UNIVERSITÉ PARIS-EST

THÈSE

présentée pour obtenir

LE GRADE DE DOCTEUR EN SCIENCES DE L'UNIVERSITÉ PARIS-EST

Spécialité : Mathématiques

par

Patrick HOSCHEIT

Processus à valeurs dans les arbres aléatoires continus

Soutenue le 10 décembre 2012 devant un jury composé de :

M. Romain ABRAHAM	directeur de thèse
M. Jean-François DELMAS	directeur de thèse
M. Grégory MIERMONT	rapporteur
M. Anton WAKOLBINGER	rapporteur
M. Philippe CHASSAING	
M. Amaury LAMBERT	

In the broad light of day mathematicians check their equations and their proofs, leaving no stone unturned in their search for rigour. But, at night, under the full moon, they dream, they float among the stars and wonder at the miracle of the heavens. They are inspired. Without dreams there is no art, no mathematics, no life.

—Sir Michael Atiyah

Remerciements

Résumé

Dans cette thèse, nous présentons le fruit de trois années de recherche autour du thème des arbres aléatoires continus et de différentes façons de les découper. Ce travail a donné lieu à trois articles soumis pour publication à des revues scientifiques à comité de lecture :

- [ADH12a] Romain Abraham, Jean-François Delmas and Patrick Hoscheit, *Exit times for an increasing Lévy tree-valued process*, prépublication arXiv :1202.5463 (2012), 1-33
- [ADH12b] Romain Abraham, Jean-François Delmas and Patrick Hoscheit, *A note on Gromov-Hausdorff-Prokhorov distance between (locally) compact measure spaces*, prépublication arXiv :1202.5464 (2012), 1-19
- [Hos12] Patrick Hoscheit, *Fluctuations for the number of records on random binary trees* (2012), *soumis*

Ces trois articles constituent, après une longue introduction au sujet, les trois chapitres de cette thèse. Dans l'appendice, un résultat technique est présenté, qui étend légèrement un résultat de [ADV10]. La preuve ne présente que peu de différences par rapport à la version publiée, mais nous l'incluons par souci d'exhaustivité.

Topologie des arbres réels mesurés

Les arbres aléatoires continus sont des espaces métriques aléatoires appartenant à la classe des *arbres réels* (ou $\mathbf{R}$-arbres), c'est-à-dire des espaces métriques (T, d) tels que

(i) Pour tous $s, t \in T$, il existe une unique isométrie $f_{s,t}$ de $[0, d(s,t)]$ vers T telle que $f_{s,t}(0) = s$ et $f_{s,t}(d(s,t)) = t$.

(ii) Pour tous $s, t \in T$, si q est une injection continue de $[0,1]$ dans T telle que $q(0) = s$ et $q(1) = t$, alors $q([0,1]) = f_{s,t}([0, d(s,t)])$.

Les arbres réels sont naturellement munis d'une mesure borélienne, dite mesure de *longueur* ℓ, qui représente en quelque sorte la mesure de Lebesgue sur T. De plus, certains arbres, notamment les arbres de Lévy définis plus bas, possèdent une mesure $\mathbf{m}$, dite de *masse* qui représente une mesure « uniforme » sur les feuilles de l'arbre.

Pour pouvoir définir et étudier des variables aléatoires à valeurs dans les arbres réels, la question de la topologie des espaces d'arbres se pose naturellement. La topologie la plus usuelle,

sur l'espace $\mathbb{T}_{\mathrm{cpct}}$ des classes d'isométrie des arbres réels *compacts*, est la topologie de Gromov-Hausdorff, proposée par M. Gromov ([Gro07]) dans le contexte de la théorie géométrique des groupes. Il s'agit d'une topologie métrisable dont l'idée maîtresse est de comparer deux arbres réels compacts grâce à leur distance de Hausdorff, quitte à les plonger tous deux dans un espace métrique commun. Cette topologie a trouvé de nombreuses applications, notamment pour l'étude de limites d'échelle de grands objets aléatoires.

Dans ce travail, nous nous sommes intéressés particulièrement à l'étude de topologies sur les espaces métriques mesurés. En effet, pour étudier certaines propriétés des arbres, il peut être intéressant de disposer de mesures permettant de choisir des points au hasard dans l'arbre. Par exemple, si T_n est un arbre choisi uniformément parmi tous les arbres binaires enracinés à n noeuds, quelle est la distance moyenne entre la racine et une feuille ? Autant le choix d'une feuille uniforme dans un arbre discret ne pose pas de problème, autant il n'en va pas de même pour les arbres aléatoires continus. Néanmoins, pour une large classe d'entre eux, il est possible de construire des mesures portées par l'arbre ; il est donc pertinent de chercher des topologies sur les espaces d'arbres qui prennent en compte cette dimension.

La littérature existante sur le sujet est déjà assez vaste, grâce à des liens avec la théorie du transport optimal ([Vil09]) qui ont conduit différents auteurs à s'y intéresser ([Fuk87, Stu06a]). Différentes approches existent ; certaines négligent les aspects géométriques des arbres au profit des mesures qu'ils portent, telle la topologie de Gromov-faible ([Gro07, GPW08]), d'autres combinent les deux, telle la topologie de Gromov-Hausdorff mesurée ([Eva08, Mie09]). Toutefois, aucune de ces topologies ne permet de prendre en compte les arbres aléatoires de Lévy dans leur généralité : en effet, dans le cas surcritique, les arbres de Lévy ne sont pas compacts avec probabilité positive, et la mesure qu'ils portent n'est que σ-finie.

Dans l'article [ADH12b], nous avons donc développé une topologie sur la classe $\mathbb{T}$ des **R**-arbres enracinés, localement compacts et complets, munis de mesures boréliennes localement finies (finies sur les parties bornées). L'idée est de partir d'une distance de type Gromov-Hausdorff-Prokhorov sur la classe des arbres réels compacts, munis de mesures finies, et d'étendre cette distance à $\mathbb{T}$ par localisation. Plus précisément, soit

$$d^{\mathrm{c}}_{\mathrm{GHP}}(T_1,T_2)=\inf_{(Z,\Phi_1,\Phi_2)}\left(d^Z(\Phi_1(\emptyset_1),\Phi_2(\emptyset_2))+d^Z_H(\Phi_1(\mathcal{T}_1),\Phi_2(\mathcal{T}_2))+d^Z_{\mathrm{Pr}}((\Phi_1)_*\mu_1,(\Phi_2)_*\mu_2)\right),$$

la distance de Gromov-Hausdorff-Prokhorov *compacte* entre deux tels arbres (enracinés et mesurés) $(\mathcal{T}_1,\emptyset_1,\mu_1)$ et $(\mathcal{T}_2,\emptyset_2,\mu_2)$. Dans cette expression, d_{H} désigne la distance de Hausdorff entre parties fermées d'un espace polonais Z, d_{Pr} désigne la distance de Prokhorov entre mesures finies sur Z, et la borne inférieure porte sur tous les plongements isométriques Φ_1,Φ_2 de T_1,T_2 dans un espace métrique polonais (Z,d^Z). La distance d^c_{GHP} métrise en un sens la topologie de la convergence faible de mesures finies sur des espaces compacts. La distance de Gromov-Hausdorff-Prokhorov sur $\mathbb{T}$ est alors définie par

$$d_{\mathrm{GHP}}(\mathcal{T}_1,\mathcal{T}_2)=\int_0^\infty \mathrm{e}^{-r}\left(1\wedge d^{\mathrm{c}}_{\mathrm{GHP}}(B_{\mathcal{T}_1}(\emptyset_1,r),B_{\mathcal{T}_2}(\emptyset_2,r))\right)\,dr.$$

On vérifie que cette expression définit bien une distance sur $\mathbb{T}$, et que la topologie ainsi définie est une topologie d'espace métrique polonais, ce qui constitue un préalable indispensable à l'étude de processus stochastiques sur $\mathbb{T}$. La preuve de ce dernier résultat nécessite l'utilisation d'un critère de précompacité dans l'espace $(\mathbb{T}, d_{\mathrm{GHP}})$ dont la démonstration, relativement technique, constitue le coeur du chapitre 2.

Arbres aléatoires continus

Les arbres aléatoires que nous étudions appartiennent pour l'essentiel à la classe des *arbres de Lévy*, qui sont des modèles d'arbres découverts par Le Gall et Le Jan ([LL98a, LL98b]). La motivation initiale est de donner un sens à la généalogie des processus de branchement à espace d'états continu (CSBP). Ceux-ci sont des processus de Markov $(Z_t, t \geq 0)$ à valeurs dans $\mathbf{R}_+$, représentant l'évolution de la taille d'une grande population, dont les individus sont infinitésimaux. Les CSBP sont caractérisés par une fonction ψ, leur *mécanisme de branchement*, qui est telle que

$$\psi(u) = \alpha u + \beta u^2 + \int_{(0,\infty)} \left(\mathrm{e}^{-ux} - 1 + ux\mathbf{1}_{\{x<1\}}\right) \, \Pi(dx), \tag{1}$$

où $\alpha \in \mathbf{R}$, $\beta \geq 0$ et où Π est une mesure σ-finie sur $\mathbf{R}_+$, telle que $\int (1 \wedge x^2)\Pi(dx) < \infty$. Les trajectoires des CSBP peuvent être mises en relation, via la *transformation de Lamperti*, avec les trajectoires d'un processus de Lévy spectralement positif. On peut donc identifier une composante brownienne et une composante de sauts dans un CSBP général.

Il en va de même pour les arbres de Lévy : leurs points de branchements sont soit de degré 3 (branchement *binaire*), correspondant à la partie « brownienne » du processus de branchement, soit de degré infini, correspondant aux sauts du processus de branchement. Une question fondamentale pour l'étude des arbres de Lévy est la question de la criticalité : la population qu'ils modélisent s'éteint-elle presque sûrement ou bien peut-elle survivre en temps infini ? Sous des conditions raisonnables (*condition de Grey*), le critère est simple : si $\psi'(0+) = 0$ (cas *critique*) ou si $\psi'(0+) > 0$ (cas *sous-critique*), le processus s'éteint presque sûrement (avec probabilité 1, il existe $t_0 \geq 0$ tel que $Z_t = 0$ pour tout $t \geq t_0$). Par contre, si $-\infty \leq \psi'(0+) < 0$, (cas *sur-critique*), avec probabilité positive, le processus ne s'éteint jamais. L'arbre de Lévy correspondant sera donc compact dans les cas sous-critique et critique, alors qu'avec probabilité positive, dans le cas sur-critique, il n'est pas compact, seulement localement compact.

Divers travaux se sont intéressés à la question de la *fragmentation* des arbres de Lévy (ou de leurs dérivés, tels l'arbre brownien d'Aldous ou l'arbre α-stable), notamment Aldous-Pitman ([AP98a]) et Miermont ([Mie03, Mie04]). La procédure la plus générale a été décrite par Abraham, Delmas et Voisin ([ADV10]) et consiste à combiner une fragmentation homogène sur le squelette de l'arbre, similaire à la fragmentation d'Aldous et Pitman, avec une fragmentation biaisée par la taille des noeuds infinis. Sous certaines hypothèses, il est alors possible de montrer que l'arbre ainsi élagué est encore un arbre de Lévy, avec un mécanisme de branchement décrit explicitement.

