Hereditary Tree Growth and Lévy Forests

Thomas Duquesne

Université Pierre et Marie Curie (Paris 6)

Joint work with Matthias Winkel

We introduce the notion of hereditary property and we also consider reduction of trees by a given hereditary property.

We explain a general tightness criterion for trees.

We then consider families of Galton-Watson trees with exponential edge lengths that are consistent under hereditary reduction.

We show that they always converge to Lévy forests that have been introduced by Le Gall and Le Jan in 1998 and that extend the Continuum Random Tree of Aldous.

Definition

Real Trees: a metric space (T, d) is a real tree if any two points $\sigma, \sigma' \in T$ are connected by a unique injective path denoted by $[\![\sigma, \sigma']\!]$, which furthermore is isometric to the interval $[0, d(\sigma, \sigma')]$.

• $[\![\sigma, \sigma']\!]$ is the geodesic joining σ to σ'.

• Real trees generalize graph-trees. Informally they are obtained by glueing intervals (with their metric) without creating loops.

• We shall restrict to **compact real trees** with a distinguished point $\rho \in T$ called the *root*.

Definition

Let (T_1, d_1, ρ_1) and (T_2, d_2, ρ_2) be two compact rooted real trees. Their **pointed Gromov-Hausdorff distance** is given by

$$\delta(T_1, T_2) = \inf \{ d_{\mathrm{Haus}}(f_1(T_1), f_2(T_2)) \vee d(f_1(\rho_1), f_2(\rho_2)) \} .$$

Here the infimum is taken over all $(f_1, f_2, (E, d))$, where (E, d) is a metric space, where $f_i : T_i \to E$, $i = 1, 2$, are isometric embeddings and where d_{Haus} stands for the Hausdorff distance on the set of compact subsets of E.

• $\delta(T_1, T_2)$ only depends on the pointed isometry classes of T_1 and T_2.

• Denote $\mathbb{T}$ the set of pointed isometry classes of compact rooted real trees. Then, $(\mathbb{T}, \delta)$ is Polish (Evans-Pitman-Winter / Gromov).

- **Notation:** $\widetilde{T}$ stands for the pointed isometry class of (T, d, ρ).

- **The subtree above** $\sigma \in T$ is $\theta_\sigma T = \{\sigma' \in T : \sigma \in [\![\rho, \sigma']\!]\}$.

Note that $(\theta_\sigma T, d, \sigma)$ is also a compact rooted real tree. We denote by $\widetilde{\theta_\sigma T}$ its pointed isometry class in $\mathbb{T}$.

Definition

A Borel subset $A \subset \mathbb{T}$ is **hereditary** iff for any compact rooted real tree (T, d, ρ) and for any $\sigma \in T$,

$$\widetilde{\theta_\sigma T} \in A \Longrightarrow \widetilde{T} \in A .$$

- If σ enjoys property A, the ancestor (the root) also does.

Definition

Let $A \subset \mathbb{T}$ be hereditary. The A-**reduced subtree of** T is

$$R_A(T) = \text{Closure of } \{\rho\} \cup \{\sigma \in T : \widetilde{\theta}_\sigma T \in A\}.$$

• $(R_A(T), d, \rho)$ is a compact rooted real tree. Its pointed isometry class only depends on $\widetilde{T}$. So R_A can be viewed as a function from $\mathbb{T}$ to $\mathbb{T}$.

• The A-reduction $R_A : \mathbb{T} \to \mathbb{T}$ is Borel-measurable.

• **Composition:** let A, A' be two hereditary properties. We set

$$A' \circ A = \left\{ \widetilde{T} \in \mathbb{T} : R_A(\widetilde{T}) \in A' \right\} .$$

Then

$$R_{A'} \circ R_A = R_{A' \circ A} .$$

Important example: the length erasure. The height of (T, d, ρ) is

$$H(T) := \sup_{\sigma \in T} d(\rho, \sigma) .$$

For any $h \in [0, \infty)$,

$$A_h := \{ \widetilde{T} \in \mathbb{T} : H(\widetilde{T}) \geq h \} \text{ is hereditary.}$$

R_{A_h} is the h-**length erasure**. Note that $R_{A_h}(T)$ is a discrete tree with edge lengths.

$$A_{h'} \circ A_h = A_{h'+h} \quad \text{and} \quad R_{A_{h'}} \circ R_{A_h} = R_{A_{h'+h}} .$$

• Leaf erasure (or trimming) is a nice approximation tool (Kesten, Neveu, Le Jan, Evans-Pitman-Winter).

• The h-**erased profile of** T is given for any $a \in [0, \infty)$ by

$$Z_a^{(h)}(T) = \#\{\sigma \in T : H(\theta_\sigma T) \geq h \text{ and } d(\rho, \sigma) = a\} .$$

Namely it is the number of points in $R_{A_h}(T)$ at distance a from the root.

