} B_{[nt]}^n, t \geq 0\right) \xrightarrow[n\to\infty]{\text{(d)}} (Y_t, t \geq 0).$$

In particular, we have,

$$(Y_t^{n_k}, t \geq 0) \xrightarrow[k\to\infty]{\text{(d)}} (Y_t, t \geq 0). \tag{2}$$

As (2) holds for every strictly increasing subsequence of $(m_n)_{n\geq 1}$, we obtain the desired result. □

2.1.2. *Discrete-state branching processes.* A (continuous-time) discrete-state branching process (in short DSBP) is a continuous-time Markov chain $Y = (Y_t, t \geq 0)$ with values in $\mathbb{Z}_+$ whose transition probabilities $(P_t(i, j), t \geq 0)_{i\geq 0, j\geq 0}$ satisfy the following branching property: For every $i \in \mathbb{Z}_+$, $t \geq 0$ and $|s| \leq 1$,

$$\sum_{j=0}^{\infty} P_t(i, j) s^j = \left(\sum_{j=0}^{\infty} P_t(1, j) s^j\right)^i.$$

We exclude the trivial case where $P_t(i,i)=1$ for every $t \geq 0$ and $i \in \mathbb{Z}_+$. Then, there exist $a>0$ and an offspring distribution γ with $\gamma(1)=0$ such that the generator of Y can be written of the form

$$Q=\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \dots \\ a\gamma(0) & -a & a\gamma(2) & a\gamma(3) & a\gamma(4) & \dots \\ 0 & 2a\gamma(0) & -2a & 2a\gamma(2) & a\gamma(3) & \dots \\ 0 & 0 & 3a\gamma(0) & -3a & 3a\gamma(2) & \dots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}.$$

Furthermore, it is well known that Y becomes extinct almost surely if and only if γ is critical or subcritical. We refer the reader to [3] and [13] for more details.

2.2. *Deterministic trees.*

2.2.1. *The space* $(\mathbb{T}, \mathbb{d}_{\mathrm{GH}})$ *of rooted compact* $\mathbb{R}$*-trees.* We start with a basic definition.

DEFINITION 2.1. A metric space $(\mathcal{T}, d)$ is an $\mathbb{R}$-tree if the following two properties hold for every $\sigma_1, \sigma_2 \in \mathcal{T}$:

(i) There is an isometric map f_{σ_1,σ_2} from $[0, d(\sigma_1,\sigma_2)]$ into $\mathcal{T}$ such that $f_{\sigma_1,\sigma_2}(0)=\sigma_1$ and $f_{\sigma_1,\sigma_2}(d(\sigma_1,\sigma_2))=\sigma_2$.

(ii) If q is a continuous injective map from $[0,1]$ into $\mathcal{T}$ such that $q(0)=\sigma_1$ and $q(1)=\sigma_2$, we have

$$q([0,1]) = f_{\sigma_1,\sigma_2}([0, d(\sigma_1,\sigma_2)]).$$

A rooted $\mathbb{R}$-tree is an $\mathbb{R}$-tree with a distinguished vertex $\rho=\rho(\mathcal{T})$ called the root.

In what follows, $\mathbb{R}$-trees will always be rooted.

Let $(\mathcal{T}, d)$ be an $\mathbb{R}$-tree with root ρ, and $\sigma, \sigma_1, \sigma_2 \in \mathcal{T}$. We write $[\![\sigma_1, \sigma_2]\!]$ for the range of the map f_{σ_1,σ_2}. In particular, $[\![\rho, \sigma]\!]$ is the path going from the root to σ and can be interpreted as the ancestral line of σ.

The height $\mathcal{H}(\mathcal{T})$ of the $\mathbb{R}$-tree $\mathcal{T}$ is defined by $\mathcal{H}(\mathcal{T}) = \sup\{d(\rho,\sigma) : \sigma \in \mathcal{T}\}$. In particular, if $\mathcal{T}$ is compact, its height $\mathcal{H}(\mathcal{T})$ is finite.

Two rooted $\mathbb{R}$-trees $\mathcal{T}$ and $\mathcal{T}'$ are called equivalent if there is a root-preserving isometry that maps $\mathcal{T}$ onto $\mathcal{T}'$. We denote by $\mathbb{T}$ the set of all equivalence classes of rooted compact $\mathbb{R}$-trees. We often abuse notation and identify a rooted compact $\mathbb{R}$-tree with its equivalence class.

The set $\mathbb{T}$ can be equipped with the pointed Gromov–Hausdorff distance, which is defined as follows. If (E, δ) is a metric space, we use the notation δ_{Haus} for the usual Hausdorff metric between compact subsets of E. Then, if $\mathcal{T}$ and $\mathcal{T}'$ are two rooted compact $\mathbb{R}$-trees with respective roots ρ and ρ', we define the distance $\mathbb{d}_{\mathrm{GH}}(\mathcal{T}, \mathcal{T}')$ as

$$\mathbb{d}_{\mathrm{GH}}(\mathcal{T}, \mathcal{T}') = \inf\{\delta_{\mathrm{Haus}}(\phi(\mathcal{T}), \phi'(\mathcal{T}')) \vee \delta(\phi(\rho), \phi'(\rho'))\},$$

where the infimum is over all isometric embeddings $\phi : \mathcal{T} \to E$ and $\phi' : \mathcal{T}' \to E$ into a common metric space (E, δ). We see that $\mathbb{d}_{\mathrm{GH}}(\mathcal{T}, \mathcal{T}')$ only depends on the equivalence classes of $\mathcal{T}$ and $\mathcal{T}'$. According to Theorem 2 in [7], $\mathbb{d}_{\mathrm{GH}}$ defines a metric on $\mathbb{T}$ that makes it complete and separable. Furthermore, the distance $\mathbb{d}_{\mathrm{GH}}$ can often be evaluated in the following way. First recall that if $(\mathcal{T}, d)$ and $(\mathcal{T}', d')$ are two rooted compact $\mathbb{R}$-trees, a correspondence between $\mathcal{T}$ and $\mathcal{T}'$ is a subset $\mathcal{R}$ of $\mathcal{T} \times \mathcal{T}'$ such that for every $\sigma \in \mathcal{T}$ (resp. $\sigma' \in \mathcal{T}'$), there exists $\sigma' \in \mathcal{T}'$ (resp. $\sigma \in \mathcal{T}$) such that $(\sigma, \sigma') \in \mathcal{R}$. The distorsion of the correspondence $\mathcal{R}$ is then defined by

$$\mathrm{dis}(\mathcal{R}) = \sup\{|d(\sigma_1, \sigma_2) - d'(\sigma_1', \sigma_2')| : (\sigma_1, \sigma_1'), (\sigma_2, \sigma_2') \in \mathcal{R}\}.$$

Then if ρ and ρ' denote the respective roots of $\mathcal{T}$ and $\mathcal{T}'$, Lemma 2.3 in [7] ensures that

$$\mathbb{d}_{\mathrm{GH}}(\mathcal{T}, \mathcal{T}') = \tfrac{1}{2} \inf\{\mathrm{dis}(\mathcal{R}) : \mathcal{R} \in \mathcal{C}(\mathcal{T}, \mathcal{T}'), (\rho, \rho') \in \mathcal{R}\}, \tag{3}$$

where $\mathcal{C}(\mathcal{T}, \mathcal{T}')$ denotes the set of all correspondences between $\mathcal{T}$ and $\mathcal{T}'$.

We equip $\mathbb{T}$ with its Borel σ-field. If $\mathcal{T} \in \mathbb{T}$, we set $\mathcal{T}_{\le t} = \{\sigma \in \mathcal{T} : d(\rho, \sigma) \le t\}$ for every $t \ge 0$. Plainly, $\mathcal{T}_{\le t}$ is an $\mathbb{R}$-tree which is naturally rooted at ρ. Note that the mapping $\mathcal{T} \mapsto \mathcal{T}_{\le t}$ from $\mathbb{T}$ into $\mathbb{T}$ is Lipschitz for the Gromov–Hausdorff metric.

2.2.2. *The $\mathbb{R}$-tree coded by a function.* We now recall a construction of rooted compact $\mathbb{R}$-trees which is described in [6]. Let $g : [0, +\infty) \to [0, +\infty)$ be a continuous function with compact support satisfying $g(0) = 0$. We exclude the trivial case where g is identically zero. For every $s, t \ge 0$, we set

$$m_g(s, t) = \inf_{r \in [s \wedge t, s \vee t]} g(r),$$
$$d_g(s, t) = g(s) + g(t) - 2m_g(s, t).$$

We define an equivalence relation $\sim$ on $[0, +\infty)$ by declaring that $s \sim t$ if and only if $d_g(s, t) = 0$ [or equivalently if and only if $g(s) = g(t) = m_g(s, t)$]. Let $\mathcal{T}_g$ be the quotient space

$$\mathcal{T}_g = [0, +\infty)/\sim.$$

Then, d_g induces a metric on $\mathcal{T}_g$ and we keep the notation d_g for this metric. According to Theorem 2.1 of [6], the metric space $(\mathcal{T}_g, d_g)$ is a compact $\mathbb{R}$-tree. By convention, its root is the equivalence class of 0 for $\sim$ and is denoted by ρ_g.

2.2.3. *Subtrees of a tree above a fixed level.* Let $(\mathcal{T}, d) \in \mathbb{T}$ and $t \ge 0$. Denote by $\mathcal{T}^{i,\circ}, i \in I$ the connected components of the open set $\mathcal{T}_{>t} = \{\sigma \in \mathcal{T} : d(\rho, \sigma) > t\}$. Let $i \in I$. Then the ancestor of $\sigma \in \mathcal{T}^{i,\circ}$ at level t, that is, the unique vertex on the line segment $[\![\rho, \sigma]\!]$ at distance t from ρ, must be the same for all $\sigma \in \mathcal{T}^{i,\circ}$.

We denote by σ_i this common ancestor and set $\mathcal{T}^i = \mathcal{T}^{i,\circ} \cup \{\sigma_i\}$. Then $\mathcal{T}^i$ is a compact rooted $\mathbb{R}$-tree which is naturally rooted at σ_i. The trees $\mathcal{T}^i, i \in I$ are called the subtrees of $\mathcal{T}$ above level t. We now consider, for every $h > 0$,

$$Z(t, t+h)(\mathcal{T}) = \#\{i \in I : \mathcal{H}(\mathcal{T}^i) > h\}.$$

By a compactness argument, we can easily verify that $Z(t, t+h)(\mathcal{T}) < \infty$.

2.2.4. *Discrete trees.* We start with some formalism for discrete trees. We first introduce the set U defined by

$$U = \bigcup_{n \geq 0} \mathbb{N}^n,$$

where by convention $\mathbb{N}^0 = \{\varnothing\}$. An element of U is a sequence $u = u^1 \cdots u^n$, and we set $|u| = n$ so that $|u|$ represents the generation of u. In particular, $|\varnothing| = 0$. If $u = u^1 \cdots u^n$ and $v = v^1 \cdots v^m$ belong to U, we write $uv = u^1 \cdots u^n v^1 \cdots v^m$ for the concatenation of u and v. In particular, $\varnothing u = u \varnothing = u$. The mapping $\pi : U \setminus \{\varnothing\} \to U$ is defined by $\pi(u^1 \cdots u^n) = u^1 \cdots u^{n-1}$ [$\pi(u)$ is the father of u]. Let π^k be the kth iterative of the mapping π. Note that $\pi^k(u) = \varnothing$ if $k = |u|$.

A rooted ordered tree θ is a finite subset of U such that:

(i) $\varnothing \in \theta$,
(ii) $u \in \theta \setminus \{\varnothing\} \Rightarrow \pi(u) \in \theta$,
(iii) for every $u \in \theta$, there exists a number $k_u(\theta) \geq 0$ such that $uj \in \theta$ if and only if $1 \leq j \leq k_u(\theta)$.

We denote by $\mathcal{A}$ the set of all rooted ordered trees. If $\theta \in \mathcal{A}$, we write $\mathcal{H}(\theta)$ for the height of θ, that is $\mathcal{H}(\theta) = \max\{|u| : u \in \theta\}$. And for every $u \in \theta$, we define $\tau_u \theta \in \mathcal{A}$ by $\tau_u \theta = \{v \in U : uv \in \theta\}$. This is the tree θ shifted at u.