Un cas particulier de l'élagage des arbres de Lévy vérifie une propriété de consistance : si ψ est un mécanisme de branchement, on définit, pour $\theta > 0$, le mécanisme de branchement

$$\psi_\theta(u) = \psi(u+\theta) - \psi(\theta).$$

En utilisant l'élagage d'Abraham-Delmas-Voisin sur un arbre de Lévy de mécanisme ψ, on peut obtenir un arbre de Lévy, de mécanisme de branchement ψ_θ. De plus, la famille $(\psi_\theta, \theta > 0)$ est consistante, au sens où

$$\psi_{\theta+\theta'} = (\psi_\theta)_{\theta'} = (\psi_{\theta'})_\theta,$$

ce qui permet, *via* la procédure d'élagage, de construire un processus $(\mathcal{T}_\theta, \theta \geq 0)$, à valeurs dans $\mathbb{T}$ et indexé par θ de telle façon que pour tout $\theta \geq 0$, l'arbre $\mathcal{T}_\theta$ soit un arbre de Lévy de mécanisme de branchement ψ_θ. L'étude de ce processus, commencée dans [AD12a], se heurte à l'absence d'une bonne description de ses trajectoires, en particulier de ses transitions infinitésimales.

C'est dans ce contexte que nous avons introduit dans [ADH12c] une construction alternative du processus d'élagage utilisant des mesures ponctuelles de Poisson à valeurs dans $\mathbb{T}$. Cette construction permet de montrer que les transitions du processus d'élagage sont telles qu'à l'instant $\theta > 0$, conditionnellement à $\mathcal{T}_\theta$, l'arbre $\mathcal{T}_{\theta-}$ est obtenu en greffant un nouvel arbre de Lévy, de mécanisme de branchement ψ_θ, sur une feuille de $\mathcal{T}_\theta$ uniformément choisie. Formellement, on peut écrire le générateur infinitésimal du processus de croissance obtenu en retournant le temps, $(\mathcal{T}_{-\theta},\ \theta \in (-\infty, 0])$, de la façon suivante : si $\theta < 0$, si F est mesurable bornée sur $\mathbb{T}$ et si $\mathcal{T}$ est un élément de $\mathbb{T}$, alors

$$(\mathcal{L}_\theta F)(\mathcal{T}) = \int_{\mathcal{T}} \mathbf{m}^{\mathcal{T}}(ds) \int_{\mathbb{T}} \mathbf{N}^{\psi_\theta}[dT] \big(F(\mathcal{T} \circledast (T, s)) - F(\mathcal{T})\big).$$

Dans cette expression, $\mathbf{m}^{\mathcal{T}}$ désigne la mesure de masse de l'arbre $\mathcal{T}$ (mesure « uniforme » sur les feuilles), la notation $\mathcal{T} \circledast (T, s)$ désigne l'arbre obtenu en greffant T sur la feuille $s \in \mathcal{T}$, et la mesure $\mathbf{N}^{\psi_\theta}[dT]$ correspond à la loi de l'arbre greffé à l'instant θ (voir l'équation (3.40) pour une expression explicite). Ce théorème est prouvé dans le chapitre 3 en appliquant de façon itérative la propriété de Markov « spéciale » démontrée dans [AD12a], ce qui révèle au passage une intéressante décomposition des arbres de Lévy en générations.

Nous appliquons ensuite cette description trajectorielle à l'étude du comportement du processus d'élagage à certains temps d'arrêt, qui sont les instants où le processus franchit une hauteur donnée :

$$A_h = \sup\{\theta,\ H_{\max}(\mathcal{T}_\theta) > h\}, \quad 0 < h \leq \infty.$$

Nous décrivons précisément la loi du couple $(\mathcal{T}_{A_h-}, \mathcal{T}_{A_h})$ au moyen d'une décomposition spinale, par rapport à l'épine dorsale $[\![\emptyset, \mathrm{x}]\!]$, où x est la feuille sur laquelle vient se greffer l'arbre responsable du franchissement de la hauteur h. Cette décomposition est à rapprocher des classiques décompositions de Bismut et de Williams.

Coupures et processus de records

En 1970, Meir et Moon s'intéressent au problème suivant : étant donné un arbre T_n discret, enraciné, à n arêtes, on choisit uniformément l'une des arêtes, que l'on retire. On réitère ensuite la procédure sur la composante connexe contenant la racine et on note $X(T_n)$ le nombre (aléatoire) de coupures ainsi effectuées jusqu'à ce que la racine soit isolée. Lorsque T_n est un arbre aléatoire, le comportement de $X(T_n)$ est connu dans certains cas. Notamment, lorsque T_n est un arbre de Galton-Watson critique, de variance σ^2 finie, conditionné à avoir n arêtes, Panholzer ([Pan06]) et Janson ([Jan06]) ont montré

$$\lim_{n\to\infty} \frac{X(T_n)}{\sigma\sqrt{n}} = \mathscr{R},$$

où $\mathscr{R}$ est une variable aléatoire de loi Rayleigh, c'est-à-dire, distribuée suivant la densité $x\exp(-x^2/2)\mathbf{1}_{[0,\infty)}(x)$. Cette loi est connue pour être la loi de la hauteur d'une feuille choisie uniformément dans l'arbre brownien d'Aldous, qui est la limite d'échelle des arbres T_n convenablement renormalisés. Plusieurs travaux ont été menés pour relier cette variable aléatoire à un processus de fragmentation sur l'arbre continu.

La manière la plus naturelle de fragmenter l'arbre brownien d'Aldous a été découverte par Aldous et Pitman ([AP98a]) et consiste à fragmenter le squelette de l'arbre suivant un processus de Poisson à valeurs dans l'arbre. Il est démontré dans l'article original que cette fragmentation est un exemple de fragmentation *auto-similaire*, d'indice 1/2, de mesure de dislocation binaire, et sans érosion.

Cette fragmentation a été utilisée notamment par Bertoin et Miermont ([BM12]) pour construire un arbre cut($\mathscr{T}$) qui code la généalogie de la fragmentation d'Aldous-Pitman, et qui est en même temps limite d'échelle d'arbres qui codent pour la fragmentation discrète des arbres T_n. Le temps nécessaire à isoler la racine correspond alors à la hauteur d'une feuille uniforme dans ces arbres. Comme cut($\mathscr{T}$) est encore distribué comme un arbre brownien, on retrouve la limite d'échelle de Janson (et des résultats plus généraux sur le temps nécessaire à la séparation de k points de l'arbre).

Une approche différente a été choisie par Abraham et Delmas ([AD11]), qui examinent l'effet de la fragmentation d'Aldous-Pitman sur des sous-arbres de l'arbre brownien $\mathscr{T}$, obtenus en sélectionnant des feuilles uniformes. Notons T_n^* le sous-arbre de $\mathscr{T}$ engendré par n feuilles uniformes et la racine, privé de l'arête adjacente à la racine. La loi de T_n^* est connue depuis Aldous ([Ald91a]) : il s'agit d'un arbre binaire enraciné, à n feuilles et à $2n-2$ arêtes.

Abraham et Delmas considèrent alors le processus de *séparation* (ou de records) défini de la façon suivante : pour tout $s\in\mathscr{T}$, on pose $\theta(s)$ l'instant auquel le point s est séparé de la racine dans la fragmentation. Le processus $(\theta(s), s\in\mathscr{T})$ est alors un processus de saut pur défini sur l'arbre, qui croît vers $+\infty$ à mesure que l'on se rapproche de la racine.

L'analogue du nombre de coupures $X(T_n)$ est alors le nombre de sauts du processus θ sur T_n^*. Abraham et Delmas démontrent que

$$\lim_{n\to\infty} \frac{X(\mathrm{T}_n^*)}{\sqrt{2n}} = \int_{\mathscr{T}} \theta(s)\mathbf{m}(ds).$$

De plus, la quantité $\Theta = \int_{\mathscr{T}} \theta(s)\mathbf{m}(ds)$ est bien de loi Rayleigh, retrouvant ainsi le résultat de Janson.

Dans le chapitre 4, nous présentons une étude des fluctuations de $X(\mathrm{T}_n^*)/\sqrt{2n}$ autour de sa limite, sous forme d'un théorème central limite :

$$\lim_{n\to\infty} n^{1/4}\left(\frac{X_n^*}{\sqrt{2n}} - \Theta\right) = Z,$$

en loi, où Z admet pour fonction caractéristique $\mathbb{E}_\infty^{(1)}\left[\exp(itZ)\right] = \mathbb{E}_\infty^{(1)}\left[\exp(-t^2\Theta/\sqrt{2})\right]$. La preuve de ce théorème repose essentiellement sur l'utilisation de martingales, qui sont naturellement définies à partir des processus ponctuels de Poisson qui interviennent dans la définition de la fragmentation d'Aldous-Pitman. En particulier, si $X(T)$ est le nombre de sauts[1] de θ sur un sous-arbre $T \subset \mathscr{T}$, on peut montrer par des arguments de martingales que

$$\mathbb{E}\left[X(T)\right] = \mathbb{E}\left[\int_T \theta(s)\ \ell(ds)\right]$$
$$\mathbb{E}\left[\left(X(T) - \int_T \theta(s)\ \ell(ds)\right)^2\right] = \mathbb{E}\left[\int_T \theta(s)\ \ell(ds)\right].$$

Ceci permet de relier le nombre de sauts de θ sur T_n^* à la valeur moyenne de θ sur T_n^*. Heuristiquement, tout repose alors sur la convergence faible des mesures uniformes $\ell(ds)/\ell(\mathrm{T}_n)\mathbf{1}_{\mathrm{T}_n}(s)$ vers la mesure de masse $\mathbf{m}(ds)$ de l'arbre (Aldous). Nous présentons dans le chapitre 4 des méthodes permettant de contrôler précisément ces convergences pour en examiner les fluctuations.

[1]On peut avoir $X(T) = \infty$, en particulier si T contient la racine

Table of contents

CHAPTER 1

Introduction

In this thesis, we present several contributions of the author to the theory of continuum random trees. The following papers, each forming a chapter of this thesis, have been submitted to peer-reviewed journals. In the appendix, a technical result proven by the author is added for the sake of completeness, slightly generalizing a result appearing in [ADV10].

- [ADH12a] Romain Abraham, Jean-François Delmas and Patrick Hoscheit, *Exit times for an increasing Lévy tree-valued process*, Arxiv preprint arXiv:1202.5463 (2012), 1-33
- [ADH12b] Romain Abraham, Jean-François Delmas and Patrick Hoscheit, *A note on Gromov-Hausdorff-Prokhorov distance between (locally) compact measure spaces*, Arxiv preprint arXiv:1202.5464 (2012), 1-19
- [Hos12] Patrick Hoscheit, *Fluctuations for the number of records on random binary trees* (2012)

Motivation

This short introduction aims to paint a broad picture of the field, which has seen an enormous development in the last ten years, by giving some historical perspective and by introducing some of the research topics related to continuum random tree theory. We will usually not try to give precise definitions of the terms used, referring to the appropriate papers for more details.