• Note that $Z_a^{(h)}(T) \in \mathbb{N}$ and $a \mapsto Z_a^{(h)}(T)$ is left-continuous with right limit.

• The h-erased profile allows to control $N_h(T)$, that is the minimal number of open balls necessary to cover T:

$$N_{2h}(T) \;\leq\; 1 + \sum_{1 \leq k \leq \frac{1}{h} H(T)} Z_{kh}^{(h)}(T) .$$

Proof: Let $\sigma \in T$ be such that $d(\rho, \sigma) \geq h$. Then there exists a unique $k \in \mathbb{N}$ such that

$$(k+1)h \leq d(\rho, \sigma) < (k+2)h \; .$$

Thus, there exists a unique $\sigma_0 \in T$ such that

$$\sigma_0 \in [\![\rho, \sigma]\!] \quad \text{and} \quad d(\rho, \sigma_0) = kh \; .$$

Now, observe that

$$H(\theta_{\sigma_0}(T)) \geq h \quad \text{and} \quad \sigma \in \mathrm{Ball}(\sigma_0, 2h) \; ,$$

which implies the inequality.

Growth. Let (T_1, d_1, ρ_1) and (T_2, d_2, ρ_2) be two compact rooted real trees. If there exists an isometry $f : T_1 \to T_2$ such that $f(\rho_1) = \rho_2$, we say that T_1 can be embedded in T_2.

It only depends on the pointed isometry classes of T_1 and T_2 and we denote it by

$$\widetilde{T}_1 \preceq \widetilde{T}_2 \; .$$

- $\preceq$ is a partial order on $\mathbb{T}$.
- Note that if $\widetilde{T}_1 \preceq \widetilde{T}_2$, then

$$Z_a^{(h)}(T_1) \leq Z_a^{(h)}(T_2) \quad \text{and} \quad H(T_1) \leq H(T_2) \; .$$

Theorem

Let $\widetilde{\mathcal{T}}_n$, $n \in \mathbb{N}$, be a sequence of random compact rooted real trees. Assume the following.

*(**a**) The laws of $H(\widetilde{\mathcal{T}}_n)$, $n \in \mathbb{N}$, are tight on $[0, \infty)$.*

*(**b**) For any fixed $a, h \in (0, \infty)$ the laws of $Z_a^{(h)}(\widetilde{\mathcal{T}}_n)$, $n \in \mathbb{N}$, are tight on $\mathbb{N}$.*

Then, the laws of $\widetilde{\mathcal{T}}_n$, $n \in \mathbb{N}$ are tight on $\mathbb{T}$. The converse is true.

- *In addition assume the following.*

*(**c**) For any $n \in \mathbb{N}$, a.s. $\widetilde{\mathcal{T}}_n \preceq \widetilde{\mathcal{T}}_{n+1}$.*

Then, there exists a random compact rooted real tree $\widetilde{\mathcal{T}}$, such that

$$a.s. \qquad \delta(\widetilde{\mathcal{T}}_n, \widetilde{\mathcal{T}}) \xrightarrow[n\to\infty]{} 0 \; .$$

Proof of the last point. By a projective argument, we can find a Polish space (E, d_E) such that

$$\mathcal{T}_n \subset \mathcal{T}_{n+1} \quad \text{and} \quad \rho = \rho_n \, , \quad n \in \mathbb{N} \, .$$

We only need to prove that $\bigcup_{n\in\mathbb{N}} \mathcal{T}_n$ is totally bounded. By (**a**) and (**b**), we get for any $a, h \in (0, \infty)$, a.s.

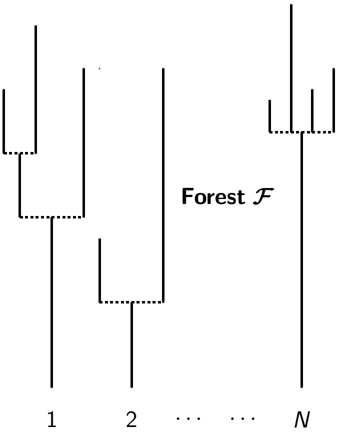
$$\sup_{n\in\mathbb{N}} H(\mathcal{T}_n) \; < \; \infty \quad \text{and} \quad \sup_{n\in\mathbb{N}} Z_a^{(h)}(\mathcal{T}_n) \; < \; \infty \, .$$

Now observe that

$$N_{2h}\Big(\bigcup_{n\in\mathbb{N}} \mathcal{T}_n \Big) \leq 1 + \sum_{1\leq k\leq \frac{1}{h}\sup_{n\in\mathbb{N}} H(\mathcal{T}_n)} \sup_{n\in\mathbb{N}} Z_{kh}^{(h)}(\mathcal{T}_n) \; < \; \infty \, .$$

Then take $\mathcal{T}$ as the closure of $\bigcup_{n\in\mathbb{N}} \mathcal{T}_n$ and note that $\mathcal{T}_n \to \mathcal{T}$ for the Hausdorff on the set of compact subsets of E.