Let us define an equivalence relation on $\mathcal{A}$ by setting $\theta \sim \theta'$ if and only if we can find a permutation φ_u of the set $\{1, \ldots, k_u(\theta)\}$ for every $u \in \theta$ such that $k_u(\theta) \geq 1$, in such a way that

$$\theta' = \{\varnothing\} \cup \{\varphi_\varnothing(u^1)\varphi_{u^1}(u^2) \cdots \varphi_{u^1 \ldots u^{n-1}}(u^n) : u^1 \cdots u^n \in \theta, n \geq 1\}.$$

In other words $\theta \sim \theta'$ if they correspond to the same unordered tree. Let $\mathbb{A} = \mathcal{A}/\sim$ be the associated quotient space and let $\mathbb{p} : \mathcal{A} \to \mathbb{A}$ be the canonical projection. It is immediate that if $\theta \sim \theta'$, then $k_\varnothing(\theta) = k_\varnothing(\theta')$. So, for every $\xi \in \mathbb{A}$, we may define $k_\varnothing(\xi) = k_\varnothing(\theta)$ where θ is any representative of ξ. Let us fix $\xi \in \mathbb{A}$ such that $k_\varnothing(\xi) = k > 0$ and choose a representative θ of ξ. We can define $\{\xi^1, \ldots, \xi^k\} = \{\mathbb{p}(\tau_1\theta), \ldots, \mathbb{p}(\tau_k\theta)\}$ as the unordered family of subtrees of ξ above the first generation. Then, if $F : \mathcal{A}^k \to \mathbb{R}_+$ is any symmetric function, we have

$$\begin{aligned}(4)\quad &(\#\mathbb{p}^{-1}(\xi))^{-1} \sum_{\theta \in \mathbb{p}^{-1}(\xi)} F(\tau_1\theta, \ldots, \tau_k\theta) \\ &\quad = (\#\mathbb{p}^{-1}(\xi^1))^{-1} \cdots (\#\mathbb{p}^{-1}(\xi^k))^{-1} \sum_{\theta_1 \in \mathbb{p}^{-1}(\xi^1)} \cdots \sum_{\theta_k \in \mathbb{p}^{-1}(\xi^k)} F(\theta_1, \ldots, \theta_k).\end{aligned}$$

Note that the right-hand side of (4) is well defined since it is symmetric in $\{\xi^1, \dots, \xi^k\}$. The identity (4) is a simple combinatorial fact, whose proof is left to the reader.

A marked tree is a pair $T = (\theta, \{h_u\}_{u\in\theta})$ where $\theta \in \mathcal{A}$ and $h_u \geq 0$ for every $u \in \theta$. We denote by $\mathcal{M}$ the set of all marked trees. We can associate with every marked tree $T = (\theta, \{h_u\}_{u\in\theta}) \in \mathcal{M}$, an $\mathbb{R}$-tree $\mathcal{T}^T$ in the following way. Let $\mathbb{R}^\theta$ be the vector space of all mappings from θ into $\mathbb{R}$. Write $(e_u, u \in \theta)$ for the canonical basis of $\mathbb{R}^\theta$. We define $l_\varnothing = 0$ and $l_u = \sum_{k=1}^{|u|} h_{\pi^k(u)} e_{\pi^k(u)}$ for $u \in \theta$. Let us set

$$\mathcal{T}^T = \bigcup_{u\in\theta} [l_u, l_u + h_u e_u].$$

The tree $\mathcal{T}^T$ is a connected union of line segments in $\mathbb{R}^\theta$. It is equipped with the distance d_T which is the restriction to $\mathcal{T}^T$ of the $l_1(\theta)$-distance on $\mathbb{R}^\theta$, and can be rooted at $\rho(\mathcal{T}^T) = 0$ so that it becomes a rooted compact $\mathbb{R}$-tree. Consider for example the marked tree $T = (\theta, \{h_u\}_{u\in\theta})$ where $\theta = \{\varnothing, 1, 2, 21\}$, $h_\varnothing = h_{21} = 1$, $h_1 = 2$ and $h_2 = 3$. Then we have

$$\mathcal{T}^T = [0, e_\varnothing] \cup [e_\varnothing, e_\varnothing + 2e_1] \cup [e_\varnothing, e_\varnothing + 3e_2] \cup [e_\varnothing + 3e_2, e_\varnothing + 3e_2 + e_{22}].$$

If $\theta \in \mathcal{A}$, we write $\mathcal{T}^\theta$ for the $\mathbb{R}$-tree $\mathcal{T}^T$ where $T = (\theta, \{h_u\}_{u\in\theta})$ with $h_\varnothing = 0$ and $h_u = 1$ for every $u \in \theta \setminus \{\varnothing\}$, and we write d_θ for the associated distance. Notice in particular that $\mathcal{H}(\theta) = \mathcal{H}(\mathcal{T}^\theta)$. We then set $m_\varnothing = 0$ and $m_u = \sum_{k=0}^{|u|-1} e_{\pi^k(u)} = l_u + e_u$ for every $u \in \theta \setminus \{\varnothing\}$.

It is easily checked that $\mathcal{T}^\theta = \mathcal{T}^{\theta'}$ if $\theta \sim \theta'$. Thus for every $\xi \in \mathbb{A}$, we may write $\mathcal{T}^\xi$ for the tree $\mathcal{T}^\theta$ where θ is any representative of ξ.

2.2.5. *A discrete approximation of an* $\mathbb{R}$-*tree.* We will now explain how to approximate a general tree $\mathcal{T}$ in $\mathbb{T}$ by a discrete type tree. Let $\varepsilon > 0$ and set $\mathbb{T}^{(\varepsilon)} = \{\mathcal{T} \in \mathbb{T} : \mathcal{H}(\mathcal{T}) > \varepsilon\}$. We associate with every $\mathcal{T}$ in $\mathbb{T}^{(\varepsilon)}$ an element $\xi^\varepsilon(\mathcal{T})$ of $\mathbb{A}$ consisting of the root and of all points in $\mathcal{T}$ which are within distance $k\varepsilon$, $k \in \mathbb{N}$, to the root and which have a subtree above of height greater that ε (see Figure 1 below). More precisely we can construct $\xi^\varepsilon(\mathcal{T})$ by induction in the following way:

- If $\mathcal{T} \in \mathbb{T}^{(\varepsilon)}$ satisfies $\mathcal{H}(\mathcal{T}) \leq 2\varepsilon$, we set $\xi^\varepsilon(\mathcal{T}) = \mathbb{p}(\{\varnothing\})$.
- Let n be a positive integer. Assume that we have defined $\xi^\varepsilon(\mathcal{T})$ for every $\mathcal{T} \in \mathbb{T}^{(\varepsilon)}$ such that $\mathcal{H}(\mathcal{T}) \leq (n+1)\varepsilon$. Let $\mathcal{T}$ be an $\mathbb{R}$-tree such that $(n+1)\varepsilon < \mathcal{H}(\mathcal{T}) \leq (n+2)\varepsilon$. We set $k = Z(\varepsilon, 2\varepsilon)(\mathcal{T})$ and we denote by $\mathcal{T}^1, \dots, \mathcal{T}^k$ the k subtrees of $\mathcal{T}$ above level ε with height greater than ε. Then $\varepsilon < \mathcal{H}(\mathcal{T}^i) \leq (n+1)\varepsilon$ for every $i \in \{1, \dots, k\}$, so we can define $\xi^\varepsilon(\mathcal{T}^i)$. Let us choose a representative θ^i of $\xi^\varepsilon(\mathcal{T}^i)$ for every $i \in \{1, \dots, k\}$. We set

  $$\xi^\varepsilon(\mathcal{T}) = \mathbb{p}(\{\varnothing\} \cup 1\theta^1 \cup \cdots \cup k\theta^k),$$

  where $i\theta^i = \{iu : u \in \theta^i\}$. Clearly this does not depend on the choice of the representatives θ_i.

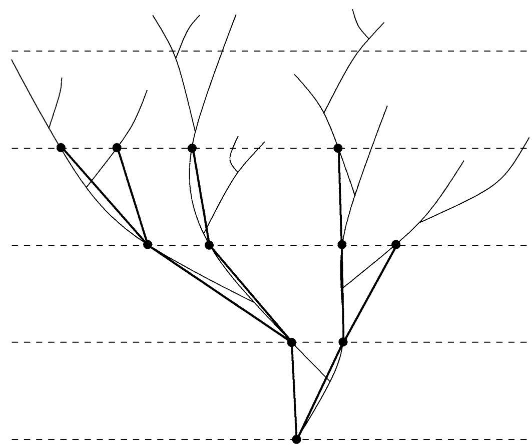

FIG. 1. *Construction of* $\xi^\varepsilon(\mathcal{T})$.

If $r > 0$ and $\mathcal{T}$ is a compact rooted $\mathbb{R}$-tree with metric d, we write $r\mathcal{T}$ for the same tree equipped with the metric rd.

LEMMA 2.2. *For every* $\varepsilon > 0$ *and every* $\mathcal{T} \in \mathbb{T}^{(\varepsilon)}$, *we have*

$$\mathbb{d}_{\mathrm{GH}}\big(\varepsilon\mathcal{T}^{\xi^\varepsilon(\mathcal{T})}, \mathcal{T}\big) \le 2\varepsilon. \tag{5}$$

PROOF. Let $\varepsilon > 0$ and $\mathcal{T} \in \mathbb{T}$. Let θ be any representative of $\xi^\varepsilon(\mathcal{T})$. Recall the notation $(m_u, u \in \theta)$. We can construct a mapping $\phi : \theta \to \mathcal{T}$ such that:

(i) for every $\sigma \in \mathcal{T}$, there exists $u \in \theta$ such that $0 \le d(\rho, \sigma) - d(\rho, \phi(u)) \le 2\varepsilon$ where ρ denotes the root of $\mathcal{T}$,

(ii) for every $u \in \theta$, $d(\rho, \phi(u)) = \varepsilon|u|$,

(iii) for every $u, u' \in \theta$, $0 \le \varepsilon d_\theta(m_u, m_{u'}) - d(\phi(u), \phi(u')) \le 2\varepsilon$.

To be specific, we always take $\phi(\varnothing) = \rho$, which suffices for the construction if $\mathcal{H}(\mathcal{T}) \le 2\varepsilon$. If $(n+1)\varepsilon < \mathcal{H}(\mathcal{T}) \le (n+2)\varepsilon$ for some $n \ge 1$, we have as above

$$\theta = \{\varnothing\} \cup 1\theta^1 \cup \cdots \cup k\theta^k,$$

where $\theta^1, \dots, \theta^k$ are representatives of respectively $\xi^\varepsilon(\mathcal{T}^1), \dots, \xi^\varepsilon(\mathcal{T}^k)$, if $\mathcal{T}^1, \dots, \mathcal{T}^k$ are the subtrees of $\mathcal{T}$ above level ε with height greater than ε. With an obvious notation we define $\phi(ju) = \phi_j(u)$ for every $j \in \{1, \dots, k\}$ and $u \in \theta^j$. Properties (i)–(iii) are then easily checked by induction.

Thus we can construct a correspondence $\mathcal{R} \in \mathcal{C}(\mathcal{T}, \mathcal{T}^{\xi^\varepsilon(\mathcal{T})})$ such that $(\rho, m_\varnothing) \in \mathcal{R}$ and which satisfies the following two properties:

- For every $\sigma \in \mathcal{T}$, there exists $u \in \theta$ such that $(\sigma, m_u) \in \mathcal{R}$ and $0 \le d(\rho, \sigma) - d(\rho, \phi(u)) \le 2\varepsilon$,

- for every $\sigma \in \mathcal{T}^{\xi^\varepsilon(\mathcal{T})}$, there exists $u \in \theta$ such that $(\phi(u), \sigma) \in \mathcal{R}$ and $\varepsilon d_\theta(\sigma, m_u) \leq \varepsilon$.

We easily check that $\mathrm{dis}(\mathcal{R}) \leq 4\varepsilon$. The result then follows from (3). □

2.3. *Lévy trees.* Roughly speaking, a Lévy tree is a $\mathbb{T}$-valued random variable which is associated with a CSBP in such a way that it describes the genealogy of a population evolving according to this CSBP.

2.3.1. *The measure* Θ_ψ. We consider on a probability space $(\mathbf{\Omega}, \mathbf{P})$ a ψ-CSBP $Y = (Y_t, t \geq 0)$, where the function ψ is of the form (1), and we suppose that Y becomes extinct almost surely. This condition is equivalent to

$$\int_1^\infty \frac{du}{\psi(u)} < \infty. \tag{6}$$

This implies that at least one of the following two conditions holds:

$$\beta > 0 \ \text{ or } \ \int_{(0,1)} r\pi(dr) = \infty. \tag{7}$$

The Lévy tree associated to Y will be defined as the tree coded by the so-called height process, which is a functional of the Lévy process with Laplace exponent ψ. Let us denote by $X = (X_t, t \geq 0)$ a Lévy process on $(\mathbf{\Omega}, \mathbf{P})$ with Laplace exponent ψ. This means that X is a Lévy process with no negative jumps, and that for every $\lambda, t \geq 0$,

$$\mathbf{E}(\exp(-\lambda X_t)) = \exp(t\psi(\lambda)).$$

Then, X does not drift to $+\infty$ and has paths of infinite variation [by (7)].