The first formal definition of a continuum random tree goes back to the three seminal papers by Aldous ([Ald91a, Ald91b, Ald93]), as well as the work by Le Gall ([Le 91]) on trees in the representation of measure-valued processes. Aldous offers several equivalent definitions of what he calls the *compact continuum tree* (later the *Brownian tree*), as well as several convergence results. Among these, the following is of particular interest since it led to important generalizations. We state it as in [Ald93].

Theorem 1 (Aldous, 1993). *Let T_n be a conditioned Galton-Watson tree whose offspring distribution ξ satisfies* $\mathbb{E}[\xi] = 1$ *(ξ is* critical*)*, $0 < \mathrm{Var}(\xi) = \sigma^2 < \infty$ *and* $\mathrm{GCD}\big(\{j \in \mathbf{N},\ \mathbb{P}(\xi = j) > 0\}\big) = 1$. *Rescale* *the edges of T_n to have length $\sigma n^{-1/2}$. Let $f_n : [\![1, 2n-1]\!] \to \mathbf{R}_+$ be the search-depth process for T_n. Define $\bar{f}_n = [0,1] \to \mathbf{R}_+$ by*

$$\bar{f}_n(i/2n) = f_n(i)\ ;\ 1 \le i \le 2n-1\ ;\ \bar{f}_n(0) = \bar{f}_n(1) = 0,$$

with linear interpolation between these values. Then, $(\bar{f}_n(t),\ 0 \le t \le 1) \to (2B_t,\ 0 \le t \le 1)$ in distribution on $C([0,1])$, where B is a standard normalized Brownian excursion.

The class of conditioned, finite-variance Galton-Watson trees is known to contain several classical models of trees. For instance, uniform rooted plane binary trees are conditioned Galton-Watson tree with $\xi(0) = \xi(2) = 1/2$; uniform rooted labelled trees (Cayley trees) correspond to the case $\xi(k) = \mathrm{e}^1/k!,\ k \ge 0$ (for a complete account of the theory, see [Jan12]). It is remarkable that the limiting process (or the tree encoded by it) depends on ξ only through its variance σ^2. This kind of universal behavior is strongly reminiscent of the universality of Brownian motion as a scaling limit for random walks (and indeed, Aldous's result relies crucially on Donsker's invariance theorem).

This technique of encoding trees by continuous excursions proved very convenient and quite suitable for generalizations. Most importantly, in 1998, Le Gall and Le Jan ([LL98b]) defined the general *height process* which encodes the genealogy of conservative continuous-state branching processes by an excursion $H : [0,\sigma] \to \mathbf{R}_+$. This in turn led to the definition of *Lévy trees* by Duquesne and Le Gall ([DL02, DL05]). Lévy trees are a natural generalization of Aldous's Brownian tree, being the scaling limits of *unconditioned* Galton-Watson trees. They are natural genealogical trees for continuous-state branching processes (CSBP), and, through the snake construction by Le Gall ([Le 91, Le 93a]), they constitute a powerful tool for the study of measure-valued branching processes. Indeed, a superprocess with general branching mechanism ψ can always be represented as a cloud of Lévy snakes. This is very useful when investigating the connections between superprocesses and nonlinear partial differential equations of the form

$$\frac{1}{2}\Delta u = \psi(u),$$

where ψ is the Laplace exponent of a spectrally positive Lévy process, see for instance [Le 94, Le 95] for the case $\psi(u) = u^2$. We refer to the seminal papers by Dynkin ([Dyn91, Dyn93]) for more information on the link between superprocesses and partial differential equations.

Another motivation for studying continuum random trees is combinatorial. Indeed, many asymptotic properties of combinatorial trees can be readily explained by features of the limiting continuum random tree. For instance, using powerful singularity analysis methods, Flajolet and Odlyzko ([FO80]) proved in 1982 that the height $H(T_n)$ of a uniform binary tree with n nodes has asymptotic distribution given by

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{H(T_n)}{2\sqrt{n}} \in dx\right) = 4x \sum_{k\ge1} k^2 \left(2k^2x^2 - 3\right) \mathrm{e}^{-k^2x^2}\ dx,$$

known as the *Theta distribution.* Using Aldous's theorem about the limiting distribution of T_n, we can easily show that in fact

$$\lim_{n\to\infty} \frac{H(T_n)}{2\sqrt{n}} = \max_{0\le t\le1} B_t,$$

where B_t is a normalized Brownian excursion. Thus, the Theta distribution can be very naturally accounted for as the maximum of a normalized Brownian excursion, which is interpreted as the maximal height of a leaf in the Brownian CRT. Many more applications of this kind are collected in [Drm09].

The interest for continuum random trees in combinatorics is not restricted to asymptotic random tree theory alone: in recent years, a considerable literature has emerged on the topic of *random maps*. These are random graph structures embedded in the two-dimensional sphere (or higher-genera manifolds, see [Bet10]). Through the use of clever bijections (such as the Cori-Vauquelin-Schaeffer bijection, see [CS04], or the Bouttier-di Francesco-Guitter bijection, see [BdFG04]), the theory of planar maps is strongly connected with the theory of random trees. Therefore, the asymptotic structure of planar maps is usually studied through the Brownian CRT. A recent breakthrough was recently made independently by Le Gall ([Le 12]) and Miermont ([Mie12]) who proved that there exists a universal scaling limit for a large class of random maps, named the *Brownian map*, which is related to Aldous's CRT. A good survey of the theory can be found in [LM12].

The asymptotic theory of random trees contains other scaling limits than the Brownian CRT. Indeed, just as the Brownian CRT can be seen as a conditioned version of some specific Lévy tree (associated with the branching mechanism $\psi(u) = u^2/2$), there are non-Brownian continuum random trees that arise as scaling limits of natural tree models. For instance, the *stable tree*, with stability parameter $\alpha \in (1,2)$ is also a conditioned Lévy tree (with branching mechanism $\psi(u) = u^\alpha$), and it was proven by Duquesne ([Duq03]) that it corresponds to the scaling limit of conditioned Galton-Watson trees with a critical offspring distribution having infinite variance, lying in the attraction domain of an α-stable distribution. Other conditionings that lead to the stable CRT in the limit were considered in [Kor12a].

As these few examples show, there is a strong interest in continuum random trees, since they arise naturally as scaling limits of many discrete models, whether it be branching processes, random tree models or random maps. The contour process approach pioneered by Aldous was completed by a more geometrical approach, starting with [EPW05]. The main idea is to consider random trees (discrete and continuum) as metric spaces, endowed with several measures. This enables among other things the detailed analysis of their geometrical features. For instance, the fine topological properties of Lévy trees were extensively studied by Duquesne ([DL05, DL06, Duq10, Duq12]).

In this framework, scaling limits of random discrete trees can therefore be seen as convergence results in the appropriate space of metric spaces. This approach also enables to define and study *tree-valued processes*, which will be our main interest. In the next sections, we will present the general theory (in its present state, which is far from complete) of continuum random trees seen as metric spaces, theory which will be used to define the Lévy tree-valued pruning process of Abraham and Delmas ([AD12a]), as well as several other tree-valued processes. Finally, we will give a summary of the fragmentation theory of self-similar CRTs and its relation to certain cutting-edge edge-cutting procedures on discrete trees.

Although most—if not all—of the results presented here can be found elsewhere, we hope that this synthetic presentation of the theory might prove useful.

1.1 Topologies on tree spaces

This section will present a panorama of the different topologies that are used in continuum random tree theory. We refer to the excellent books [BBI01] (Chapters 7 and 8) and [Vil09] (Chapter 27) for pedagogical introductions to the subject.

The Gromov-Hausdorff topology

Although we will eventually turn to tree-like spaces, the basic theory of Gromov-Hausdorff convergence requires only a metric structure. We want to define what it means for two metric spaces (X,d) and (X',d') to be "close". If A and B are compact subspaces of some common Polish[1] metric space Z, then a good way to compare A and B is given by the Hausdorff distance:

$$d_{\mathrm{H}}^{Z}(A,B) = \inf\left\{\varepsilon > 0;\ A \subset B^{\varepsilon} \text{ and } B \subset A^{\varepsilon}\right\},$$

where $A^{\varepsilon} = \left\{x \in Z;\ \inf_{y\in A} d^{Z}(x,y) < \varepsilon\right\}$ is the ε-halo set of A. We want to make use of this definition to compare "shapes" of trees. Even better, we want to define a distance between tree-like structures that might not even be embedded in a common metric space. In order to do this, we will use the procedure by Gromov:

Definition 2 (Gromov-Hausdorff distance). *Let (X,d) and (X',d') be two compact metric spaces. The Gromov-Hausdorff distance between X and X' is defined by*

$$d_{\mathrm{GH}}\left(X,X'\right) = \inf_{(Z,d^{Z},\Phi,\Phi')} d_{\mathrm{H}}^{Z}\left(\Phi(X),\Phi'(X')\right),$$

where the infimum is taken over all Polish metric spaces (Z,d^{Z}) and over all isometric embeddings $\Phi : X \hookrightarrow Z$ and $\Phi' : X' \hookrightarrow Z$.

Like most concepts in this section, the Gromov-Hausdorff distance was developed in the framework of geometric group theory. Gromov first defined (and used) the notion in his groundbreaking work on polynomial-growth groups ([Gro81]).

The Gromov-Hausdorff distance obviously assigns distance 0 to two isometric compact metric spaces. Therefore, we will consider the space $\mathbb{K}$ of isometry classes of compact metric spaces. On this space, it can then be proven that d_{GH} is in fact a metric (a definite-positive, symmetric function satisfying the triangle inequality). The ensuing metric topology will be called *Gromov-Hausdorff topology*. A natural question is to wonder what its continuous functions look like, and how it compares to other topologies on $\mathbb{K}$.

A first remark in that direction is that Gromov-Hausdorff convergence does not preserve topological invariants such as the genus. For instance, a sphere $\mathbb{S}^2$ with a very "small" handle attached to it (which is a surface of genus 1) will converge in the Gromov-Hausdorff sense, as the handle becomes smaller and smaller, towards the sphere $\mathbb{S}^2$ itself (of genus 0). There

[1] A Polish metric space is a complete and separable metric space (X,d). A topological space is Polish if its topology can be metrized by a complete and separable metric. Bourbaki uses the terminology "Polishable space" (which is not the same as a polishable space).

are stronger topologies on $\mathbb{K}$, such as the uniform topology or the Lipschitz topology, that do preserve topological invariants of this kind.