Growth Processes



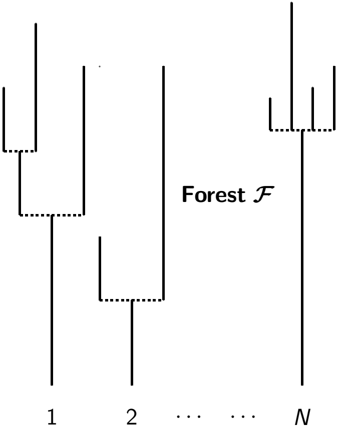
• Let $\xi = (\xi(k))_{k\geq 0}$ be such that

$$\sum_{k\geq 0} k\xi(k) \leq 1 \text{ and } \xi(0)+\xi(1) < 1 \ .$$

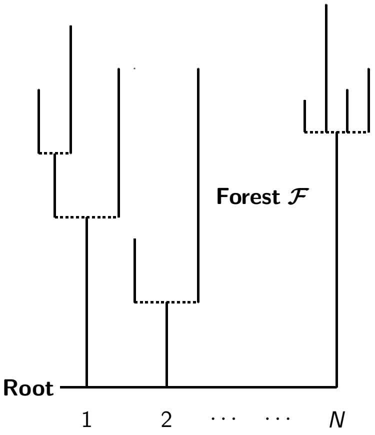
Notation: $\varphi_\xi(r) = \sum_{k\geq 0} \xi(k) r^k$.

• We consider a forest $\mathcal{F}$ of N independent Galton-Watson trees with *exponentially distributed edge lengths* with mean $\frac{1}{c}$ and *offspring distribution* ξ. Here,

$$N \overset{\text{(law)}}{=} \text{Poisson}(x)$$



- $\mathcal{F}$: N independent Galton-Watson trees with offspring distribution ξ and exponential edge lengths with mean $\frac{1}{c}$.
- We view such a forest as a *single tree* by merging all the ancestors to a single point that is called the root.



• $\mathcal{F}$: N independent Galton-Watson trees with offspring distribution ξ and exponential edge lengths with mean $\frac{1}{c}$.

• We view such a forest as a *single tree* by merging all the ancestors to a single point that is called the root.

• We view $\mathcal{F}$ as a random element of $\mathbb{T}$ and we say that it is a

$$\mathcal{F} \text{ is a } \mathrm{GW}(\xi, c, x)\text{-real forest} .$$

Theorem

Let $A \subset \mathbb{T}$ be hereditary. Assume that

$$\alpha := \mathbf{P}\big(\textit{1 single } GW(\xi, c)\textit{-real tree} \notin A\big) \in (0,1)\,.$$

Then, $R_A(\mathcal{F})$ is a $GW(\xi^{(\alpha)}, c^{(\alpha)}, x^{(\alpha)})$-real forest, where

$$x^{(\alpha)} = (1-\alpha)x\,, \quad c^{(\alpha)} = \big(1 - \varphi'_\xi(\alpha)\big)c$$

and

$$\varphi_{\xi^{(\alpha)}}(r) = r + \frac{\varphi_\xi(\alpha + (1-\alpha)r) - \alpha - (1-\alpha)r}{(1-\alpha)(1-\varphi'_\xi(\alpha))}\,, \quad r \in [0,1]\,.$$

If $\xi^{(\alpha)} = \xi$ for any $\alpha \in (0,1)$, then there exists $\gamma \in (1,2]$ such that

$$\varphi_\xi(r) = r + \frac{1}{\gamma}(1-r)^\gamma\,.$$

GW-laws are stable under hereditary reduction. It is natural to consider families of GW-tree that consistent under hereditary reduction.

Definition

Growth process: $\mathcal{F}_\lambda$, $\lambda \in [0, \infty)$ such that

(**a**) $\mathcal{F}_\lambda$ is a GW($\xi_\lambda, c_\lambda, x_\lambda$)-real forest.

(**b**) For any $\lambda < \lambda'$, there exists $A_{\lambda,\lambda'} \subset \mathbb{T}$ hereditary and such that a.s.

$$R_{A_{\lambda,\lambda'}}(\mathcal{F}_{\lambda'}) = \mathcal{F}_\lambda \ .$$

Examples: leaf-erasure, mass-erasure, site and leaf percolation ...