We can define the height process $H = (H_t, t \geq 0)$ by the following approximation:

$$H_t = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_0^t \mathbb{1}_{\{X_s \leq I_t^s + \varepsilon\}} \, ds,$$

where $I_t^s = \inf\{X_r : s \leq r \leq t\}$ and the convergence holds in probability (see Chapter 1 in [5]). Informally, we can say that H measures the size of the set $\{s \in [0, t] : X_{s-} \leq I_t^s\}$. Thanks to condition (6), we know that the process H has a continuous modification (see Theorem 1.4.3 in [5]). From now on, we consider only this modification.

Let us now set $I_t = \inf\{X_s : 0 \leq s \leq t\}$ for every $t \geq 0$, and consider the process $X - I = (X_t - I_t, t \geq 0)$. We recall that $X - I$ is a strong Markov process, for which the point 0 is regular. The process $-I$ is a local time for $X - I$ at level 0. We write N for the associated excursion measure. We let $\Delta(de)$ be the "law" of $(H_s, s \geq 0)$ under N. This makes sense because the values of the height process in an excursion of $X - I$ away from 0 only depend on that excursion (see Section

1.2 in [5]). Then, $\Delta(de)$ is a σ-finite measure on $C([0,\infty))$, and is supported on functions with compact support such that $e(0) = 0$.

The Lévy tree is the tree $(\mathcal{T}_e, d_e)$ coded by the function e, in the sense of Section 2.2.2, under the measure $\Delta(de)$. We denote by Θ_ψ the σ-finite measure on $\mathbb{T}$ which is the "law" of the Lévy tree, that is the image of $\Delta(de)$ under the measurable mapping $e \mapsto \mathcal{T}_e$.

2.3.2. *A discrete approximation of the Lévy tree.* Let us now recall that the Lévy tree is the limit in the Gromov–Hausdorff distance of suitably rescaled Galton–Watson trees.

We start by recalling the definition of Galton–Watson trees which was given informally in the introduction above. Let γ be a critical or subcritical offspring distribution. We exclude the trivial case where $\gamma(1) = 1$. Then, there exists a unique probability measure Π_γ on $\mathcal{A}$ such that:

(i) For every $p \ge 0$, $\Pi_\gamma(k_\varnothing = p) = \gamma(p)$,

(ii) for every $p \ge 1$ with $\gamma(p) > 0$, under $\Pi_\gamma(\cdot \mid k_\varnothing = p)$, the shifted trees $\tau_1\theta, \dots, \tau_p\theta$ are independent and distributed according to the probability measure Π_γ.

Recall that if $r > 0$ and $\mathcal{T}$ is a compact rooted $\mathbb{R}$-tree with metric d, we write $r\mathcal{T}$ for the same tree equipped with the metric rd. The following result is Theorem 4.1 in [6].

THEOREM 2.3. *Let $(\gamma_n)_{n\ge1}$ be a sequence of critical or subcritical offspring distributions. For every $n \ge 1$, let us denote by X^n a Galton–Watson process with offspring distribution γ_n, started at $X^n_0 = n$. Let $(m_n)_{n\ge1}$ be a nondecreasing sequence of positive integers converging to infinity. We define a sequence of processes $(Y^n)_{n\ge1}$ by setting, for every $t \ge 0$ and $n \ge 1$,*

$$Y^n_t = n^{-1} X^n_{[m_n t]}.$$

Assume that, for every $t \ge 0$, $(Y^n_t)_{n\ge1}$ converges in distribution to Y_t where $Y = (Y_t, t \ge 0)$ is a ψ-CSBP which becomes extinct almost surely. Assume furthermore that for every $\delta > 0$,

$$\liminf_{n\to\infty} \mathbb{P}(Y^n_\delta = 0) > 0.$$

Then, for every $a > 0$, the law of the $\mathbb{R}$-tree $m_n^{-1}\mathcal{T}^\theta$ under the probability measure $\Pi_{\gamma_n}(\cdot \mid \mathcal{H}(\theta) \ge [am_n])$ converges as $n \to \infty$ to the probability measure $\Theta_\psi(\cdot \mid \mathcal{H}(\mathcal{T}) > a)$ in the sense of weak convergence of measures in the space $\mathbb{T}$.

3. Proof of Theorem 1.1. Let Θ be an infinite measure on $(\mathbb{T}, d_{GH})$ satisfying the assumptions of Theorem 1.1. Clearly Θ is σ-finite.

We start with two important lemmas that will be used throughout this section. Let us first define $v:(0,\infty)\to(0,\infty)$ by $v(t)=\Theta(\mathcal{H}(\mathcal{T})>t)$ for every $t>0$. Recall that for every $t>0$, we denote by Θ^t the probability measure $\Theta(\cdot \mid \mathcal{H}(\mathcal{T})>t)$.

LEMMA 3.1. *The function* v *is nonincreasing*, *continuous and verifies*

$$v(t)\underset{t\to 0}{\longrightarrow}\infty \quad \text{and} \quad v(t)\underset{t\to\infty}{\longrightarrow}0.$$

PROOF. We only have to prove the continuity of v. To this end, we argue by contradiction and assume that there exists $t>0$ such that $\Theta(\mathcal{H}(\mathcal{T})=t)>0$. Let $s>0$ and $u\in(0,t)$ such that $v(u)>v(t)$. From the regenerative property (R), we have

$$\begin{aligned}
&\Theta^s(\mathcal{H}(\mathcal{T})=s+t)\\
&\quad=\Theta^s\big(\Theta^s(\mathcal{H}(\mathcal{T})=s+t \mid Z(s,s+u))\big)\\
&\quad=\Theta^s\big(\big(\Theta^u(\mathcal{H}(\mathcal{T})\le t)\big)^{Z(s,s+u)}-\big(\Theta^u(\mathcal{H}(\mathcal{T})<t)\big)^{Z(s,s+u)}\big)\\
&\quad\ge\Theta^s\big(\Theta^u(\mathcal{H}(\mathcal{T})=t)\big(\Theta^u(\mathcal{H}(\mathcal{T})\le t)\big)^{Z(s,s+u)-1}\big)\\
&\quad=\frac{\Theta(\mathcal{H}(\mathcal{T})=t)}{v(u)}\Theta^s\left(\left(1-\frac{v(t)}{v(u)}\right)^{Z(s,s+u)-1}\right)\\
&\quad>0.
\end{aligned}$$

We have shown that $\Theta(\mathcal{H}(\mathcal{T})=t+s)>0$ for every $s>0$. This is absurd since Θ is σ-finite. □

LEMMA 3.2. *For every* $t>0$ *and* $0<a<b$, *the conditional law of the random variable* $Z(t,t+b)$, *under the probability measure* Θ^t *and given* $Z(t,t+a)$, *is a binomial distribution with parameters* $Z(t,t+a)$ *and* $v(b)/v(a)$ (*where we define the binomial distribution with parameters* 0 *and* $p\in[0,1]$ *as the Dirac measure* δ_0).

PROOF. This is a straightforward consequence of the regenerative property (R). □

3.1. *The CSBP derived from* Θ. In this section, we consider a random forest of trees derived from a Poisson point measure with intensity Θ. We associate with this forest a family of Galton–Watson processes. We then construct local times at every level $a>0$ as limits of the rescaled Galton–Watson processes. Finally we show that the local time process is a CSBP.

Let us now fix the framework. We consider a probability space $(\Omega, \mathbb{P})$ and on this space a Poisson point measure $\mathcal{N} = \sum_{i \in I} \delta_{\mathcal{T}_i}$ on $\mathbb{T}$, whose intensity is the measure Θ.

3.1.1. *A family of Galton–Watson trees.* We start with some notation that we need in the first lemma. We consider on another probability space $(\Omega', \mathbb{P}')$, a collection $(\theta_\xi, \xi \in \mathbb{A})$ of independent $\mathcal{A}$-valued random variables such that for every $\xi \in \mathbb{A}$, θ_ξ is distributed uniformly over $\mathbb{p}^{-1}(\xi)$. In what follows, to simplify notation, we identify an element ξ of the set $\mathbb{A}$ with the subset $\mathbb{p}^{-1}(\xi)$ of $\mathcal{A}$. Recall the notation $\mathbb{T}^{(\varepsilon)}$ and the definition of $\xi^\varepsilon(\mathcal{T})$ before Lemma 2.2.

LEMMA 3.3. *Let us define for every $\varepsilon > 0$, a mapping $\theta^{(\varepsilon)}$ from $\mathbb{T}^{(\varepsilon)} \times \Omega'$ into $\mathcal{A}$ by*

$$\theta^{(\varepsilon)}(\mathcal{T}, \omega) = \theta_{\xi^\varepsilon(\mathcal{T})}(\omega).$$

Then for every positive integer p, the law of the random variable $\theta^{(\varepsilon)}$ under the probability measure $\Theta^{p\varepsilon} \otimes \mathbb{P}'$ is $\Pi_{\mu_\varepsilon}(\cdot \mid \mathcal{H}(\theta) \geq p-1)$ where μ_ε denotes the law of $Z(\varepsilon, 2\varepsilon)$ under Θ^ε.

PROOF. Since $\{\mathcal{H}(\mathcal{T}) > p\varepsilon\} \times \Omega' = \{\mathcal{H}(\theta^{(\varepsilon)}) \geq p-1\}$ for every $p \geq 1$, it suffices to show the result for $p = 1$. Let k be a nonnegative integer. According to the construction of $\xi^\varepsilon(\mathcal{T})$, we have

$$\Theta^\varepsilon \otimes \mathbb{P}'\big(k_\varnothing\big(\theta^{(\varepsilon)}\big) = k\big) = \Theta^\varepsilon\big(Z(\varepsilon, 2\varepsilon) = k\big) = \mu_\varepsilon(k).$$

Let us fix $k \geq 1$ with $\mu_\varepsilon(k) > 0$. Let $F : \mathcal{A}^k \to \mathbb{R}_+$ be a symmetric function. Then we have

$$\begin{aligned}
&\Theta^\varepsilon \otimes \mathbb{P}'\big(F\big(\tau_1\theta^{(\varepsilon)}, \dots, \tau_k\theta^{(\varepsilon)}\big) \,|\, k_\varnothing\big(\theta^{(\varepsilon)}\big) = k\big) \\
&\quad = \Theta^\varepsilon \otimes \mathbb{P}'\left(\sum_{\theta \in \xi^\varepsilon(\mathcal{T})} F(\tau_1\theta, \dots, \tau_k\theta)\mathbb{1}_{\{\theta_{\xi^\varepsilon(\mathcal{T})} = \theta\}} \,\middle|\, Z(\varepsilon, 2\varepsilon) = k\right) \\
&\quad = \Theta^\varepsilon\left((\#\xi^\varepsilon(\mathcal{T}))^{-1} \sum_{\theta \in \xi^\varepsilon(\mathcal{T})} F(\tau_1\theta, \dots, \tau_k\theta) \,\middle|\, Z(\varepsilon, 2\varepsilon) = k\right). \qquad (8)
\end{aligned}$$

On the event $\{Z(\varepsilon, 2\varepsilon) = k\}$, we write $\mathcal{T}^1, \dots, \mathcal{T}^k$ for the k subtrees of $\mathcal{T}$ above level ε with height greater than ε. Then, formula (4) and the regenerative property (R) yield

$$\begin{aligned}
&\Theta^\varepsilon\left((\#\xi^\varepsilon(\mathcal{T}))^{-1} \sum_{\theta \in \xi^\varepsilon(\mathcal{T})} F(\tau_1\theta, \dots, \tau_k\theta) \,\middle|\, Z(\varepsilon, 2\varepsilon) = k\right) \\
&\quad = \Theta^\varepsilon\Big((\#\xi^\varepsilon(\mathcal{T}^1))^{-1} \cdots (\#\xi^\varepsilon(\mathcal{T}^k))^{-1}
\end{aligned}$$

$$\times \sum_{\theta_1 \in \xi^\varepsilon(\mathcal{T}^1)} \cdots \sum_{\theta_k \in \xi^\varepsilon(\mathcal{T}^k)} F(\theta_1, \ldots, \theta_k) \Big| Z(\varepsilon, 2\varepsilon) = k \Big)$$

$$= \int \Theta^\varepsilon(d\mathcal{T}_1) \cdots \Theta^\varepsilon(d\mathcal{T}_k) (\#\xi^\varepsilon(\mathcal{T}_1))^{-1} \cdots (\#\xi^\varepsilon(\mathcal{T}_k))^{-1}$$

$$\times \sum_{\theta_1 \in \xi^\varepsilon(\mathcal{T}_1)} \cdots \sum_{\theta_k \in \xi^\varepsilon(\mathcal{T}_k)} F(\theta_1, \ldots, \theta_k)$$

$$= \int \Theta^\varepsilon \otimes \mathbb{P}'(d\mathcal{T}_1, d\omega'_1) \cdots \Theta^\varepsilon \otimes \mathbb{P}'(d\mathcal{T}_k, d\omega'_k)$$

$$\times F\big(\theta^{(\varepsilon)}(\mathcal{T}_1, \omega'_1), \ldots, \theta^{(\varepsilon)}(\mathcal{T}_k, \omega'_k)\big),$$

as in (8). We have thus proved that

$$\Theta^\varepsilon \otimes \mathbb{P}'\big(F\big(\tau_1\theta^{(\varepsilon)}, \ldots, \tau_k\theta^{(\varepsilon)}\big) | k_\varnothing\big(\theta^{(\varepsilon)}\big) = k\big)$$

$$= \int \Theta^\varepsilon \otimes \mathbb{P}'(d\mathcal{T}_1, d\omega'_1) \cdots \Theta^\varepsilon \otimes \mathbb{P}'(d\mathcal{T}_k, d\omega'_k) \tag{9}$$

$$\times F\big(\theta^{(\varepsilon)}(\mathcal{T}_1, \omega'_1), \ldots, \theta^{(\varepsilon)}(\mathcal{T}_k, \omega'_k)\big).$$