The definition of the Gromov-Hausdorff distance, although relatively easy to picture and to compute, does however not give a good insight into the topology it defines. There are several different formulations of Gromov-Hausdorff convergence that are more topological in nature, two of which we will describe now. First, the notion of ε-isometries: a map $f : X \to Y$ is an ε-isometry if

$$\sup_{x,x' \in X} \left| d_Y(f(x), f(x')) - d_X(x, x') \right| \leq \varepsilon$$

and if $f(X)$ is an ε-net in Y (meaning that every point in Y is at most at distance ε from a point in $f(X)$). Of course, a 0-isometry is a classical isometry. Yet also for $\varepsilon > 0$, the existence of ε-isometries gives some information about distance in the Gromov-Hausdorff sense:

Proposition 3. *Let X, Y be two compact metric spaces, and let $\varepsilon > 0$. Then, if $d_{\mathrm{GH}}(X, Y) < \varepsilon$, there exists a 2ε-isometry $X \to Y$. Conversely, if there exists an ε-isometry $X \to Y$, then $d_{\mathrm{GH}}(X, Y) < 2\varepsilon$.*

Hence, convergence in the Gromov-Hausdorff sense can be reduced to the existence of ε-isometries with arbitrarily small ε. This description, as well as the simple fact that the set of ***finite*** metric spaces is dense in $\mathbb{K}$, leads to the other important characterization of Gromov-Hausdorff convergence, which shows that it is basically all about convergence of finite sets:

Proposition 4. *Let $(X_n,\ n \geq 1)$ and X_∞ be compact metric spaces. Then X_n converges to X_∞ in the Gromov-Hausdorff topology if and only if for every $\varepsilon > 0$, we can find finite ε-nets $S^{(\varepsilon)} \subset X$ and $S_n^{(\varepsilon)} \subset X_n$ such that $S_n^{(\varepsilon)}$ converges to $S^{(\varepsilon)}$ in the Gromov-Hausdorff sense. Moreover, these sets can be chosen in such a way that, for sufficiently large n, the sets $S_n^{(\varepsilon)}$ have same cardinality as $S^{(\varepsilon)}$.*

As far as Gromov-Hausdorff convergence of *finite* sets is concerned, it can be shown that it is equivalent to stronger convergences, such as the uniform convergence, which means that all distances in the set converge. Therefore, if X_n converges in Gromov-Hausdorff sense to X, and if $S \subset X$ is a finite set, we can find finite subsets $S_n \subset X_n$ converging in a very strong sense to S. This shows an important invariant for Gromov-Hausdorff convergence: if a property can be expressed in terms of the distances between a finite number of points, then it will be preserved in the Gromov-Hausdorff limit. We will see below an application of this fact. Finally, let us describe a pre-compactness criterion in the Gromov-Hausdorff space:

Theorem 5 (Gromov). *Let $\mathscr{X}$ be a class of compact metric spaces, such that*

- *There is a constant D such that for any $X \in \mathscr{X}$, we have $\mathrm{diam}(X) \leq D$*
- *For every $\varepsilon > 0$, there exists a constant $N(\varepsilon) \geq 1$ such that for every $X \in \mathscr{X}$, there exists an ε-net $X^{(\varepsilon)}$ with at most $N(\varepsilon)$ elements.*

Then, the class $\mathscr{X}$ is pre-compact in the Gromov-Hausdorff topology. In other words, any sequence in $\mathscr{X}$ contains a subsequence that converges in the Gromov-Hausdorff sense.

In the next section, we will see how the Gromov-Hausdorff convergence can be used in the context of trees, and how it can be extended to non-compact spaces.

Real trees: the locally compact case

The geometrical objects known as *real trees* (or **R**-trees) have been extensively studied, most notably in the context of geometrical group theory ([Bes02, Chi01, Dre84]). They will be the natural state-space for the random trees we want to consider.

Definition 6 (Real tree). *A metric space (T,d) is a real tree if the following properties are satisfied:*

(i) For every $s,t \in T$, there is a unique isometric map $f_{s,t}$ from $[0,d(s,t)]$ to T such that $f_{s,t}(0)=s$ and $f_{s,t}(d(s,t))=t$.

(ii) For every $s,t \in T$, if q is a continuous injective map from $[0,1]$ to T such that $q(0)=s$ and $q(1)=t$, then $q([0,1]) = f_{s,t}([0,d(s,t)])$.

A summary of the properties of real trees in a probabilistic context can be found in [Eva08]. There is a characterization of real trees, the so-called *four-point condition*, stating that real trees are the only complete, path-connected metric spaces X that satisfy, for all $\mathrm{x}_1,\mathrm{x}_2,\mathrm{x}_3,\mathrm{x}_4 \in X$,

$$d(\mathrm{x}_1,\mathrm{x}_2)+d(\mathrm{x}_3,\mathrm{x}_4) \leq (d(\mathrm{x}_1,\mathrm{x}_3)+d(\mathrm{x}_2,\mathrm{x}_4)) \vee (d(\mathrm{x}_3,\mathrm{x}_2)+d(\mathrm{x}_1,\mathrm{x}_4)).$$

This condition is exactly of the type specified earlier, that is, a property of the distances between a finite number of vertices in the space. Therefore, the set $\mathbb{T}_{\mathrm{cpct}}$ of compact real trees is a closed subspace of $\mathbb{K}$ under Gromov-Hausdorff convergence. Much more can actually be said about its topology, using the pre-compactness criterion for the Gromov-Hausdorff topology. Using the following theorem, we can then define $\mathbb{T}_{\mathrm{cpct}}$-valued random processes with nice properties.

Theorem 7 (Evans, Pitman, Winter [EPW05]). *The set $\mathbb{T}_{\mathrm{cpct}}$ of (isometry classes of) compact real trees, endowed with the Gromov-Hausdorff distance, is a Polish metric space.*

Let us now describe a standard way of defining and studying trees, the *contour process*. Let f be an *excursion*, that is, a continuous, nonnegative function with compact support $[0,\sigma]$ such that $f(0)=f(\sigma)=0$. We can then define a real tree structure in the following way: on the set $[0,\sigma]$, let $\sim_f$ be the equivalence relation such that

$$x \sim_f y \Leftrightarrow f(x)=f(y)=\min_{u\in[x\wedge y, x\vee y]} f(u), \qquad x,y\in[0,\sigma].$$

Then, define a semimetric on $[0,\sigma]$ by

$$d_f(x,y) = f(x)+f(y)-2\min_{u\in[x\wedge y, x\vee y]} f(u), \qquad x,y\in[0,\sigma].$$

Then, $d_f(x,y)=0$ if and only if $x\sim_f y$, so that d_f defines a true metric on the quotient set $[0,\sigma]_{/\sim}$. It can then be checked (see [DL06]) that the metric space $\mathcal{T}_f = ([0,\sigma]_{/\sim}, d_f)$ is a (compact) real tree. We will usually *root* the tree $\mathcal{T}_f$ by distinguishing the equivalence class of 0, which we will note $\varnothing = \overline{0}$. We use the notation p_f for the projection map

$$p_f : [0,\sigma] \to \mathcal{T}_f \quad x \mapsto \overline{x}.$$

This construction was actually well-known in the combinatorial community, since it provides an easy way to encode tree structures in a functional way. When starting from a tree, a contour process can, at least informally, be recovered by imagining a traveler starting at the root and exploring the tree at unit speed, ***always turning left***. Of course, this notion of left-turning needs to be made precise, which we will not try to do here. For a very detailed analysis of the relation between real trees and contour process descriptions, see [Duq06].

However, we can give an example by considering the so-called ***plane trees***, which are discrete trees in which the neighbor vertices of a given vertex are ordered. Let

$$\mathscr{U} = \bigcup_{n\geq 0} \left(\mathbf{N}^*\right)^n,$$

with the convention $(\mathbf{N}^*)^0 = \{\varnothing\}$. The set $\mathscr{U}$ is Ulam's ***universal tree***, and plane trees are finite subsets $T \subset \mathscr{U}$ such that

- the empty word belongs to T: $\varnothing \in T$;
- if a nonempty word belongs to T, its ancestor too : for all $u = (u_1,\dots,u_n) \in T\setminus\{\varnothing\}$, $\overleftarrow{u} = (u_1,\dots,u_{n-1}) \in T$;
- for all $u = (u_1,\dots,u_n) \in T$, there exists an integer $k(u) \geq 0$ such that $(u_1,\dots,u_n,i)$ belongs to T if and only if $1 \leq i \leq k(u)$ (if $k(u) = 0$, none of the children of u belong to T).

For instance, subcritical and critical Galton-Watson trees can be represented by plane trees ([Nev86]). Note that plane trees are a particular instance of compact real trees, rooted at $\varnothing$, with edges of unit length connecting every $u \in T$ to its ancestor $\overleftarrow{u}$. We can also recover a continuous excursion encoding for them by exploring the tree at unit speed in lexicographic order (this is sometimes called the *depth-first search process*).

If T has n vertices, noted $(u_1 = \varnothing,\dots,u_n)$ in lexicographic order, it is then easy to see that the contour process is an excursion $C(T) : [0,2n-2] \to \mathbf{R}_+$. We can also consider the *Łukasiewicz path* associated to T by defining recursively:

$$W_i(T) = \begin{cases} 0 & \text{if } i = 0,\\ W_{i-1}(T) + k(u_{i-1}) - 1 & \text{if } 1 \leq i \leq n,\\ 0 & \text{if } i > n. \end{cases}$$

It is a fundamental fact ([Le 05]) that the Łukasiewicz path of a Galton-Watson tree with subcritical or critical offspring distribution ξ is distributed as a random walk $(W_0,\dots,W_\zeta)$, started at 0, with jump probabilities $\mathbb{P}(W_{n+1} - W_n = k) = \xi(k+1)$ for $k \geq -1$, and stopped at $\zeta = \inf\{n \geq 0,\ W_n = -1\}$. This connection with a simple random walk gives (via Donsker's theorem) a heuristic for the rescaling in Aldous's theorem, and has proved a very useful tool in a number of contexts ([Kor12a, Kor12b, LM11, Mar08]).

Using the contour process correspondence between continuous excursions and real trees, the convergence from Theorem 1 can be reformulated in terms of Gromov-Hausdorff convergence. Indeed, the map $f \mapsto \mathscr{T}_f$, defined on the set of excursions, can be shown to be

Lipschitz continuous for the topology of uniform convergence on compact subsets of $\mathbf{R}_+$: it is proven in [DL05] that, if f and g are two excursions, then

$$d_{\mathrm{GH}}\left(\mathscr{T}_f, \mathscr{T}_g\right) \leq 2\|f-g\|_\infty,$$

where $\|\cdot\|_\infty$ is the uniform norm. Therefore, Aldous's theorem implies that the Galton-Watson trees T_n encoded by their rescaled contour processes f_n converge in the Gromov-Hausdorff topology to the tree encoded by the normalized Brownian excursion.