Theorem

Assume that $c_\lambda \longrightarrow \infty$ *as* $\lambda \to \infty$. *Then, up to reparametrisation, there exists* $\Psi : [0, \infty) \to [0, \infty)$ *such that*

$$x_\lambda = x\lambda, \quad c_\lambda = \Psi'(\lambda), \quad \varphi_{\xi_\lambda}(r) = r + \frac{\Psi(\lambda(1-r))}{\lambda \Psi'(\lambda)}$$

and Ψ *has the following Lévy-Khintchine form*

$$\Psi(\lambda) = a\lambda + b\lambda^2 + \int_{(0,\infty)} \left(e^{-\lambda r} - 1 + \lambda r \right) \pi(dr) \,,$$

with $a, b \in [0, \infty)$ *and* $\int_{(0,\infty)} \min(r, r^2)\, \pi(dr) < \infty$.

Explanation: Since $R_{A_{\lambda_0,\lambda}}(\mathcal{F}_\lambda) = \mathcal{F}_{\lambda_0}$, there is $\alpha_\lambda \in (0,1)$, such that

$$\varphi_{\xi_{\lambda_0}}(r) = r + \frac{\varphi_{\xi_\lambda}(\alpha_\lambda + (1-\alpha_\lambda)r) - \alpha_\lambda - (1-\alpha_\lambda)r}{(1-\alpha_\lambda)(1-\varphi'_{\xi_\lambda}(\alpha_\lambda))}, \quad r \in [0,1] .$$

Now $c_\lambda \longrightarrow \infty$ implies that $\alpha_\lambda \longrightarrow 1$. Then, $\varphi_{\xi_{\lambda_0}}$ can be analytically extended on $(-\infty, 1)$. We set

$$\Psi(r) = \varphi_{\xi_{\lambda_0}}(1-r) - 1 + r , \quad r \in [0,\infty)$$

and note that for any $k \geq 2$,

$$(-1)^k \Psi^{(k)}(r) \geq 0, \quad r \in (0,\infty)$$

and we apply Bernstein to $\Psi^{(2)}$.

Theorem

Assume that $c_\lambda \longrightarrow \infty$ *as* $\lambda \to \infty$ *and that* $\int^\infty \frac{d\lambda}{\Psi(\lambda)} < \infty$. *Then*

$$a.s. \quad \delta(\mathcal{F}_\lambda, \mathcal{F}) \xrightarrow[\lambda\to\infty]{} 0$$

where $\mathcal{F}$ *is a* (Ψ, x)*-Lévy forest.*

Moreover, for any $a \in [0, \infty)$, *a.s.*

$$\frac{1}{\lambda}\#\{\sigma \in \mathcal{F}_\lambda : d(\rho, \sigma) = a\} \xrightarrow[\lambda\to\infty]{} Z_a$$

where $(Z_a)_{a\in[0,\infty)}$ *is a Continuous States-space Branching Process with branching mechanism* Ψ *and initial value* $Z_0 = x$.

Proof: The law of $H(\mathcal{F}_\lambda)$ is known explicitly in terms of λ and Ψ (elementary computation). We check that

$$H(\mathcal{F}_\lambda) \xrightarrow[\lambda \to \infty]{} \text{r.v. in } [0, \infty)$$

Then, $R_{A_h}(\boldsymbol{\mathcal{F}}_\lambda)$ is a GW$(\xi_{h,\lambda}, c_{h,\lambda}, x_{h,\lambda})$-real forest, where the parameters $\xi_{h,\lambda}, c_{h,\lambda}, x_{h,\lambda}$ are explicitly computed in terms of Ψ, λ and h (elementary computation).

Thus $a \mapsto Z_a^{(h)}(\boldsymbol{\mathcal{F}}_\lambda)$ is a Markov Branching process whose law is explicitely known and we check that

$$Z_a^{(h)}(\boldsymbol{\mathcal{F}}_\lambda) \xrightarrow[\lambda \to \infty]{} \text{r.v. in } \mathbb{N}.$$

Then, the convergence theorem applies.

• The assumption $\int^{\infty} \frac{d\lambda}{\Psi(\lambda)} < \infty$ is a necessary and sufficient condition for convergence. It corresponds to the tightness of the height $H(\mathcal{F}_\lambda)$.

This condition holds in the stable cases where $\Psi(\lambda) = \lambda^\gamma$, $\gamma \in (1, 2]$.

• The case where $\gamma = 2$ corresponds to the Brownian forest that is coded by a reflected Brownian motion and that, roughly speaking, corresponds to Aldous's CRT. The contour function of the Brownian forest is coded by a a reflected Brownian motion stopped when its local time at 0 reaches x.

• All the results extend to supercritical cases.

DZIĘKUJE !