Note that for every permutation φ of the set $\{1, \ldots, k\}$, the two k-tuples $(\tau_{\varphi(1)}\theta^{(\varepsilon)}, \ldots, \tau_{\varphi(k)}\theta^{(\varepsilon)})$ and $(\tau_1\theta^{(\varepsilon)}, \ldots, \tau_k\theta^{(\varepsilon)})$ have the same distribution under $\Theta^\varepsilon \otimes \mathbb{P}'$. Then, (9) means that the law of $\theta^{(\varepsilon)}$ under $\Theta^\varepsilon \otimes \mathbb{P}'$ satisfies the branching property of the Galton–Watson trees. This completes the proof of the desired result. □

Recall that $\sum_{i \in I} \delta_{\mathcal{T}_i}$ is a Poisson point measure on $\mathbb{T}$ with intensity Θ. Let us now set, for every $t, h > 0$,

$$\mathcal{Z}(t, t+h) = \sum_{i \in I} Z(t, t+h)(\mathcal{T}_i). \tag{10}$$

Notice that the above sum is almost surely finite. For every $\varepsilon > 0$, we define a process $\mathcal{X}^\varepsilon = (\mathcal{X}^\varepsilon_k, k \geq 0)$ on $(\Omega, \mathbb{P})$ by the formula

$$\mathcal{X}^\varepsilon_k = \mathcal{Z}\big(k\varepsilon, (k+1)\varepsilon\big), \qquad k \geq 0. \tag{11}$$

PROPOSITION 3.4. *For every $\varepsilon > 0$, the process $\mathcal{X}^\varepsilon$ is a Galton–Watson process whose initial distribution is the Poisson distribution with parameter $v(\varepsilon)$ and whose offspring distribution is μ_ε.*

PROOF. We first observe that $\mathcal{X}^\varepsilon_0 = \mathcal{N}(\{\mathcal{H}(\mathcal{T}) > \varepsilon\})$ is Poisson with parameter $\Theta(\mathcal{H}(\mathcal{T}) > \varepsilon) = v(\varepsilon)$. Then let p be a positive integer. We know from a classical property of Poisson measures that, under the probability measure $\mathbb{P}$ and conditionally on the event $\{\mathcal{X}^\varepsilon_0 = p\}$, the atoms of $\mathcal{N}$ that belong to the set $\mathbb{T}^\varepsilon$ are distributed as p independent variables with distribution Θ^ε. Furthermore, it follows from Lemma 3.3 that under Θ^ε, the process $(Z(k\varepsilon,(k+1)\varepsilon))_{k\geq 0}$ is a Galton–Watson process started at one with offspring distribution μ_ε. This completes the proof. $\square$

As a consequence, we get the next proposition, which we will use throughout this work.

PROPOSITION 3.5. *For every $t>0$ and $h>0$, we have $\Theta(Z(t,t+h))\leq v(h)$.*

PROOF. Since compact $\mathbb{R}$-trees have finite height, the Galton–Watson process $\mathcal{X}^\varepsilon$ dies out $\mathbb{P}$ a.s. This implies that μ_ε is critical or subcritical so that $(\mathcal{X}^\varepsilon_k, k\geq 0)$ is a supermartingale. Let $t,h>0$. We can find $\varepsilon>0$ and $k\in\mathbb{N}$ such that $t=k\varepsilon$ and $\varepsilon\leq h$. Thus we have,

$$\Theta\big(Z(t,t+\varepsilon)\big)=\Theta\big(Z\big(k\varepsilon,(k+1)\varepsilon\big)\big)=\mathbb{E}(\mathcal{X}^\varepsilon_k)\leq\mathbb{E}(\mathcal{X}^\varepsilon_0)=v(\varepsilon). \tag{12}$$

Using Lemma 3.2 and (12), we get

$$\Theta\big(Z(t,t+h)\big)=\Theta\bigg(Z(t,t+\varepsilon)\frac{v(h)}{v(\varepsilon)}\bigg)\leq v(h).$$

$\square$

3.1.2. *A local time process.*

PROPOSITION 3.6. *For every $t\geq 0$, there exists a random variable L_t on the space $\mathbb{T}$ such that Θ a.e.,*

$$\frac{Z(t,t+h)}{v(h)}\underset{h\to 0}{\longrightarrow}L_t.$$

PROOF. Let us start with the case $t=0$. As $Z(0,h)=\mathbb{1}_{\{\mathcal{H}(\mathcal{T})>h\}}$ for every $h>0$, Lemma 3.1 gives $v(h)^{-1}Z(0,h)\to 0$ Θ a.e. as $h\to 0$, so we set $L_0=0$.

Let us now fix $t>0$. Thanks to Lemma 3.1, we can define a decreasing sequence $(\varepsilon_n)_{n\geq 1}$ by the condition $v(\varepsilon_n)=n^4$ for every $n\geq 1$. We claim that there exists a random variable L_t on the space $\mathbb{T}$ such that, Θ a.e.,

$$\frac{Z(t,t+\varepsilon_n)}{n^4}\underset{n\to\infty}{\longrightarrow}L_t. \tag{13}$$

Indeed, using Lemma 3.2, we have, for every $n\geq 1$,

$$\begin{aligned}
&\Theta^t\bigg(\bigg|\frac{Z(t,t+\varepsilon_n)}{n^4}-\frac{Z(t,t+\varepsilon_{n+1})}{(n+1)^4}\bigg|^2\bigg)\\
&\quad=\Theta^t\bigg(\frac{1}{n^8}\Theta^t\bigg(\bigg|Z(t,t+\varepsilon_n)-\frac{n^4}{(n+1)^4}Z(t,t+\varepsilon_{n+1})\bigg|^2\bigg|Z(t,t+\varepsilon_{n+1})\bigg)\bigg)\\
&\quad\leq\Theta^t\bigg(\frac{Z(t,t+\varepsilon_{n+1})}{4n^8}\bigg).
\end{aligned}$$

From Proposition 3.5 and the definition of ε_{n+1} we obtain

$$\Theta^t\left(\left|\frac{Z(t,t+\varepsilon_n)}{n^4}-\frac{Z(t,t+\varepsilon_{n+1})}{(n+1)^4}\right|^2\right)\le\frac{(n+1)^4}{4v(t)n^8}. \tag{14}$$

Thanks to the Cauchy–Schwarz inequality, we get

$$\Theta^t\left(\left|\frac{Z(t,t+\varepsilon_n)}{n^4}-\frac{Z(t,t+\varepsilon_{n+1})}{(n+1)^4}\right|\right)\le\frac{(n+1)^2}{2n^4\sqrt{v(t)}}\le\frac{2}{n^2\sqrt{v(t)}}. \tag{15}$$

The bound (15) implies

$$\Theta\left(\sum_{n=1}^{\infty}\left|\frac{Z(t,t+\varepsilon_n)}{n^4}-\frac{Z(t,t+\varepsilon_{n+1})}{(n+1)^4}\right|\right)<\infty.$$

In particular, Θ a.e.,

$$\sum_{n=1}^{\infty}\left|\frac{Z(t,t+\varepsilon_n)}{n^4}-\frac{Z(t,t+\varepsilon_{n+1})}{(n+1)^4}\right|<\infty.$$

Our claim (13) follows.

For every $h\in(0,\varepsilon_1]$, we can find $n\ge 1$ such that $\varepsilon_{n+1}\le h\le\varepsilon_n$. Then, we have $Z(t,t+\varepsilon_n)\le Z(t,t+h)\le Z(t,t+\varepsilon_{n+1})$ Θ a.e., and $n^4\le v(h)\le(n+1)^4$ so that

$$\frac{Z(t,t+\varepsilon_n)}{(n+1)^4}\le\frac{Z(t,t+h)}{v(h)}\le\frac{Z(t,t+\varepsilon_{n+1})}{n^4}.$$

We then deduce from (13) that Θ a.e.,

$$\frac{Z(t,t+h)}{v(h)}\underset{h\to 0}{\longrightarrow}L_t$$

which completes the proof. $\square$

DEFINITION 3.1. We define a process $\mathcal{L}=(\mathcal{L}_t,t\ge 0)$ on $(\Omega,\mathbb{P})$ by setting $\mathcal{L}_0=1$ and for every $t>0$,

$$\mathcal{L}_t=\sum_{i\in I}L_t(\mathcal{T}_i).$$

Notice that $L_t(\mathcal{T})=0$ if $\mathcal{H}(\mathcal{T})\le t$ so that the above sum is almost surely finite. Recall the definition of $\mathcal{Z}$ which is given by (10).

COROLLARY 3.7. *For every $t\ge 0$, we have $\mathbb{P}$ a.s.*

$$\frac{\mathcal{Z}(t,t+h)}{v(h)}\underset{h\to 0}{\longrightarrow}\mathcal{L}_t.$$

Moreover, this convergence holds in $\mathbb{L}^1(\mathbb{P})$ uniformly in $t\in[0,\infty)$.

PROOF. The first assertion for $t=0$ is a consequence of the definition of $\mathcal{L}_0$ together with simple estimates for the Poisson distribution and the case $t>0$ is an immediate consequence of Proposition 3.6.

Let us focus on the second assertion. From the second moment formula for Poisson measures, we get, for every $t \geq 0$ and $n \geq 1$,

$$\begin{aligned}
\mathbb{E}\biggl(\biggl(\frac{\mathcal{Z}(t,t+\varepsilon_n)}{n^4}-\frac{\mathcal{Z}(t,t+\varepsilon_{n+1})}{(n+1)^4}\biggr)^2\biggr)
&= \Theta\biggl(\biggl(\frac{Z(t,t+\varepsilon_n)}{n^4}-\frac{Z(t,t+\varepsilon_{n+1})}{(n+1)^4}\biggr)^2\biggr)\\
&\quad+\biggl(\Theta\biggl(\frac{Z(t,t+\varepsilon_n)}{n^4}-\frac{Z(t,t+\varepsilon_{n+1})}{(n+1)^4}\biggr)\biggr)^2\\
&= \Theta\biggl(\biggl(\frac{Z(t,t+\varepsilon_n)}{n^4}-\frac{Z(t,t+\varepsilon_{n+1})}{(n+1)^4}\biggr)^2\biggr),
\end{aligned}$$

where the last equality follows from Lemma 3.2. Now, we have

$$\begin{aligned}
\Theta\biggl(\biggl(\frac{Z(0,\varepsilon_n)}{n^4}-\frac{Z(0,\varepsilon_{n+1})}{(n+1)^4}\biggr)^2\biggr) &= \Theta\biggl(\biggl(\frac{\mathbb{1}_{\{\mathcal{H}(\mathcal{T})>\varepsilon_n\}}}{n^4}-\frac{\mathbb{1}_{\{\mathcal{H}(\mathcal{T})>\varepsilon_{n+1}\}}}{(n+1)^4}\biggr)^2\biggr)\\
&= \frac{1}{n^4}-\frac{1}{(n+1)^4}
\end{aligned}$$

and for every $t>0$, thanks to the bound (14),

$$\Theta\biggl(\biggl(\frac{Z(t,t+\varepsilon_n)}{n^4}-\frac{Z(t,t+\varepsilon_{n+1})}{(n+1)^4}\biggr)^2\biggr) \leq \frac{(n+1)^4}{4n^8}.$$

So for every $t \geq 0$ and $n \geq 1$, we have from the Cauchy–Schwarz inequality

$$\mathbb{E}\biggl(\biggl|\frac{\mathcal{Z}(t,t+\varepsilon_n)}{n^4}-\frac{\mathcal{Z}(t,t+\varepsilon_{n+1})}{(n+1)^4}\biggr|\biggr) \leq \frac{(n+1)^2}{n^4}. \tag{16}$$

Then $n^{-4}\mathcal{Z}(t,t+\varepsilon_n) \to \mathcal{L}_t$ in $\mathbb{L}^1$ as $n \to \infty$ and, for every $n \geq 2$,

$$\mathbb{E}\biggl(\biggl|\frac{\mathcal{Z}(t,t+\varepsilon_n)}{n^4}-\mathcal{L}_t\biggr|\biggr) \leq \sum_{k=n}^{\infty}\frac{(k+1)^2}{k^4} \leq \sum_{k=n}^{\infty}\frac{4}{k^2} \leq \frac{8}{n}.$$

In the same way as in the proof of (16), we have the following inequality: If $h \in (0,\varepsilon_1]$, $t \geq 0$ and n is a positive integer such that $\varepsilon_{n+1} \leq h \leq \varepsilon_n$,

$$\mathbb{E}\biggl(\biggl|\frac{\mathcal{Z}(t,t+\varepsilon_n)}{n^4}-\frac{\mathcal{Z}(t,t+h)}{v(h)}\biggr|\biggr) \leq \frac{\sqrt{v(h)}}{n^4} \leq \frac{16}{\sqrt{v(h)}}.$$

Then, for every $h \in (0,\varepsilon_2]$ and $t \geq 0$, we get

$$\mathbb{E}\biggl(\biggl|\frac{\mathcal{Z}(t,t+h)}{v(h)}-\mathcal{L}_t\biggr|\biggr) \leq 16\bigl(v(h)^{-1/2}+v(h)^{-1/4}\bigr),$$

which completes the proof. $\square$

We will now establish a regularity property of the process $(\mathcal{L}_t, t \geq 0)$.