However, the theory of Gromov-Hausdorff convergence only accounts for *compact* real trees, whereas many natural examples we want to consider are not compact. For instance, supercritical Galton-Watson trees are infinite with positive probability. Although the definition of the Gromov-Hausdorff distance still makes sense for noncompact spaces, it is far too restrictive. Indeed, consider the example of two-dimensional spheres $(r_n\mathbb{S}^2)$ embedded in $\mathbf{R}^3$, with some sequence of radii $(r_n\ n \geq 1)$ converging to infinity. This sequence does not converge to any compact metric space in the Gromov-Hausdorff sense, yet when looking at a bounded neighbourhood of any point in $\mathbb{S}^2$, this neighbourhood will look more and more like a subset of a two-dimensional plane. In this sense, we would like to say that *locally*, the spheres $(r_n\mathbb{S}^2,\ n \geq 1)$ converge to the two-dimensional plane. The notion of convergence which makes this precise is the *pointed* Gromov-Hausdorff convergence.

In order to make things more tractable, it is convenient to work with *boundedly compact* metric spaces (spaces in which the balls of finite radius are compact). This is not too restrictive when working with real trees, since the Hopf-Rinow theorem guarantees that all locally compact and complete real trees are indeed boundedly compact. Furthermore, we always assume that the spaces are *pointed*, which means that some specific point (the root, in our case) is distinguished. Then, following Duquesne and Winkel ([DW07]), we may define a metric on isometry classes of pointed, locally compact and complete real trees by:

$$d_{\mathrm{pGH}}\left((\mathscr{T}_1, d_1, \varnothing_1), (\mathscr{T}_2, d_2, \varnothing_2)\right) = \sum_{k\geq 1} 2^{-k} d_{\mathrm{GH}}\left(B_{\mathscr{T}_1}(\varnothing_1, k), B_{\mathscr{T}_2}(\varnothing_2, k)\right),$$

where $B_X(x,r)$ denotes the closed ball of radius r around $x \in X$. For more information about the pointed Gromov-Hausdorff topology, we refer to [DW07] or to chapter 8 in [BBI01]. For our purpose, what is important is that the set $\mathbb{T}_{\mathrm{loc}}$ of (pointed isometry classes of) pointed, locally compact and complete real trees, endowed with the pointed Gromov-Hausdorff distance d_{pGH}, is again a Polish metric space.

A natural question is to ask for the relationship between the two metrics we just defined: on the one hand, the Gromov-Hausdorff distance between compact metric spaces and on the other hand, the pointed Gromov-Hausdorff distance between pointed, locally compact and complete real trees. When restricting the metric d_{pGH} to the set of rooted compact real trees, the two topologies are equivalent, in the sense that if $(T_n, \varnothing_n)$ converges to $(T, \varnothing)$ in the pointed Gromov-Hausdorff sense, then T_n converges to T in the classical Gromov-Hausdorff sense. Conversely, if $T_n \to T$ in the Gromov-Hausdorff sense, and if $\varnothing$ is any point in T, there exists a sequence of points $\varnothing_n \in T_n$ such that $(T_n, \varnothing_n)$ converges in the pointed Gromov-Hausdorff topology to $(T, \varnothing)$.

The Gromov-weak topology

The concept of Gromov-Hausdorff convergence is strong enough to capture some of the geometrical properties of converging sequences of spaces. However, it is also natural to consider random metric spaces on which *measures* are defined. Therefore, topologies that take these measures into account are needed. In order to compare measures, a standard tool is the *Prokhorov metric*, which is defined by

$$d_{\mathrm{Pr}}(\mu,\nu)=\inf\left\{\varepsilon>0;\ \mu(A)\leq\nu(A^{\varepsilon})+\varepsilon \text{ and } \nu(A)\leq\mu(A^{\varepsilon})+\varepsilon\ \text{ for any closed set } A\right\},$$

when μ and ν are finite measures[2] defined on some common Polish metric space Z. It is well-known that the Prokhorov metric metrizes the topology of weak convergence for finite measures on Z (*i.e.* convergence against bounded continuous functions), and that the set of *probability* measures on Z is closed for this topology. A beautiful theorem of Strassen relates the Prokhorov distance between two probability measure and the existence of a good *coupling*: if (S,d) is a separable metric space, and if μ,ν are probability measures on S, then $d_{\mathrm{Pr}}(\mu,\nu)\leq\varepsilon$ if and only if there exists a measure M on $S\times S$ such that

- $M(dx\times S)=\mu(dx)$ and $M(S\times dy)=\nu(dy)$
- We have $M(\{d(x,y)\geq\varepsilon\})\leq\varepsilon$.

In other words, if $d_{\mathrm{Pr}}(\mu,\nu)\leq\varepsilon$, there exists a $S\times S$-valued r.v. (X,Y) that has μ and ν as marginals, and such that $\mathbb{P}(d(X,Y)\geq\varepsilon)\leq\varepsilon$.

To paraphrase Villani ([Vil09]), there are essentially two viewpoints when dealing with metric spaces endowed with measures. The first is to focus solely on the measures and their support and so, to ignore zero-measure sets that might be relevant for the geometry of the space. This is best captured by the *Gromov-weak* topology. The second point of view is to consider that both the geometrical structure and the measures are important, and to say that two spaces are close when their metric structures are Hausdorff-close and when the measures they carry are simultaneously Prokhorov-close. The topologies in that framework are the so-called *Gromov-Hausdorff-Prokhorov* topologies, for which we present a contribution by the author in the next section.

In the probabilistic context, the Gromov-weak topology was explored in [GPW08] with applications to Λ-coalescent tree-like spaces in mind. The relevant class of spaces are *metric measure spaces*, that is, Polish metric spaces (X,d) endowed with a probability measure μ. Just like in the Gromov-Hausdorff case, we do not distinguish between two spaces (X,d,μ) and (X',d',μ') such that there is an isometry $\mathrm{Supp}(\mu)\to\mathrm{Supp}(\mu')$ that transports μ onto μ'. The set of (classes of) metric measure spaces is usually noted $\mathbb{M}$.

On the class of metric measure spaces, there is a metric defined by

$$d_{GP}\left((X,d,\mu),(X',d',\mu')\right)=\inf_{(Z,\Phi,\Phi')} d^{Z}_{\mathrm{Pr}}((\Phi)_*\mu,(\Phi')_*\mu'),$$

[2] All the measures considered in this introduction are of course assumed to be *Borel* measures. In order to keep adjectives to a minimum, we will not systematically emphasize this point, but it should be clear in the reader's mind that we always implicitly make this assumption.

where the infimum is taken over all isometric embeddings $\Phi : X \hookrightarrow Z$ and $\Phi' : X' \hookrightarrow Z$ into a common Polish metric space Z, and where $\Phi_* \mu$ is the push-forward of the measure μ by the map Φ.

The topology on $\mathbb{M}$ generated by d_{GP} is described in great detail in [GPW08], where it is proven in particular that it coincides with the Gromov-weak topology studied[3] by Gromov ([Gro81]). The fundamental idea behind the Gromov-weak topology is the following: a sequence of metric measure spaces (X_n, d_n, μ_n) converges if and only if all subspaces spanned by a finite number of points, sampled according to μ_n, converge. In some sense, the Gromov-weak topology is "weaker" than Gromov-Hausdorff convergence : recall that *all* properties pertaining to finite subspaces are continuous in the Gromov-Hausdorff topology. For example, as was already mentioned, the maximal distance to the root is continuous in the Gromov-Hausdorff topology, but not in the Gromov-weak topology. A good continuous function in the Gromov-weak sense would be the mean distance to the root, *i.e.* the distance from the root to a vertex sampled according to the measure.

An additional feature of the Gromov-Prokhorov metric, besides metrizing the Gromov-weak topology, is that the ensuing metric space $(\mathbb{M}, d_{\mathrm{GP}})$ is again a Polish metric space, so that $\mathbb{M}$-valued random processes can be defined and studied with powerful analytical methods (see for instance [GPW12]).

The Gromov-Hausdorff approach and the Gromov-Prokhorov approach can be combined, forming, necessarily, the Gromov-Hausdorff-Prokhorov metric. The relevant class of spaces is the class $\mathbb{M}_{\mathrm{cpct}}$ of *compact* metric measure spaces, carrying probability measures. They are considered up to measure-preserving isometries (true isometries, obviously, not just isometric maps between the supports of the measures). Following Miermont ([Mie09]), if (X, d, μ) and (X', d', μ') are classes in $\mathbb{M}_{\mathrm{cpct}}$, we define a metric by

$$d_{\mathrm{GHP}}^{\mathrm{cpct}}(X, X') = \inf_{(\Phi, \Phi, Z)} \left(d_{\mathrm{H}}^{Z}(\Phi(X), \Phi'(X')) \vee d_{\mathrm{Pr}}^{Z}((\Phi)_* \mu, (\Phi')_* \mu') \right),$$

where the infimum is taken over all isometric embeddings $\Phi : X \hookrightarrow Z$ and $\Phi' : X' \hookrightarrow Z$ into a common Polish metric space Z.

The topology induced by $d_{\mathrm{GHP}}^{\mathrm{cpct}}$ was described by Villani ([Vil09]) as *measured Gromov-Hausdorff topology* : if $(\mathscr{X}_n = (X_n, d_n, \mu_n), n \geq 1)$ is a sequence in $\mathbb{M}_{\mathrm{cpct}}$ and if $\mathscr{X} = (X, d, \mu) \in \mathbb{M}_{\mathrm{cpct}}$, it can be checked that $d_{\mathrm{GHP}}^{\mathrm{cpct}}(\mathscr{X}_n, \mathscr{X})$ converges to 0 if and only if there exists a sequence of measurable ε_n-isometries $f_n : \mathscr{X}_n \to \mathscr{X}$ such that $\varepsilon_n \to 0$ and such that the transported measures $(f_n)_* \mu_n$ converge weakly to μ. There are other metrics on $\mathbb{M}_{\mathrm{cpct}}$ that metrize the same topology, see for instance [EW06]. It can also be checked that $(\mathbb{M}_{\mathrm{cpct}}, d_{\mathrm{GHP}}^{\mathrm{cpct}})$ is a Polish metric space.

Of course, for the reasons mentioned earlier, when restricted to $\mathbb{M}_{\mathrm{cpct}}$, the topology induced by $d_{\mathrm{GHP}}^{\mathrm{cpct}}$ is stronger than the one induced by d_{GP}. It should however be stressed that the two approaches are actually not as different as they look. Indeed, the main difference is how these two topologies deal with sets of measure 0 that are somehow "significant" in a metric sense. When restricting ourselves to metric measure spaces (X, d, μ) such that $\mathrm{Supp}(\mu) = X$,

[3] Gromov describes another metric, the so-called $\underline{\square}_1$-metric, that also metrizes the Gromov-weak topology, see [LÏ2].

the two approaches essentially coincide. For instance, when the measure satisfies a so-called ***doubling*** property, it can be shown that d_{GP} convergence is essentially equivalent to d_{GHP}-convergence. However, the spaces with which we are concerned do ***not*** satisfy this kind of conditions, so that we need to make a choice. In the next paragraph, we describe a Gromov-Hausdorff-Prokhorov topology for general (locally compact) real trees, with boundedly finite measures.

The locally compact and boundedly finite case

In this paragraph, we shall summarize the results of Chapter 2, which contains the paper [ADH12b], submitted for publication.

The main purpose of this work is to provide a natural state-space for Lévy tree-valued processes. They might not be compact (in the supercritical case) and they might carry σ-finite measures. Therefore, we use the space $\mathbb{T}$ of rooted, locally compact, complete real trees, carrying locally finite Borel measures, in short, *w-trees*. These spaces will always be considered up to rooted isometries that preserve the measure.

It is a consequence of the Hopf-Rinow theorem that if $\mathscr{T}$ is an element of $\mathbb{T}$, then for every $r > 0$, the closed ball $B_{\mathscr{T}}(\emptyset, r)$ is a rooted compact real tree and the restriction of μ to $B_{\mathscr{T}}(\emptyset, r)$ is a finite Borel measure. This allows to define a metric on these spaces: if $T_1 = (\mathscr{T}_1, d_1, \emptyset_1, \mu_1)$ and $T_2 = (\mathscr{T}_2, d_2, \emptyset_2, \mu_2)$ are two rooted compact real trees, endowed with finite Borel measures, then define

$$d^{\mathrm{c}}_{\mathrm{GHP}}(T_1, T_2) = \inf_{(Z,\Phi_1,\Phi_2)} \left(d^Z(\Phi_1(\emptyset_1), \Phi_2(\emptyset_2)) + d_H^Z(\Phi_1(\mathscr{T}_1), \Phi_2(\mathscr{T}_2)) + d_{\mathrm{Pr}}^Z((\Phi_1)_*\mu_1, (\Phi_2)_*\mu_2) \right),$$

where the infimum is taken over all Polish metric spaces (Z, d^Z) and over all isometric embeddings $\Phi_1 : \mathscr{T}_1 \hookrightarrow Z$ and $\Phi_2 : \mathscr{T}_2 \hookrightarrow Z$. We then use the powerful idea of localization to define a metric on the space $\mathbb{T}$ in the following way:

$$d_{\mathrm{GHP}}(\mathscr{T}_1, \mathscr{T}_2) = \int_0^\infty \mathrm{e}^{-r} \left(1 \wedge d^{\mathrm{c}}_{\mathrm{GHP}}(B_{\mathscr{T}_1}(\emptyset_1, r), B_{\mathscr{T}_2}(\emptyset_2, r))\right) \, dr.$$

It can be checked that d_{GHP} defines a metric on $\mathbb{T}$. When restricted to the set of *compact* real trees, endowed with finite Borel measures, the topology generated by d_{GHP} then coincides with the topology defined by $d^{\mathrm{c}}_{\mathrm{GHP}}$.

On $\mathbb{T}$, the topology generated by d_{GHP} coincides with the ***pointed measured Gromov-Hausdorff*** topology defined in non-metric terms by Fukaya ([Fuk87]) as well as Lott and Villani ([LV09]), in the following sense: if $(\mathscr{T}_n = (T_n, \emptyset_n, d_n, \mu_n), n \geq 1)$ is a sequence in $\mathbb{T}$ and if $\mathscr{T} \in \mathbb{T}$, then $d_{\mathrm{GHP}}(\mathscr{T}_n, \mathscr{T}) \to 0$ if and only if there are sequences $R_n \to \infty$ and $\varepsilon_n \to 0$, as well as pointed, measurable ε_n-isometries

$$f_n : B_{T_n}(\emptyset_n, R_n) \to B_T(\emptyset, R_n)$$

such that the transported measures $(f_n)_*\mu_n$ converge ***vaguely***[4] to μ. This is an important point: in the d_{GHP} metric, the measures converge vaguely, which means that loss of mass might occur.

[4] Recall that vague convergence on a locally compact space is convergence against all continuous functions vanishing outside a compact set.

Indeed, consider the set of locally compact, complete real trees endowed with *finite* measures. On this set, we can naturally extend the Gromov-Prokhorov metric d_{GP}, which gives a topology that is neither stronger (because of geometrical structure) nor weaker than the Gromov-Hausdorff-Prokhorov topology (for instance, because the set of locally compact, complete real trees endowed with *probability* measures is closed for d_{GP} and not for d_{GHP}).

However, when looking at *compact* real trees carrying probability measures, the topology induced by d_{GHP} is the same as the topology induced by $d_{\mathrm{GHP}}^{\mathrm{cpct}}$, in the sense that if $d_{\mathrm{GHP}}(T_n, T) \to 0$, then the unrooted trees T_n converge in the $d_{\mathrm{GHP}}^{\mathrm{cpct}}$ sense to T. Conversely, if $d_{\mathrm{GHP}}^{\mathrm{cpct}}(T_n, T) \to 0$, then for any root $\emptyset \in T$, there exists a sequence of roots $\emptyset_n \in T_n$ such that $(T_n, \emptyset_n)$ converges in the d_{GHP} sense to $(T, \emptyset)$.

To make things clearer, it is enlightening to think about the convergence of the measures alone. The d_{GHP} topology "corresponds" to vague convergence of boundedly finite measures on locally compact spaces, whereas the d_{GP} and $d_{\mathrm{GHP}}^{\mathrm{cpct}}$ topologies corresponds to weak convergence for probability measures. When considering only compact spaces carrying finite measures, the vague and weak topologies coincide.

In order to study the d_{GHP}-topology, the following criterion describing the compact sets of $(\mathbb{T}, d_{\mathrm{GHP}})$ is very useful.

Theorem 8 (Abraham, Delmas, H. [ADH12b]). *Let $\mathscr{C}$ be a subset of $\mathbb{T}$, such that for every $r \geq 0$:*

(i) For every $\varepsilon > 0$, there exists a finite integer $N(r, \varepsilon) \geq 1$, such that for any $(\mathscr{T}, d, \emptyset, \mu) \in \mathbb{T}$, there is an ε-net of $B_{\mathscr{T}}(\emptyset, r)$ with cardinal less than $N(r, \varepsilon)$.

(ii) There exists $M(r) < \infty$ such that for every $(\mathscr{T}, d, \emptyset, \mu) \in \mathscr{C}$, we have $\mu(B_{\mathscr{T}}(\emptyset, r)) < M(r)$.

Then $\mathscr{C}$ is relatively compact: every sequence in $\mathscr{C}$ admits a sub-sequence that converges in the d_{GHP} topology.

This pre-compactness criterion is instrumental in the proof of the completeness of the metric space $(\mathbb{T}, d_{\mathrm{GHP}})$, since a pre-compact Cauchy sequence always converges. This, along with the density of "finite" length spaces with rational edge-lengths entails :

Theorem 9 (Abraham, Delmas, H. [ADH12b]). *The metric space $(\mathbb{T}, d_{\mathrm{GHP}})$ is complete and separable.*

This provides a correct framework for Lévy tree-valued processes for trees that are not necessarily compact, as we shall see in the following section. Indeed, we can also prove the following inequality, which shows that for any excursion f, the map

$$\left([0,\sigma_f], \mathrm{Leb}_{|[0,\sigma_f]}\right) \mapsto \left(\mathscr{T}_f = p_f([0,\sigma_f]), \mathbf{m}^{\mathscr{T}_f} = (p_f)_* \mathrm{Leb}_{|[0,\sigma_f]}\right)$$

is continuous, hence measurable in the Gromov-Hausdorff-Prokhorov topology, just as Duquesne and Le Gall's inequality showed for the Gromov-Hausdorff topology on compact metric spaces.

Proposition 10 (Abraham, Delmas, H. [ADH12c]). *If $f : [0, \sigma_f] \to \mathbf{R}_+$ and $g : [0, \sigma_g] \to \mathbf{R}_+$ are continuous excursions, then we have:*

$$d_{\mathrm{GHP}}(\mathscr{T}_f, \mathscr{T}_g) \leq 6\|f - g\|_\infty + |\sigma_f - \sigma_g|.$$

1.2 Tree-valued processes

The different topologies introduced in the previous section were introduced in order to provide appropriate state-spaces for stochastic processes. We will review the existing literature on continuum tree-valued processes ; however, we will focus mainly on the *pruning process* of Abraham-Delmas ([AD12a]), which is Lévy tree-valued.

Lévy trees

There are several possible definitions of Lévy trees. We will present here the historical definition of Le Gall-Le Jan ([LL98b, LL98a]) and Duquesne-Le Gall ([DL02, DL05]), using the characterization of trees by their contour processes, which, as we saw, behaves nicely with respect to the topologies we use on tree-spaces.

Let ψ be a real-valued function, such that there exists $\alpha \in \mathbf{R}$, $\beta \geq 0$ and a σ-finite measure Π on $(0,\infty)$ satisfying $\int (1 \wedge x^2)\Pi(dx) < \infty$, such that

$$\psi(u) = \alpha u + \beta u^2 + \int_{(0,\infty)} \left(\mathrm{e}^{-ux} - 1 + ux\mathbf{1}_{\{x<1\}}\right) \Pi(dx). \tag{1}$$

In this context, we will call ψ a *branching mechanism*. Of course, the function ψ can be seen as the Laplace exponent of a spectrally positive Lévy process. Let X be such a Lévy process, that is, a càdlàg process with independant, stationary increments such that

$$\mathbb{E}\left[\mathrm{e}^{-uX_t}\right] = \exp(t\psi(u)), \quad t \geq 0, u \geq 0.$$

We will always make the following assumptions on ψ:

Assumption 1 (Infinite variation). *We have $\beta > 0$ or $\int_{(0,1)} x\,\Pi(dx) = \infty$.*

This assumption is equivalent to the paths of the Lévy process X having infinite variation a.s. Although the construction is still possible in the finite variation case, the ensuing tree has a very different structure, which is essentially discrete[5].

Assumption 2 (Conservativity). *We have $\int_{(0,\varepsilon]} du/|\psi(u)| = \infty$ for any $\varepsilon > 0$.*

Note that Assumption 2 is always satisfied if $\psi'(0+) > -\infty$, which is equivalent to the integrability condition $\int_{(1,\infty)} x\Pi(dx) < \infty$. However, there exist branching mechanisms that are conservative and yet $\psi(0+) = -\infty$. For instance, $\psi(u) = u\log(u)$ is a very important example of such a branching mechanism, associated to the *Neveu tree*, as well as the *Bolthausen-Sznitman coalescent*. The branching mechanism ψ is said to be *subcritical*, *critical* or *supercritical*, according to whether $\psi'(0+)$ is positive, equal to zero or negative (the latter encompassing the case $\psi'(0+) = -\infty$).

A third, more restrictive assumption will sometimes be necessary. Notice that it implies Assumption 1, but the converse is not true.

[5]For a more precise description, including the connection with queuing processes, see [LL98b]

Assumption 3 (Grey condition). *We have* $\int^{\infty} du/|\psi(u)| < \infty$.