PROPOSITION 3.8. *The process* $(\mathcal{L}_t, t \geq 0)$ *admits a modification, denoted by* $(\widetilde{\mathcal{L}}_t, t \geq 0)$, *which is right-continuous with left-limits, and which has no fixed discontinuities.*

PROOF. We start with two lemmas.

LEMMA 3.9. *There exists* $\lambda \geq 0$ *such that* $\mathbb{E}(\mathcal{L}_t) = e^{-\lambda t}$ *for every* $t \geq 0$.

PROOF. We claim that the function $t \in [0, +\infty) \mapsto \mathbb{E}(\mathcal{L}_t)$ is multiplicative, meaning that for every $t, s \geq 0$, $\mathbb{E}(\mathcal{L}_{t+s}) = \mathbb{E}(\mathcal{L}_t)\mathbb{E}(\mathcal{L}_s)$. As $\mathcal{L}_0 = 1$ by definition, $\mathbb{E}(\mathcal{L}_0) = 1$. Let $t, s > 0$ and $0 < h < s$. Let us denote by $\mathcal{T}^1, \dots, \mathcal{T}^{Z(t,t+h)}$ the subtrees of $\mathcal{T}$ above level t with height greater than h. Then, using the regenerative property (R), we can write

$$\begin{aligned}\Theta\big(Z(t+s, t+s+h)\big) &= \Theta\left(\sum_{i=1}^{Z(t,t+h)} Z(s, s+h)(\mathcal{T}^i)\right) \\ &= \Theta\big(Z(t, t+h)\big)\Theta^h\big(Z(s, s+h)\big),\end{aligned}$$

which implies

$$\mathbb{E}\big(\mathcal{Z}(t+s, t+s+h)\big) = \mathbb{E}\big(\mathcal{Z}(t, t+h)\big)\mathbb{E}\left(\frac{\mathcal{Z}(s, s+h)}{v(h)}\right). \tag{17}$$

Thus, dividing by $v(h)$ and letting $h \to 0$ in (17), we get our claim from Corollary 3.7. Moreover, thanks to Proposition 3.5 and Corollary 3.7, we know that $\mathbb{E}(\mathcal{L}_t) \leq 1$ for every $t \geq 0$. Then, we obtain in particular that the function $t \in [0, \infty) \mapsto \mathbb{E}(\mathcal{L}_t)$ is nonincreasing.

To complete the proof, we have to check that $\mathbb{E}(\mathcal{L}_t) > 0$ for every $t > 0$. If we assume that $\mathbb{E}(\mathcal{L}_t) = 0$ for some $t > 0$ then $L_t = 0$, Θ a.e. Let $s, h > 0$ such that $0 < h < s$. With the same notation as in the beginning of the proof, we can write

$$\begin{aligned}\Theta\big(\mathcal{H}(\mathcal{T}) > t+s\big) &= \Theta\big(\exists i \in \{1, \dots, Z(t, t+h)\}, \mathcal{H}(\mathcal{T}^i) > s\big) \\ &= \Theta\left(1 - \left(1 - \frac{v(s)}{v(h)}\right)^{Z(t,t+h)}\right).\end{aligned} \tag{18}$$

Now, thanks to Proposition 3.6, Θ a.e.,

$$\left(1 - \frac{v(s)}{v(h)}\right)^{Z(t,t+h)} \underset{h \to 0}{\longrightarrow} \exp(-L_t v(s)) = 1.$$

Moreover, Θ a.e.,

$$1-\left(1-\frac{v(s)}{v(h)}\right)^{Z(t,t+h)}\leq \mathbb{1}_{\{\mathcal{H}(\mathcal{T})>t\}}.$$

Then, using dominated convergence in (18) as $h\to 0$, we obtain that $\Theta(\mathcal{H}(\mathcal{T})>t+s)=0$ which contradicts the assumptions of Theorem 1.1. □

LEMMA 3.10. *Let us denote by $D=\{k2^{-n},k\geq 1,n\geq 0\}$ the set of positive dyadic numbers and define $\mathcal{G}_t=\sigma(\mathcal{L}_s,s\in D,s\leq t)$ for every $t\in D$. Then $(\mathcal{L}_t,t\in D)$ is a nonnegative supermartingale with respect to the filtration $(\mathcal{G}_t,t\in D)$.*

PROOF. Let p be a positive integer, let $s_1,\dots,s_p,s,t\in D$ such that $s_1<\cdots<s_p\leq s<t$ and let $f:\mathbb{R}^p\to\mathbb{R}_+$ be a bounded continuous function. We can find a positive integer n such that 2^nt, 2^ns, and 2^ns_i for $i\in\{1,\dots,p\}$ are nonnegative integers. Recall the definition of $\mathcal{X}^\varepsilon$ which is given by (11). From Proposition 3.4, the process $\mathcal{X}^{2^{-n}}$ is a subcritical Galton–Watson process, so

$$\mathbb{E}(\mathcal{X}^{2^{-n}}_{2^nt}f(\mathcal{X}^{2^{-n}}_{2^ns_1},\dots,\mathcal{X}^{2^{-n}}_{2^ns_p}))\leq \mathbb{E}(\mathcal{X}^{2^{-n}}_{2^ns}f(\mathcal{X}^{2^{-n}}_{2^ns_1},\dots,\mathcal{X}^{2^{-n}}_{2^ns_p})).$$

Therefore we have also,

$$\begin{aligned}&\mathbb{E}\left(\frac{\mathcal{Z}(t,t+2^{-n})}{v(2^{-n})}f\left(\frac{\mathcal{Z}(s_1,s_1+2^{-n})}{v(2^{-n})},\dots,\frac{\mathcal{Z}(s_p,s_p+2^{-n})}{v(2^{-n})}\right)\right)\\ &\quad\leq \mathbb{E}\left(\frac{\mathcal{Z}(s,s+2^{-n})}{v(2^{-n})}f\left(\frac{\mathcal{Z}(s_1,s_1+2^{-n})}{v(2^{-n})},\dots,\frac{\mathcal{Z}(s_p,s_p+2^{-n})}{v(2^{-n})}\right)\right).\end{aligned}\tag{19}$$

We can then use Corollary 3.7 to obtain

$$\mathbb{E}(\mathcal{L}_tf(\mathcal{L}_{s_1},\dots,\mathcal{L}_{s_p}))\leq \mathbb{E}(\mathcal{L}_sf(\mathcal{L}_{s_1},\dots,\mathcal{L}_{s_p})).$$ □

We now complete the proof of Proposition 3.8. Let us set, for every $t\geq 0$,

$$\widetilde{\mathcal{G}}_t=\bigcap_{s>t,s\in D}\mathcal{G}_s.$$

From Lemma 3.10 and classical results on supermartingales, we can define a right-continuous supermartingale $(\widetilde{\mathcal{L}}_t,t\geq 0)$ with respect to the filtration $(\widetilde{\mathcal{G}}_t,t\geq 0)$ by setting, for every $t\geq 0$,

$$\widetilde{\mathcal{L}}_t=\lim_{s\downarrow t,s\in D}\mathcal{L}_s,\tag{20}$$

where the limit holds $\mathbb{P}$ a.s. and in $\mathbb{L}^1$ (see e.g., Chapter VI in [4] for more details). We claim that $(\widetilde{\mathcal{L}}_t,t\geq 0)$ is a càdlàg modification of $(\mathcal{L}_t,t\geq 0)$ with no fixed discontinuities.

We first prove that $(\widetilde{\mathcal{L}}_t, t \geq 0)$ is a modification of $(\mathcal{L}_t, t \geq 0)$. For every $t \geq 0$ and every sequence $(s_n)_{n \geq 0}$ in D such that $s_n \downarrow t$ as $n \uparrow \infty$, we have thanks to (20) and Lemma 3.9,

$$\mathbb{E}(\widetilde{\mathcal{L}}_t) = \lim_{n \to \infty} \mathbb{E}(\mathcal{L}_{s_n}) = \mathbb{E}(\mathcal{L}_t).$$

Let us now show that for every $t \geq 0$, $\mathcal{L}_t \leq \widetilde{\mathcal{L}}_t$ $\mathbb{P}$ a.s. Let $\alpha, \varepsilon > 0$ and $\delta \in (0, 1)$. Thanks to Corollary 3.7, we can find $h_0 > 0$ such that for every $h \in (0, h_0)$ and $n \geq 0$,

$$\mathbb{E}\left(\left|\frac{\mathcal{Z}(t, t+h)}{v(h)} - \mathcal{L}_t\right|\right) \leq \varepsilon\alpha \quad \text{and} \quad \mathbb{E}\left(\left|\frac{\mathcal{Z}(s_n, s_n+h)}{v(h)} - \mathcal{L}_{s_n}\right|\right) \leq \varepsilon\alpha.$$

From Lemma 3.1, we may choose $h \in (0, h_0)$ and $n_0 \geq 0$ such that $s_n - t + h \leq h_0$ and $v(h) \leq (1+\delta) v(s_n - t + h)$ for every $n \geq n_0$. We notice that $\mathcal{Z}(t, s_n + h) \leq \mathcal{Z}(s_n, s_n + h)$ so that, for every $n \geq n_0$,

$$\begin{aligned}
&\mathbb{P}\big(\mathcal{L}_t > (1+\delta)\mathcal{L}_{s_n} + \varepsilon\big) \\
&\quad \leq \mathbb{P}\left(\mathcal{L}_t - \frac{\mathcal{Z}(t, s_n+h)}{v(s_n - t + h)} > (1+\delta)\mathcal{L}_{s_n} - (1+\delta)\frac{\mathcal{Z}(s_n, s_n+h)}{v(h)} + \varepsilon\right) \\
&\quad \leq 2\varepsilon^{-1}\mathbb{E}\left(\left|\frac{\mathcal{Z}(t, s_n+h)}{v(s_n - t + h)} - \mathcal{L}_t\right|\right) + 2\varepsilon^{-1}(1+\delta)\mathbb{E}\left(\left|\frac{\mathcal{Z}(s_n, s_n+h)}{v(h)} - \mathcal{L}_{s_n}\right|\right) \\
&\quad \leq 6\alpha.
\end{aligned}$$

We have thus shown that

$$\mathbb{P}\big(\mathcal{L}_t > (1+\delta)\mathcal{L}_{s_n} + \varepsilon\big) \underset{n \to \infty}{\longrightarrow} 0. \tag{21}$$

So, $\mathbb{P}(\mathcal{L}_t - (1+\delta)\widetilde{\mathcal{L}}_t > \varepsilon) = 0$ for every $\varepsilon > 0$, implying that $\mathcal{L}_t \leq (1+\delta)\widetilde{\mathcal{L}}_t$ $\mathbb{P}$ a.s. This leads us to the claim $\mathcal{L}_t \leq \widetilde{\mathcal{L}}_t$ $\mathbb{P}$ a.s. Since we saw that $\mathbb{E}(\mathcal{L}_t) = \mathbb{E}(\widetilde{\mathcal{L}}_t)$, we have $\mathcal{L}_t = \widetilde{\mathcal{L}}_t$ $\mathbb{P}$ a.s. for every $t \geq 0$.

Now, $(\widetilde{\mathcal{L}}_t, t \geq 0)$ is a right-continuous supermartingale. Thus, $(\widetilde{\mathcal{L}}_t, t \geq 0)$ is also left-limited and we have $\mathbb{E}(\widetilde{\mathcal{L}}_t) \leq \mathbb{E}(\widetilde{\mathcal{L}}_{t-})$ for every $t > 0$. Moreover, we can prove in the same way as we did for (21) that, for every $t > 0$ and every sequence $(s_n, n \geq 0)$ in D such that $s_n \uparrow t$ as $n \uparrow \infty$,

$$\mathbb{P}\big(\widetilde{\mathcal{L}}_{s_n} > (1+\delta)\widetilde{\mathcal{L}}_t + \varepsilon\big) \underset{n \to \infty}{\longrightarrow} 0,$$

implying that $\widetilde{\mathcal{L}}_{t-} \leq \widetilde{\mathcal{L}}_t$ $\mathbb{P}$ a.s. So, $\mathcal{L}_t = \mathcal{L}_{t-}$ $\mathbb{P}$ a.s. for every $t > 0$ meaning that $(\widetilde{\mathcal{L}}_t, t \geq 0)$ has no fixed discontinuities. □

From now on, we will only deal with this càdlàg modification and we denote it by $(\mathcal{L}_t, t \geq 0)$.