The contour process of Lévy trees, called the ***height process***, is defined, in the critical or subcritical case, under Assumptions 1 and 2, in the following way. For any $t \geq 0$, let $(X_s^{(t)} = X_t - X_{(t-s)-},\ 0 \leq s \leq t)$ be its time-reversed process and let $S^{(t)}$ be the supremum process of $X^{(t)}$. The local time process at 0 of $S^{(t)} - X^{(t)}$ (suitably normalized) will be called $L^{(t)}$. Then, there exists ([DL02]) a l.c.s modification of the process $(L_t^{(t)},\ t \geq 0)$, which we will note $(H_t,\ t \geq 0)$. It is important to notice that

$$H_t = 0 \Leftrightarrow X_t = \inf_{u \in [0,t]} X_u, \tag{2}$$

since if X_t is a current minimum time, then $S_s^{(t)} = 0$ for all $0 \leq s \leq t$ and $X_s^{(t)} < 0$ for all $0 < s \leq t$ by definition. Therefore, we can take $-I_t = -\inf_{u \in [0,t]} X_u$ as a local time process for H at level 0. If $x > 0$, we note T_x the first time at which the height process has accumulated more than x local time at 0. The distribution of the stopped height process $(H_t,\ 0 \leq t \leq T_x)$ will be noted $\mathbb{P}_x^{\psi}$.

Now, $X - I$ is a Markov process, for which 0 is a regular point. Therefore, we can consider its excursion measure above 0, suitably normalized, which we will note $\mathbb{N}^{\psi}$. It is not hard to see, as in (2), that H_t only depends on the excursion of $X - I$ that straddles t. Therefore, it is possible to define H under the excursion measure $\mathbb{N}^{\psi}$. Of course, $\mathbb{N}^{\psi}$ is not a probability measure, but a σ-finite measure defined on the set of positive excursions $[0,\sigma] \to \mathbf{R}_+$. Under $\mathbb{P}_x^{\psi}$, we will also note $\sigma = T_x$ the length of the excursion of the height process. The distribution of σ is then specified by its Laplace transform:

$$\mathbb{P}_x^{\psi}\left[\mathrm{e}^{-\lambda\sigma}\right] = \exp\left(-x\mathbb{N}^{\psi}[1 - \exp(-\lambda\sigma)]\right) = \exp\left(-x\psi^{-1}(\lambda)\right), \quad \lambda > 0.$$

If $a > 0$, a local time process $(L_s^a(H),\ 0 \leq s \leq \sigma)$ at level a can be constructed for H, under $\mathbb{P}_x^{\psi}$ or under $\mathbb{N}^{\psi}$ (see [DL02] for more details). The total amount of local time spent by the height process at level $a > 0$, $L_\sigma^a(H)$, is the continuum analog of the total generation size in a Galton-Watson process. Indeed, we will see below that the local time processes of the height process also have a branching structure, for which the height process defines a compatible genealogy.

It should be noted that the height process is not, in general, a Markov process, except in the case where $\Pi = 0$, where it is distributed as a (β-scaled) Brownian motion with drift $-\alpha$, reflected at 0. This makes it rather unwieldy to work with; however, it is a simple functional of a more general Markov process, the ***exploration process***. For a detailed study of the height process, the exploration process and much, much more, see [DL02].

In order to define the associated real tree, a bit more regularity is required. Indeed, the contour process needs to be continuous for Proposition 10 to apply[6]. Under Assumptions 1 and 2, the continuity of the height process is actually equivalent to Assumption 3, which leads to the following definition.

[6] The case where H is not continuous, but where Assumptions 1 and 2 still hold is interesting. The height process is then a.s. a Darboux, l.c.s. function, which is unbounded on every interval. Although the definition of the tree using H is still possible in that case, the tree will no longer be locally compact, which requires new topological insights. The Neveu tree, encoded by the height process with branching mechanism $\psi(u) = u\log(u)$ is in that case.

Definition 11 (Lévy tree). *Let ψ be a critical or subcritical branching mechanism satisfying Assumptions 1, 2 and 3, and let $x>0$. The Lévy tree with branching mechanism ψ is the real tree defined, under $\mathbb{P}^{\psi}_x$ or under $\mathbb{N}^{\psi}$, by the ψ-height process $(H_t,\ 0\le t\le\sigma)$, through the contour process description.*

The distribution of this tree is still written $\mathbb{P}^{\psi}_x(d\mathscr{T})$ or $\mathbb{N}^{\psi}[d\mathscr{T}]$. Several constructions exist to extend this definition to the supercritical case for ψ. Abraham and Delmas ([AD12a]) use a Girsanov relation on the truncated critical Lévy tree to define the truncated supercritical Lévy tree. Let us briefly recall the construction. If ψ is a supercritical branching mechanism, let θ^* be the largest root of ψ'. The branching mechanism $\psi_{\theta^*}(u)=\psi(u+\theta^*)$ is then critical, so that we can consider the distribution $\mathbb{P}^{\psi_{\theta^*}}_x$ (resp. $\mathbb{N}^{\psi_{\theta^*}}$) of the ψ_{θ^*}-Lévy tree.

Definition 12 (Supercritical Lévy trees). *Let $a\ge 0$. The distribution $\mathbb{P}^{\psi,a}_x$ (resp. $\mathbb{N}^{\psi,a}$) of the supercritical height process truncated at level a is defined with respect to the* critical *truncated distribution by the absolute continuity relation(s) :*

$$\frac{d\mathbb{P}^{\psi,a}_x}{d\mathbb{P}^{\psi_{\theta^*},a}_x}=\exp\left(-\theta^* x+\theta^* L^a_\sigma(H)+\psi(\theta^*)\int_0^a L^b_\sigma(H)\,db\right),$$
$$\frac{d\mathbb{N}^{\psi,a}}{d\mathbb{N}^{\psi_{\theta^*},a}}=\exp\left(\theta^* L^a_\sigma(H)+\psi(\theta^*)\int_0^a L^b_\sigma(H)\,db\right).$$

Using a Kolmogorov argument, it is then possible to define the distribution $\mathbb{P}^{\psi}_x(d\mathscr{T})$ (resp. $\mathbb{N}^{\psi}[d\mathscr{T}]$) of the supercritical Lévy tree. Although this tree might no longer be compact (in the supercritical case, we have $\sigma=\infty$ with positive probability), it is locally compact by construction, and thus fits in the framework of locally compact, complete length spaces of the previous section. We will list below several properties of Lévy trees, regardless of their criticality. Most of the results in this section come either from the book [DL02] or the seminal paper [DL06] which translated the theory of the exploration process into the language of real trees.

Let us now describe more precisely what happens when cutting a Lévy tree at level $a>0$. Let

$$\mathcal{N}^{\mathscr{T}}_a(dx,d\mathscr{T}')=\sum_{i\in I_a}\delta_{(x_i,\mathscr{T}^i)}(dx,d\mathscr{T}')$$

be the point measure associated to the set of subtrees $(\mathscr{T}^i,\ i\in I_a)$ started at level a, that is, the (closure of the) connected components of the set $\{s\in\mathscr{T},\ d(\emptyset,s)>a\}$, as well as the points $(x_i,\ i\in I_a)$ at which they are attached to the level set $\mathscr{T}(a)=\{s\in\mathscr{T},\ d(\emptyset,s)=a\}$. Using the local times at level a of the height process $(L^a_s(H),\ s\ge 0)$, it is possible to define measures $(\ell^a(ds),\ a>0)$, called *local time measures*, which satisfy, for every fixed $a>0$, $\mathbb{N}^{\psi}$-a.e. :

- The measure ℓ^a is supported on the level set $\mathscr{T}(a)=\{s\in\mathscr{T},\ d(\emptyset,s)=a\}$
- If ϕ is a bounded, continuous function on $\mathscr{T}$, then

$$\langle\ell^a,\phi\rangle=\lim_{\varepsilon\downarrow 0}\frac{1}{b(\varepsilon)}\int\phi(x)\mathbf{1}_{\{h(\mathscr{T}')\ge\varepsilon\}}\mathcal{N}^{\mathscr{T}}_a(dx,d\mathscr{T}'),$$

  with $b(\varepsilon)=\mathbb{N}^{\psi}[H(\mathscr{T})\ge\varepsilon]$.

The local time measures ℓ^a can be constructed in such a way that the mapping $a \mapsto \ell^a$ is $\mathbb{N}^\psi$-a.e. càdlàg for the weak topology on finite measures on $\mathcal{T}$. The discontinuities of this mapping present an interesting structure, since they are related to the infinitary nodes in the tree. Indeed, it is a striking feature of Lévy trees that the degree $n(x)$ of vertices $x \in \mathcal{T}$ (the number of connected components of $\mathcal{T} \setminus \{x\}$) can only be 1, 2, 3 or infinite. If $n(x) = 1$, x is a *leaf* of the tree ; the set of all leaves will be noted $\mathrm{Lf}(\mathcal{T})$. Else, we say that x belongs to the *skeleton* of the tree, which we note $\mathrm{Ske}(\mathcal{T})$. If $n(x) \geq 3$, x is a *branching point*—binary if $n(x) = 3$ and infinitary if $n(x) = \infty$. There are binary branching points in $\mathcal{T}$ if and only if $\beta > 0$ in the branching mechanism of $\mathcal{T}$ and there are infinitary branching points if and only if $\Pi \neq 0$. The formula relating discontinuities of $(\ell^a,\ a > 0)$ and infinitary branching points is as follows: if b is a discontinuity of the local time process, then there is a unique infinitary branching point x_b in $\mathcal{T}(b)$ and $\ell^b = \ell^{b-} + \Delta_b \delta_{x_b}$, where $\Delta_b > 0$ is called *width* (or *local time*) of the node x_b and can be obtained by the approximation

$$\Delta_b = \lim_{\varepsilon \to 0} \frac{1}{b(\varepsilon)} Z(x_b, \varepsilon), \tag{3}$$

where $Z(x_b, \varepsilon)$ is the number of sub-trees originating from x_b with height larger than ε.

The *branching property* of Lévy trees then states that for every $a > 0$, the conditional distribution of the point measure $\mathcal{N}_a^{\mathcal{T}}(dx, d\mathcal{T}')$ under $\mathbb{N}^\psi[d\mathcal{T} | H(\mathcal{T}) > a]$, given the truncated tree $\{s \in \mathcal{T},\ d(\emptyset, s) \leq a\}$, is that of a Poisson point measure on $\mathcal{T}(a) \times \mathbb{T}$ with intensity $\ell^a(dx)\mathbb{N}^\psi[d\mathcal{T}']$. Regarding cutting at heights, Lévy trees also enjoy a *regenerative* property. If $a, h > 0$, let $Z(a, a+h)$ be the number of subtrees of $\mathcal{T}$ started at level a with height greater than h. Then, it can be proven that, conditionally on $\{Z(a, a+h) = p\}$ for some integer p, the p subtrees are independent, distributed according to $\mathbb{N}[d\mathcal{T} | H(\mathcal{T}) > h]$. In [Wei07], Weill actually proved that this property *characterizes* subcritical and critical Lévy trees among compact continuum random trees.