3.1.3. *The CSBP.* We will prove that the suitably rescaled family of Galton–Watson processes $(\mathcal{X}^{\varepsilon})_{\varepsilon>0}$ converges to the local time $\mathcal{L}$.

Thanks to Lemma 3.1, we can define a sequence $(\eta_n)_{n\geq 1}$ by the condition $v(\eta_n) = n$ for every $n \geq 1$. We set $m_n = [\eta_n^{-1}]$ where $[x]$ denotes the integer part of x. We recall from Proposition 3.4 that $\mathcal{X}^{\eta_n}$ is a Galton–Watson process on $(\Omega, \mathbb{P})$ whose initial distribution is the Poisson distribution with parameter n. For every $n \geq 1$, we define a process $\mathcal{Y}^n = (\mathcal{Y}^n_t, t \geq 0)$ on $(\Omega, \mathbb{P})$ by the following formula,

$$\mathcal{Y}^n_t = n^{-1}\mathcal{X}^{\eta_n}_{[m_n t]}, \qquad t \geq 0.$$

PROPOSITION 3.11. *For every $t \geq 0$, $\mathcal{Y}^n_t \to \mathcal{L}_t$ in probability as $n \to \infty$.*

PROOF. The result for $t = 0$ is a consequence of the definition of $\mathcal{L}_0$ together with simple estimates for the Poisson distribution. Let $t, \delta > 0$. We can write

$$\begin{aligned}\mathbb{P}(|\mathcal{Y}^n_t - \mathcal{L}_t| > 2\delta) &\leq \mathbb{P}(|\mathcal{Y}^n_t - \mathcal{L}_{\eta_n[m_n t]}| > \delta) + \mathbb{P}(|\mathcal{L}_{\eta_n[m_n t]} - \mathcal{L}_t| > \delta)\\ &\leq \delta^{-1}\mathbb{E}(|\mathcal{Y}^n_t - \mathcal{L}_{\eta_n[m_n t]}|) + \mathbb{P}(|\mathcal{L}_{\eta_n[m_n t]} - \mathcal{L}_t| > \delta).\end{aligned}$$

Now, Corollary 3.7 and Proposition 3.8 imply respectively that

$$\mathbb{E}(|\mathcal{Y}^n_t - \mathcal{L}_{\eta_n[m_n t]}|) \underset{n\to\infty}{\longrightarrow} 0 \quad \text{and} \quad \mathbb{P}(|\mathcal{L}_{\eta_n[m_n t]} - \mathcal{L}_t| > \delta) \underset{n\to\infty}{\longrightarrow} 0,$$

which completes the proof. □

COROLLARY 3.12. *For every $t \geq 0$, the law of $\mathcal{Y}^n_t$ under $\mathbb{P}(\cdot \mid \mathcal{X}^{\eta_n}_0 = n)$ converges weakly to the law of $\mathcal{L}_t$ under $\mathbb{P}$ as $n \to \infty$.*

PROOF. For positive integers n and k, we denote by μ^n the offspring distribution of the Galton–Watson process $\mathcal{X}^{\eta_n}$, by f^n the generating function of μ^n and by f^n_k the kth iterative of f^n. Let $\lambda > 0$ and $t \geq 0$. We have,

$$\begin{aligned}\mathbb{E}(\exp(-\lambda\mathcal{Y}^n_t)) &= \sum_{p=0}^{\infty} e^{-n}\frac{n^p}{p!}(f^n_{[m_n t]}(e^{-\lambda/n}))^p\\ &= \exp(-n(1 - f^n_{[m_n t]}(e^{-\lambda/n}))).\end{aligned}$$

From Proposition 3.11, it holds that

$$\exp(-n(1 - f^n_{[m_n t]}(e^{-\lambda/n}))) \underset{n\to\infty}{\longrightarrow} \mathbb{E}(\exp(-\lambda\mathcal{L}_t)).$$

Let us set $u(t, \lambda) = -\log(\mathbb{E}[\exp(-\lambda\mathcal{L}_t)])$. It follows that,

$$n(1 - f^n_{[m_n t]}(e^{-\lambda/n})) \underset{n\to\infty}{\longrightarrow} u(t, \lambda).$$

Furthermore

$$\mathbb{E}(\exp(-\lambda\mathcal{Y}^n_t) \mid \mathcal{X}^{\eta_n}_0 = n) = (f^n_{[m_n t]}(e^{-\lambda/n}))^n.$$

Thus we obtain,

$$\mathbb{E}\big(\exp(-\lambda \mathcal{Y}_t^n) \mid \mathcal{X}_0^{\eta_n} = n\big) \underset{n\to\infty}{\longrightarrow} \exp(-u(t,\lambda)) = \mathbb{E}[\exp(-\lambda \mathcal{L}_t)]. \qquad \square$$

At this point, we can use Theorem 2.1 to assert that $(\mathcal{L}_t, t \geq 0)$ is a CSBP and that the law of $(\mathcal{Y}_t^n, t \geq 0)$ under the probability measure $\mathbb{P}(\cdot \mid \mathcal{X}_0^{\eta_n} = n)$ converges to the law of $(\mathcal{L}_t, t \geq 0)$ as $n \to \infty$ in the space of probability measures on the Skorokhod space $\mathbb{D}(\mathbb{R}_+)$. To verify the assumptions of Theorem 2.1, we need to check that there exists $\delta > 0$ such that $\mathbb{P}(\mathcal{L}_\delta > 0) > 0$. This is obvious from Lemma 3.9.

3.2. *Identification of the measure* Θ. In the previous section, we have constructed from Θ a CSBP $\mathcal{L}$, which becomes extinct almost surely. We denote by ψ the associated branching mechanism. We can consider the σ-finite measure Θ_ψ, which is the "law" of the Lévy tree associated with $\mathcal{L}$. Our goal is to show that the measures Θ and Θ_ψ coincide.

Recall that μ^n denotes the offspring distribution of the Galton–Watson process $\mathcal{X}^{\eta_n}$.

LEMMA 3.13. *For every $a > 0$, the law of the $\mathbb{R}$-tree $\eta_n \mathcal{T}^\theta$ under the probability measure $\Pi_{\mu^n}(\cdot \mid \mathcal{H}(\theta) \geq [am_n])$ converges as $n \to \infty$ to the probability measure $\Theta_\psi(\cdot \mid \mathcal{H}(\mathcal{T}) > a)$ in the sense of weak convergence of measures in the space $\mathbb{T}$.*

PROOF. We first check that, for every $\delta > 0$,

$$\liminf_{n\to\infty} \mathbb{P}(\mathcal{Y}_\delta^n = 0) > 0. \tag{22}$$

Indeed, we have

$$\mathbb{P}(\mathcal{Y}_\delta^n = 0) = \mathbb{P}\big(\mathcal{N}\big(\mathcal{H}(\mathcal{T}) > ([m_n\delta] + 1)\eta_n\big) = 0\big) = \exp\big(-v\big(([m_n\delta] + 1)\eta_n\big)\big).$$

As v is continuous, it follows that $\mathbb{P}(\mathcal{Y}_\delta^n = 0) \to \exp(-v(\delta))$ as $n \to \infty$ implying (22).

We recall that the law of $\mathcal{Y}^n$ under the probability measure $\mathbb{P}(\cdot \mid \mathcal{X}_0^{\eta_n} = n)$ converges to the law of $(\mathcal{L}_t, t \geq 0)$. Then, thanks to (22), we can apply Theorem 2.3 to get that, for every $a > 0$, the law of the $\mathbb{R}$-tree $m_n^{-1}\mathcal{T}^\theta$ under the probability measure $\Pi_{\mu^n}(\cdot \mid \mathcal{H}(\theta) \geq [am_n])$ converges to the probability measure $\Theta_\psi(\cdot \mid \mathcal{H}(\mathcal{T}) > a)$ in the sense of weak convergence of measures in the space $\mathbb{T}$. As $m_n^{-1}\eta_n \to 1$ as $n \to \infty$, we get the desired result. $\square$

We can now complete the proof of Theorem 1.1. Indeed, thanks to Lemmas 2.2 and 3.3, we can construct on the same probability space $(\mathbf{\Omega}, \mathbf{P})$, a sequence of $\mathbb{T}$-valued random variables $(\boldsymbol{\mathcal{T}}_n)_{n\geq 1}$ distributed according to $\Theta(\cdot \mid \mathcal{H}(\mathcal{T}) > ([am_n]+$

$1)\eta_n)$ and a sequence of $\mathcal{A}$-valued random variables $(\boldsymbol{\theta}_n)_{n\geq 1}$ distributed according to $\Pi_{\mu^n}(\cdot \mid \mathcal{H}(\theta) \geq [am_n])$ such that for every $n \geq 1$, $\mathbf{P}$ a.s.,

$$\mathbb{d}_{\mathrm{GH}}(\boldsymbol{\mathcal{T}}_n, \eta_n \mathcal{T}^{\boldsymbol{\theta}_n}) \leq 2\eta_n.$$

Then, using Lemma 3.13, we have $\Theta(\cdot \mid \mathcal{H}(\mathcal{T}) > ([am_n] + 1)\eta_n) \to \Theta_\psi(\cdot \mid \mathcal{H}(\mathcal{T}) > a)$ as $n \to \infty$ in the sense of weak convergence of measures on the space $\mathbb{T}$. So we get

$$\Theta\big(\cdot|\mathcal{H}(\mathcal{T}) > a\big) = \Theta_\psi\big(\cdot|\mathcal{H}(\mathcal{T}) > a\big)$$

for every $a > 0$, and thus $\Theta = \Theta_\psi$.

We conclude this part by giving the relation between the branching mechanism ψ of the CSBP $\mathcal{L}$ and the measure Θ. Recall from Section 2.1.1 the definition of the function $(u(t,\lambda), t \geq 0, \lambda \geq 0)$ and the differential equation relating ψ to $(u(t,\lambda),\ t \geq 0,\ \lambda \geq 0)$. It is proved in [5] (this is also a consequence of Definition 3.1) that for every $t > 0$ and every $\lambda \geq 0$

$$u(t,\lambda) = \Theta(1 - e^{-\lambda L_t}).$$

4. Proof of Theorem 1.2. Let Θ be a probability measure on $(\mathbb{T}, \mathbb{d}_{\mathrm{GH}})$ satisfying the assumptions of Theorem 1.2.

In this case, we define $v : [0,\infty) \to (0,\infty)$ by $v(t) = \Theta(\mathcal{H}(\mathcal{T}) > t)$ for every $t \geq 0$. Note that $v(0) = 1$ is well defined here. Recall that for every $t > 0$, we denote by Θ^t the probability measure $\Theta(\cdot \mid \mathcal{H}(\mathcal{T}) > t)$. The following two results are proved in a similar way to Lemmas 3.1 and 3.2.

LEMMA 4.1. *The function v is nonincreasing, continuous and goes to* 0 *as* $t \to \infty$.

LEMMA 4.2. *For every $t > 0$ and $0 < a < b$, the conditional law of the random variable $Z(t, t+b)$, under the probability measure Θ^t and given $Z(t, t+a)$, is a binomial distribution with parameters $Z(t, t+a)$ and $v(b)/v(a)$.*

4.1. *The DSBP derived from* Θ. We will follow the same strategy as in Section 3 but instead of a CSBP we will now construct an integer-valued branching process.

4.1.1. *A family of Galton–Watson trees.* We recall that μ_ε denotes the law of $Z(\varepsilon, 2\varepsilon)$ under the probability measure Θ^ε, and that $(\theta_\xi, \xi \in \mathbb{A})$ is a sequence of independent $\mathcal{A}$-valued random variables defined on a probability space $(\Omega', \mathbb{P}')$ such that for every $\xi \in \mathbb{A}$, θ_ξ is distributed uniformly over $\mathbb{p}^{-1}(\xi)$. The following lemma is proved in the same way as Lemma 3.3.

LEMMA 4.3. *Let us define for every $\varepsilon > 0$, a mapping $\theta^{(\varepsilon)}$ from $\mathbb{T}^{(\varepsilon)} \times \Omega'$ into $\mathcal{A}$ by*

$$\theta^{(\varepsilon)}(\mathcal{T}, \omega) = \theta_{\xi^\varepsilon(\mathcal{T})}(\omega).$$

Then for every positive integer p, the law of the random variable $\theta^{(\varepsilon)}$ under the probability measure $\Theta^{p\varepsilon} \otimes \mathbb{P}'$ is $\Pi_{\mu_\varepsilon}(\cdot \mid \mathcal{H}(\theta) \geq p-1)$.

For every $\varepsilon > 0$, we define a process $X^\varepsilon = (X_k^\varepsilon, k \geq 0)$ on $\mathbb{T}$ by the formula

$$X_k^\varepsilon = Z\big(k\varepsilon, (k+1)\varepsilon\big), \qquad k \geq 0.$$

We show in the same way as Propositions 3.4 and 3.5 the following two results.

PROPOSITION 4.4. *For every $\varepsilon > 0$, the process X^ε is under Θ a Galton–Watson process whose initial distribution is the Bernoulli distribution with parameter $v(\varepsilon)$ and whose offspring distribution is μ_ε.*

PROPOSITION 4.5. *For every $t > 0$ and $h > 0$, we have $\Theta(Z(t, t+h)) \leq v(h) \leq 1$.*

The next proposition however is particular to the finite case and will be useful in the rest of this section.

PROPOSITION 4.6. *The family of probability measures $(\mu_\varepsilon)_{\varepsilon > 0}$ converges to the Dirac measure δ_1 as $\varepsilon \to 0$. In other words,*

$$\Theta^\varepsilon\big(Z(\varepsilon, 2\varepsilon) = 1\big) \underset{\varepsilon \to 0}{\longrightarrow} 1.$$

PROOF. We first note that

$$2\Theta^\varepsilon\big(Z(\varepsilon, 2\varepsilon) \geq 1\big) - \Theta^\varepsilon(Z(\varepsilon, 2\varepsilon)) \leq \Theta^\varepsilon\big(Z(\varepsilon, 2\varepsilon) = 1\big) \leq \Theta^\varepsilon\big(Z(\varepsilon, 2\varepsilon) \geq 1\big).$$

Moreover, $\Theta^\varepsilon(Z(\varepsilon, 2\varepsilon) \geq 1) = \Theta^\varepsilon(\mathcal{H}(\mathcal{T}) > 2\varepsilon) = v(2\varepsilon)/v(\varepsilon)$ and $\Theta^\varepsilon(Z(\varepsilon, 2\varepsilon)) \leq 1$. So,

$$\frac{2v(2\varepsilon)}{v(\varepsilon)} - 1 \leq \Theta^\varepsilon\big(Z(\varepsilon, 2\varepsilon) = 1\big) \leq \frac{v(2\varepsilon)}{v(\varepsilon)}. \tag{23}$$

We let $\varepsilon \to 0$ in (23) and we use Lemma 4.1 to obtain the desired result. □

4.1.2. *Construction of the DSBP.*

PROPOSITION 4.7. *For every $t \geq 0$, there exists an integer-valued random variable L_t on the space $\mathbb{T}$ such that $\Theta(L_t) \leq 1$ and Θ a.s.,*

$$Z(t, t+h) \underset{h \downarrow 0}{\uparrow} L_t.$$

PROOF. Let $t \geq 0$. The function $h \in (0, \infty) \mapsto Z(t, t+h) \in \mathbb{Z}_+$ is nonincreasing so that there exists a random variable L_t with values in $\mathbb{Z}_+ \cup \{\infty\}$ such that, Θ a.s.,

$$Z(t, t+h) \underset{h \downarrow 0}{\uparrow} L_t.$$

Thanks to the monotone convergence theorem, we have

$$\Theta\big(Z(t, t+h)\big) \underset{h \to 0}{\longrightarrow} \Theta(L_t).$$

Now, by Proposition 4.5, $\Theta(Z(t, t+h)) \leq 1$ for every $h > 0$. Then, $\Theta(L_t) \leq 1$ which implies in particular that $L_t < \infty$ Θ a.s. $\square$

PROPOSITION 4.8. *For every $t > 0$, the following two convergences hold Θ a.s.*,

$$Z(t-h, t) \underset{h \downarrow 0}{\uparrow} L_t, \tag{24}$$

$$Z(t-h, t+h) \underset{h \downarrow 0}{\uparrow} L_t. \tag{25}$$

PROOF. Let $t > 0$ be fixed throughout this proof. By the same arguments as in the proof of Proposition 4.7, we can find a $\mathbb{Z}_+$-valued random variable $\overline{L}_t$ such that $\Theta(\overline{L}_t) \leq 1$ and $Z(t-h, t) \uparrow \overline{L}_t$ as $h \downarrow 0$, Θ a.s. If $h \in (0, t)$, we write $\mathcal{T}^1, \ldots, \mathcal{T}^{Z(t-h,t)}$ for the subtrees of $\mathcal{T}$ above level $t-h$ with height greater than h. Then, from the regenerative property (R),

$$\begin{aligned}
&\Theta\big(|Z(t, t+h) - Z(t-h, t)| \geq 1\big) \\
&\quad = \Theta\left(\Theta\left(\left|\sum_{i=1}^{Z(t-h,t)} \big(Z(h, 2h)(\mathcal{T}^i) - 1\big)\right| \geq 1 \middle| Z(t-h, t)\right)\right) \\
&\quad \leq \Theta\big(\Theta\big(|Z(h, 2h)(\mathcal{T}^i) - 1| \geq 1 \\
&\qquad\qquad \text{for some } i \in \{1, \ldots, Z(t-h, t)\} \mid Z(t-h, t)\big)\big) \\
&\quad \leq \Theta\big(Z(t-h, t)\Theta^h\big(|Z(h, 2h) - 1| \geq 1\big)\big).
\end{aligned} \tag{26}$$

Since $Z(t-h, t)\Theta^h(|Z(h, 2h) - 1| \geq 1) \leq \overline{L}_t$ Θ a.s., Proposition 4.6 and the dominated convergence theorem imply that the right-hand side of (26) goes to 0 as $h \to 0$. Thus $L_t = \overline{L}_t$ Θ a.s.

Likewise, there exists a random variable $\widehat{L}_t$ with values in $\mathbb{Z}_+$ such that, Θ a.s., $Z(t-h, t+h) \uparrow \widehat{L}_t$ as $h \downarrow 0$. Let us now notice that, for every $h > 0$, Θ a.s., $Z(t-h, t+h) \leq Z(t-h, t)$. Moreover, thanks to Lemma 4.2, we have

$$\begin{aligned}
\Theta\big(Z(t-h, t) \geq Z(t-h, t+h) + 1\big) &= 1 - \Theta\left(\left(\frac{v(2h)}{v(h)}\right)^{Z(t-h,t)}\right) \\
&\geq 1 - \Theta\left(\left(\frac{v(2h)}{v(h)}\right)^{L_t}\right).
\end{aligned} \tag{27}$$

The right-hand side of (27) tends to 0 as $h \to 0$. So $L_t = \widehat{L}_t$ Θ a.s. $\square$

We will now establish a regularity property of the process $(L_t, t \geq 0)$.

PROPOSITION 4.9. *The process $(L_t, t \geq 0)$ admits a modification which is right-continuous with left limits, and which has no fixed discontinuities.*

PROOF. We start the proof with three lemmas. The first one is proved in a similar but easier way as Lemma 3.9.

LEMMA 4.10. *There exists $\lambda \geq 0$ such that $\Theta(L_t) = e^{-\lambda t}$ for every $t \geq 0$.*

For every $n \geq 1$ and every $t \geq 0$ we set $Y_t^n = X_{[nt]}^{1/n}$.

LEMMA 4.11. *For every $t \geq 0$, $Y_t^n \to L_t$ as $n \to \infty$, Θ a.s.*

This lemma is an immediate consequence of Proposition 4.8.

LEMMA 4.12. *Let us define $G_t = \sigma(L_s, s \leq t)$ for every $t \geq 0$. Then $(L_t, t \geq 0)$ is a nonnegative supermartingale with respect to the filtration $(G_t, t \geq 0)$.*

PROOF. Let $s, t, s_1, \dots, s_p \geq 0$ such that $0 \leq s_1 \leq \cdots \leq s_p \leq s < t$ and let $f : \mathbb{R}^p \to \mathbb{R}_+$ be a bounded measurable function. For every $n \geq 1$, the offspring distribution $\mu_{1/n}$ is critical or subcritical so that $(X_k^{1/n}, k \geq 0)$ is a supermartingale. Thus we have

$$\Theta\big(X_{[nt]}^{1/n} f\big(X_{[ns_1]}^{1/n}, \dots, X_{[ns_p]}^{1/n}\big)\big) \leq \Theta\big(X_{[ns]}^{1/n} f\big(X_{[ns_1]}^{1/n}, \dots, X_{[ns_p]}^{1/n}\big)\big).$$

Lemma 4.11 yields $\Theta(L_t f(L_{s_1}, \dots, L_{s_p})) \leq \Theta(L_s f(L_{s_1}, \dots, L_{s_p}))$ since f is bounded and $X_u^{1/n} \leq L_u$ Θ a.s. for every $u \geq 0$. $\square$

Let us set, for every $t \geq 0$,

$$\widetilde{G}_t = \bigcap_{s > t} G_s.$$

Recall that D denotes the set of positive dyadic numbers. From Lemma 4.12 and classical results on supermartingales, we can define a right-continuous supermartingale $(\widetilde{L}_t, t \geq 0)$ with respect to the filtration $(\widetilde{G}_t, t \geq 0)$ by setting, for every $t \geq 0$,

$$\widetilde{L}_t = \lim_{s \downarrow t, s \in D} L_s, \tag{28}$$

where the limit holds Θ a.s. and in $\mathbb{L}^1$. In a way similar to Section 3 we can prove that $(\widetilde{L}_t, t \geq 0)$ is a càdlàg modification of $(L_t, t \geq 0)$ with no fixed discontinuities. $\square$

From now on, we will only deal with this càdlàg modification and we denote it by $(L_t, t \geq 0)$.

PROPOSITION 4.13. *$(L_t, t \geq 0)$ is a DSBP which becomes extinct Θ a.s.*

PROOF. By the same arguments as in the proof of (24), we can prove that, for every $0 < s < t$, the following convergence holds in probability under Θ,

$$Z\left(\frac{[nt]-[ns]}{n}, \frac{[nt]-[ns]+1}{n}\right) \underset{n\to\infty}{\longrightarrow} L_{t-s}. \tag{29}$$

Let $s, t, s_1, \dots, s_p \geq 0$ such that $0 \leq s_1 \leq \cdots \leq s_p \leq s < t$, $\lambda > 0$ and let $f : \mathbb{R}^p \to \mathbb{R}$ be a bounded measurable function. For every $n \geq 1$, under $\Theta^{1/n}$, $(X_k^{1/n}, k \geq 0)$ is a Galton–Watson process started at one so that

$$\begin{aligned}\Theta^{1/n}&(f(Y_{s_1}^n, \dots, Y_{s_p}^n)\exp(-\lambda Y_t^n))\\ &= \Theta^{1/n}\big(f(Y_{s_1}^n, \dots, Y_{s_p}^n)\big(\Theta^{1/n}\big(\exp\big(-\lambda X^{1/n}_{[nt]-[ns]}\big)\big)\big)^{Y_s^n}\big).\end{aligned}$$

From Lemma 4.11, (29) and dominated convergence, we get

$$\Theta(f(L_{s_1}, \dots, L_{s_p})\exp(-\lambda L_t)) = \Theta(f(L_{s_1}, \dots, L_{s_p})(\Theta(\exp(-\lambda L_{t-s})))^{L_s}).$$

Then, $(L_t, t \geq 0)$ is a continuous-time Markov chain with values in $\mathbb{Z}_+$ satisfying the branching property. Furthermore, since $\mathcal{H}(\mathcal{T}) < \infty$ Θ a.s., it is immediate that $(L_t, t \geq 0)$ becomes extinct Θ a.s. $\square$

4.2. *Identification of the probability measure Θ.* Let us now define, for every $\mathcal{T} \in \mathbb{T}$ and $t \geq 0$, $N_t(\mathcal{T}) = \#\{\sigma \in \mathcal{T} : d(\rho, \sigma) = t\}$ where we recall that ρ denotes the root of $\mathcal{T}$.

PROPOSITION 4.14. *For every $t \geq 0$, $N_t = L_t$ Θ a.s.*

Note that for every $t \geq 0$, L_t is the number of subtrees of $\mathcal{T}$ above level t.

PROOF OF PROPOSITION 4.14. Since $\Theta(\mathcal{H}(\mathcal{T}) = 0) = 0$, we have $L_0 = 1 = N_0$ Θ a.s. Let $t > 0$. First note that $N_t \geq L_t$ Θ a.s. Furthermore, thanks to Propositions 4.7, 4.8 and 4.9, for every $t > 0$, Θ a.s., there exists $h_0 > 0$ such that for every $h \in (0, h_0]$, $L_t = L_{t-h} = Z(t-h, t+h)$.

The remaining part of the argument is deterministic. We fix $t, h_0 > 0$ and a (deterministic) tree $\mathcal{T} \in \mathbb{T}$. We assume that there is a positive integer p such that for every $h \in (0, h_0]$,

$$L_t = L_{t-h} = Z(t-h, t+h) = p,$$

and we will verify that $N_t(\mathcal{T}) = p$. To this end, we argue by contradiction and assume that $N_t(\mathcal{T}) \geq p+1$. In particular we can find $p+1$ vertices $\sigma_1, \dots, \sigma_{p+1}$

such that $d(\rho,\sigma_i)=t$ for every $i\in\{1,\dots,p+1\}$. Let us denote by $\mathcal{T}^1,\dots,\mathcal{T}^p$ the p subtrees of $\mathcal{T}$ above level $t-h_0$. There exist $k,l\in\{1,\dots,p+1\}$ and $j\in\{1,\dots,p\}$ such that $\sigma_k,\sigma_l\in\mathcal{T}^j$. Let z_j be the unique vertex of $\mathcal{T}^j$ satisfying $[\![\rho,z_j]\!]=[\![\rho,\sigma_k]\!]\cap[\![\rho,\sigma_l]\!]$. We choose $c>0$ such that $d(\rho,z_j)<c<t$. Then it is not difficult to see that $\mathcal{T}$ has at least $p+1$ subtrees above level $t-h_0+c$. This is a contradiction since $L_{t-h_0+c}=p$. So $N_t(\mathcal{T})=p$, which completes the proof. □

Proposition 4.14 means that $(L_t,t\ge 0)$ is a modification of the process $(N_t,t\ge 0)$ which describes the evolution of the number of individuals in the tree. Let us denote by Q the generator of $(L_t,t\ge 0)$ which is of the form

$$Q=\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \dots\\ a\gamma(0) & -a & a\gamma(2) & a\gamma(3) & a\gamma(4) & \dots\\ 0 & 2a\gamma(0) & -2a & 2a\gamma(2) & a\gamma(3) & \dots\\ 0 & 0 & 3a\gamma(0) & -3a & 3a\gamma(2) & \dots\\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix},$$

where $a>0$ and γ is a critical or subcritical offspring distribution with $\gamma(1)=0$.

For every $t\ge 0$ we let $\mathcal{F}_t$ be the σ-field on $\mathbb{T}$ generated by the mapping $\mathcal{T}\mapsto\mathcal{T}_{\le t}$ and completed with respect to Θ. Thus $(\mathcal{F}_t,t\ge 0)$ is a filtration on $\mathbb{T}$.

LEMMA 4.15. *Let $t>0$ and $p\in\mathbb{N}$. Under Θ, conditionally on $\mathcal{F}_t$ and given $\{L_t=p\}$, the p subtrees of $\mathcal{T}$ above level t are independent and distributed according to Θ.*

PROOF. Thanks to Lemmas 2.2 and 4.3, we can construct on the same probability space $(\mathbf{\Omega},\mathbf{P})$, a sequence of $\mathbb{T}$-valued random variables $(\boldsymbol{\mathcal{T}}_n)_{n\ge1}$ distributed according to $\Theta^{1/n}$ and a sequence of $\mathbb{A}$-valued random variables $(\boldsymbol{\theta}_n)_{n\ge1}$ distributed according to $\Pi_{\mu_{1/n}}$ such that, for every $n\ge1$,

$$d_{\mathrm{GH}}(\boldsymbol{\mathcal{T}}_n,n^{-1}\mathcal{T}^{\boldsymbol{\theta}_n})\le 2n^{-1}. \tag{30}$$

For every $n\ge1$ and $k\ge0$, we define $\mathbf{X}^n_k=\#\{u\in\boldsymbol{\theta}_n:|u|=k\}$. Let $t\ge0$ and $p\ge1$, let $g:\mathbb{T}\to\mathbb{R}$ be a bounded continuous function and let $G:\mathbb{T}^p\to\mathbb{R}$ be a bounded continuous symmetric function. For $n\ge1$, on the event $\{\mathbf{X}^n_{[nt]}=p\}$, we set $\{u^n_1,\dots,u^n_p\}=\{u\in\boldsymbol{\theta}_n:|u|=[nt]\}$ and $\boldsymbol{\theta}^i_n=\tau_{u^n_i}\boldsymbol{\theta}_n$ for every $i\in\{1,\dots,p\}$. Then we can write, thanks to the branching property of Galton–Watson trees,

$$\begin{aligned}&\mathbf{E}\big(\mathbb{1}_{\{\mathbf{X}^n_{[nt]}=p\}}g\big(n^{-1}\mathcal{T}^{\boldsymbol{\theta}_n}_{\le[nt]}\big)G(n^{-1}\mathcal{T}^{\boldsymbol{\theta}^1_n},\dots,n^{-1}\mathcal{T}^{\boldsymbol{\theta}^p_n})\big)\\ &\quad=\mathbf{E}\big(\mathbb{1}_{\{\mathbf{X}^n_{[nt]}=p\}}g\big(n^{-1}\mathcal{T}^{\boldsymbol{\theta}_n}_{\le[nt]}\big)\big)(\Pi_{\mu_{1/n}})^{\otimes p}(G(n^{-1}\mathcal{T}^{\theta_1},\dots,n^{-1}\mathcal{T}^{\theta_p})),\end{aligned} \tag{31}$$

where $\theta_1,\dots,\theta_p$ denote the coordinate variables under the product measure $(\Pi_{\mu_{1/n}})^{\otimes p}$. As a consequence of (30), we see that the law of $n^{-1}\mathcal{T}^{\theta}$ under $\Pi_{\mu_{1/n}}$

converges to Θ in the sense of weak convergence of measures on the space $\mathbb{T}$. Then, thanks to Lemma 4.11, the right-hand side of (31) converges as $n \to \infty$ to

$$\Theta\big(\mathbb{1}_{\{L_t=p\}}g(\mathcal{T}_{\le t})\big)\Theta^{\otimes p}(G(\mathcal{T}_1,\dots,\mathcal{T}_p)).$$

Similarly, the left-hand side of (31) converges as $n \to \infty$ to

$$\Theta\big(\mathbb{1}_{\{L_t=p\}}g(\mathcal{T}_{\le t})G(\mathcal{T}^1,\dots,\mathcal{T}^p)\big),$$

where $\mathcal{T}^1,\dots,\mathcal{T}^p$ are the p subtrees of $\mathcal{T}$ above level t on the event $\{L_t = p\}$. This completes the proof. $\square$

Let us define $J = \inf\{t \ge 0 : L_t \ne 1\}$. Then J is an $(\mathcal{F}_t)_{t\ge 0}$-stopping time.

LEMMA 4.16. *Let $p \in \mathbb{N}$. Under Θ, given $\{L_J = p\}$, the p subtrees of $\mathcal{T}$ above level J are independent and distributed according to Θ, and are independent of J.*

PROOF. Let $p \in \mathbb{N}$, let $f : \mathbb{R}_+ \to \mathbb{R}$ be a bounded continuous function and let $G : \mathbb{T}^p \to \mathbb{R}$ be a bounded continuous symmetric function. On the event $\{L_J = p\}$, we denote by $\mathcal{T}^1,\dots,\mathcal{T}^p$ the p subtrees of $\mathcal{T}$ above level J. Let $n \ge 1$ and $k \ge 0$. On the event $\{L_{(k+1)/n} = p\}$, we denote by $\mathcal{T}^{1,(n,k)},\dots,\mathcal{T}^{p,(n,k)}$ the p subtrees of $\mathcal{T}$ above level $(k+1)/n$. On the one hand, the right continuity of the mapping $t \mapsto L_t$ gives

$$\Theta\left(\sum_{k=1}^{\infty}\mathbb{1}_{\{L_{(k+1)/n}=p\}}G\big(\mathcal{T}^{1,(n,k)},\dots,\mathcal{T}^{p,(n,k)}\big)f\big((k+1)/n\big)\mathbb{1}_{\{k/n<J\le(k+1)/n\}}\right)$$
$$\underset{n\to\infty}{\longrightarrow}\Theta\big(\mathbb{1}_{\{L_J=p\}}G(\mathcal{T}^1,\dots,\mathcal{T}^p)f(J)\big).$$

On the other hand, thanks to Lemma 4.15, we can write, for every $n \ge 1$ and $k \ge 0$,

$$\begin{aligned}&\Theta\big(\mathbb{1}_{\{L_{(k+1)/n}=p\}}G\big(\mathcal{T}^{1,(n,k)},\dots,\mathcal{T}^{p,(n,k)}\big)f\big((k+1)/n\big)\mathbb{1}_{\{k/n<J\le(k+1)/n\}}\big)\\&\quad=\Theta\big(\mathbb{1}_{\{L_{(k+1)/n}=p\}}f\big((k+1)/n\big)\mathbb{1}_{\{k/n<J\le(k+1)/n\}}\big)\Theta^{\otimes p}(G(\mathcal{T}_1,\dots,\mathcal{T}_p)).\end{aligned}$$

It follows that

$$\begin{aligned}&\Theta\big(\mathbb{1}_{\{L_J=p\}}G(\mathcal{T}^1,\dots,\mathcal{T}^p)f(J)\big)\\&\quad=\Theta\big(\mathbb{1}_{\{L_J=p\}}f(J)\big)\Theta^{\otimes p}(G(\mathcal{T}_1,\dots,\mathcal{T}_p)).\end{aligned}$$ $\square$

We can now complete the proof of Theorem 1.2. The random variable J is the first jump time of the DSBP $(L_t, t \ge 0)$ so that J is distributed according to the exponential distribution with parameter a and is independent of L_J. Thanks to Proposition 4.14, there exists $\sigma_J \in \mathcal{T}$ such that $\mathcal{T}_{\le J} = [\![\rho,\sigma_J]\!]$. Lemma 4.16 gives the last part of the description of Θ.

Another way to describe Θ is as follows: Assume that we are given on the same probability space $(\mathbf{\Omega}, \mathbf{P})$ an $\mathcal{A}$-valued random variable $\boldsymbol{\theta}$ distributed according to Π_γ and an independent sequence of independent random variables $(\mathbf{h}_u, u \in U)$ with values in $[0, \infty)$, such that each variable $\mathbf{h}_u$ is distributed according to the exponential distribution with parameter a. We set $\mathbf{T} = (\boldsymbol{\theta}, \{\mathbf{h}_u\}_{u\in\boldsymbol{\theta}})$ and $\boldsymbol{\mathcal{T}} = \mathcal{T}^{\mathbf{T}}$. Then the random variable $\boldsymbol{\mathcal{T}}$ is distributed according to Θ.

Acknowledgment. This work is a part of my Ph.D. thesis which was supervised by J. F. Le Gall and I wish to thank him for his help, support and comments.

REFERENCES

[1] ALDOUS, D. (1991). The continuum random tree. I. *Ann. Probab.* **19** 1–28. MR1085326

[2] ALDOUS, D. (1993). The continuum random tree. III. *Ann. Probab.* **21** 248–289. MR1207226

[3] ATHREYA, K. B. and NEY, P. E. (1972). *Branching Processes.* Springer, Berlin. MR0373040

[4] DELLACHERIE, C. and MEYER, P. A. (1980). *Probabilités et potentiel. Chapitres V à VIII*: *Théorie des Martingales.* Hermann, Paris. MR0566768

[5] DUQUESNE, T. and LE GALL, J. F. (2002). *Random Trees, Lévy Processes and Spatial Branching Processes. Astérisque* **281**. MR1954248

[6] DUQUESNE, T. and LE GALL, J. F. (2005). Probabilistic and fractal aspects of Lévy trees. *Probab. Theory Related Fields* **131** 553–603. MR2147221

[7] EVANS, S. N., PITMAN, J. W. and WINTER, A. (2006). Rayleigh processes, real trees and root growth with re-grafting. *Probab. Theory Related Fields* **34** 81–126. MR2221786

[8] EVANS, S. N. and WINTER, A. (2006). Subtree prune and re-graft: A reversible tree valued Markov process. *Ann. Probab.* **34** 918–961. MR2243874

[9] GRIMVALL, A. (1974). On the convergence of sequences of branching processes. *Ann. Probab.* **2** 1027–1045. MR0362529

[10] LAMPERTI, J. (1966). The limit of a sequence of branching processes. *Z. Wahrsch. Verw. Gebiete* **7** 271–288. MR0217893

[11] MIERMONT, G. (2003). Self-similar fragmentations derived from the stable tree. I Splitting at heights. *Probab. Theory Related Fields* **127** 423–454. MR2018924

[12] MIERMONT, G. (2005). Self-similar fragmentations derived from the stable tree. II Splitting at nodes. *Probab. Theory Related Fields* **131** 341–375. MR2123249

[13] NORRIS, J. (1997). *Markov Chains.* Cambridge Univ. Press. MR1600720

DMA-ENS
45 RUE D'ULM
75005 PARIS
FRANCE
E-MAIL: weill@dma.ens.fr