As was mentioned earlier, one of the main motivations to introduce Lévy trees was to provide a genealogical structure for continuous-state branching processes. This is achieved using the local time measures of the tree. Indeed, if ψ is a branching mechanism satisfying Assumptions 1 to 3, and if $x > 0$, then the total mass process

$$\mathcal{Z}_a = \langle \ell^a, 1 \rangle, \quad a > 0$$

is distributed under $\mathbb{P}_x^\psi$ as a CSBP with branching mechanism ψ, with $\mathcal{Z}_0 = x$. The local time measures ℓ^a can also be integrated to form a measure on the whole tree $\mathcal{T}$, which is called *mass measure*:

$$\mathbf{m}(ds) = \int_0^\infty \ell^a(ds)\ da.$$

A simple application of the occupation times formula shows that the mass measure is really the push-forward of Lebesgue measure on the excursion interval $[0, \sigma]$ coding for the Lévy tree. Therefore, the mass measure is a locally finite measure on $\mathcal{T}$, supported on $\mathrm{Lf}(\mathcal{T})$ (in the sense that $\mathbf{m}(\mathcal{T} \setminus \mathrm{Lf}(\mathcal{T})) = 0$). Sometimes, $\mathbf{m}$ is referred to as the uniform measure on leaves. Notice that $\mathbf{m}$ is without atoms, since the discontinuities of the total mass process $(\mathcal{Z}_a,\ a > 0)$ are countable, hence of zero Lebesgue measure. There is another interesting measure that is naturally defined on Lévy trees: the *length measure* $\ell^{\mathcal{T}}(ds)$ (or $\ell(ds)$ when the context is clear). This is the only Borel measure on $\mathcal{T}$ such that $\ell([\![x,y]\!]) = d(x,y)$. It can be thought of as Lebesgue measure on the skeleton of the tree. Indeed, $\ell(\mathrm{Lf}(\mathcal{T})) = 0$. It is also a σ-finite measure, that is always infinite, whereas the mass measure can be finite, for instance in the critical and subcritical case, where $\sigma = \mathbf{m}(\mathcal{T}) < \infty$ a.e.

As we already mentioned in the motivation section, Lévy trees are the scaling limits for large Galton-Watson trees, defined as follows. Let $(\xi^{(n)},\ n \geq 1)$ be a sequence of critical or subcritical probability measures on $\mathbf{N}$, and let, for any $n \geq 1$, $\mathrm{T}^{(n)}$ be a Galton-Watson tree with offspring distribution $\xi^{(n)}$, and $(Y_k^{(n)},\ k \geq 0)$ the associated Galton-Watson process, started at $Y_0^{(n)} = n$. Then, we can state the following theorem about the scaling limits of the $\mathrm{T}^{(n)}$:

Theorem 13 (Duquesne, Le Gall [DL05]). *Suppose there is a nondecreasing sequence $(\gamma_n,\ n \geq 1)$ of positive integers, converging to $+\infty$, such that the following convergence in distribution holds:*

$$\lim_{n\to\infty} \left(n^{-1} Y^{(n)}_{\lfloor \gamma_n t \rfloor},\ t \geq 0 \right) = (Y_t,\ t \geq 0), \tag{4}$$

in the Skorokhod topology, where the limiting process Y is a CSBP with branching mechanism ψ. Assume further that for every $\delta > 0$,

$$\liminf_{n\to\infty} \mathbb{P}\left(Y^{(n)}_{\lfloor \gamma_n \delta \rfloor} = 0 \right) > 0.$$

Then, for every $a > 0$, the tree $\gamma_n^{-1}\mathrm{T}^{(n)}$, conditioned on $\{H(\mathrm{T}^{(n)}) \geq \lfloor a\gamma_n \rfloor\}$, converges in distribution (for the Gromov-Hausdorff topology) to the Lévy tree $\mathcal{T}$ with branching mechanism ψ, distributed as $\mathbb{N}^{\psi}[d\mathcal{T} \,|\, H(\mathcal{T}) > a]$.

This theorem can be seen as a general, unconditioned version of Aldous's convergence result (Theorem 1). A particular case of Theorem 13 is the case where all the offspring measures ξ_n coincide: $\xi_n = \xi,\ n \geq 1$. In that case, the limiting CSBP must be stable, that is, ψ is necessarily of the form $\psi(u) = u^{\alpha}$ with $\alpha \in (1,2]$. If ξ has finite variance, the limiting CSBP in (4) will be a Feller diffusion, and the limiting tree a (scaled) Brownian tree.

Theorem 13 shows how Lévy trees arise as scaling limits of Galton-Watson trees with initial population size growing infinite. However, Lévy trees can also be obtained as limits of Galton-Watson trees, without rescaling, as was shown in [DW07, DW12]. Duquesne and Winkel consider families $(\mathcal{F}_\lambda,\ \lambda \geq 0)$ of Galton-Watson forests, that is, Galton-Watson trees with exponential edge-lengths, attached by the root to the real line. These families are assumed to be *consistent*, in the sense that, if $\lambda < \mu$, the forest $\mathcal{F}_\lambda$ can be obtained from the forest $\mathcal{F}_\mu$ using some reduction operation such as independent percolation on leaves with parameter λ/μ (but can be much more general, see [DW12]). Then, as $\lambda \to \infty$, the forest $\mathcal{F}_\lambda$ converges a.s. to a Lévy tree $\mathcal{T}$ with branching mechanism ψ determined by the distribution of $\mathcal{F}_1$. The mass measure $\mathbf{m}(ds)$ arises as the weak limit of the empirical measure $\lambda^{-1}\sum_{s\in \mathrm{Lf}(\mathcal{F}_\lambda)} \delta_s$ on the leaves of $\mathcal{F}_\lambda$.

Conditioned Lévy trees: Aldous's tree and stable trees

We will now turn to the theory of conditioned Lévy trees, that is, Lévy trees conditioned on $\{\sigma = 1\}$. This conditioning is degenerate, but can be made rigourous in the α-*stable* case, with $\alpha \in (1,2]$. Indeed, in this case, the paths of the Lévy process X satisfy the following scaling property: if $\gamma > 0$,

$$(\gamma^{-1/\alpha} X_{\gamma t},\ t \geq 0) \overset{(d)}{=} (X_t,\ t \geq 0).$$

The height process H constructed using the excursions of $X - I$ above 0 has a different scaling: if $\gamma > 0$,

$$(\gamma^{1/\alpha - 1} H_{\gamma t},\ t \geq 0) \overset{(d)}{=} (H_t,\ t \geq 0).$$

Using this scaling property, it is not difficult (using for instance the pathwise construction by Chaumont, see [Cha97]) to construct a regular version $(\mathbb{N}^{(u)},\ u > 0)$ of the conditional distributions of $X - I$ given $\{\sigma = u\}$, in the sense that if F is a measurable bounded functional on the space of excursions,

$$\mathbb{N}^{\psi}[F] = \int_0^{\infty} \mathbb{N}^{(u)}[F]\ \mathbb{N}^{\psi}[\sigma \in du].$$

This, and the scaling property of the height process, allows for the definition of the so-called ***normalized excursion*** of the height process, which is a process $(H^{\text{exc}}_t,\ 0 \leq t \leq 1)$ with a.s. continuous paths, which is distributed as the rescaled process $(\sigma^{1/\alpha - 1} H_{\sigma t},\ 0 \leq t \leq 1)$ under $\mathbb{N}[dH]$. The real tree encoded by H^{exc} is a compact rooted real tree, for which the mass measure is a probability measure, the ***α-stable tree.***

In the case $\alpha = 2$, we recover Aldous's CRT, encoded by the normalized Brownian excursion. If $\alpha \in (1,2)$, the stable tree is no longer binary ; all its branching points are infinitary, as is the case in the unconditioned stable Lévy tree. Many things are known about the geometry of the stable trees: for instance, its Hausdorff dimension and packing dimension are equal to $\alpha/(\alpha-1) \in [2,\infty)$. Just like Lévy trees are scaling limits of large Galton-Watson trees, their conditioned version can be expressed as scaling limits of conditioned Galton-Watson trees. Of course, Theorem 1 is the first result in this direction, showing that the Brownian CRT is the scaling limit of critical finite-variance Galton-Watson trees, conditioned on having n vertices, with an edge-rescaling by a factor $\sqrt{n}$. Aldous's result was extended to the stable case by Duquesne ([Duq03]). In this case, we need to consider Galton-Watson trees $\mathrm{T}^{(n)}$ with offspring distribution ξ lying in the attraction domain of a stable distribution with parameter $\alpha \in (1,2)$, conditioned on having n vertices. Rescaling the edges by some factor[7] a_n, we get the Gromov-Hausdorff convergence[8] $a_n^{-1}\mathrm{T}^{(n)} \to \mathcal{T}^{\alpha}$, where $\mathcal{T}^{\alpha}$ is the α-stable Lévy tree.

There are different ways of conditioning Galton-Watson trees by total population size, and still obtaining stable CRTs in the limit. Kortchemski ([Kor12a]) considers Galton-Watson trees conditioned on their total number of leaves. All of these convergence results use the

[7] The factor is such that if $U_1 + \cdots + U_n$ are iid with distribution ξ, then $a_n^{-1}(U_1 + \cdots + U_n - n)$ converges in distribution to an α-stable r.v.

[8] In this result, as in Aldous's result, it is actually convergence of contour functions that is proven, which is slightly stronger than Gromov-Hausdorff convergence.

deep connection between the contour process of a conditioned Galton-Watson tree and the Łukasiewicz random walk. Therefore, convergence results for more general random discrete trees are difficult to obtain because this connection fails. However, in a recent paper ([HM12]), Haas and Miermont showed a very general convergence result for so-called *Markov branching trees*. The limiting trees are *self-similar fragmentation trees* (see below), which are more general than conditioned Lévy trees, since this class contains the Brownian tree, as well as the α-stable trees with index $\alpha \in (1,2)$. A spectacular application of their results is the convergence of uniform unordered trees with n vertices, with out-degree bounded by m, rescaled by a factor $\sqrt{n}$, to the (scaled) Brownian CRT. In the same spirit, Rizzolo ([Riz11]) considers the case of finite-variance Galton-Watson trees, with a more general conditioning on the number of vertices with out-degree in a given subset $A \subset \mathbf{N}$. The limiting tree is again the Brownian CRT.

An alternative way of studying CRTs is by looking at their discrete subtrees. Aldous ([Ald93]) introduced the Brownian CRT as (the closure of) a projective limit of a consistent family of discrete trees $(\mathscr{R}(k),\ k \geq 1)$. Conversely, if $(\mathscr{T}, d, \emptyset, \mathbf{m})$ is a w-tree with *finite* mass measure $\mathbf{m}$, one can define its n-dimensional marginals by sampling n iid leaves $(\mathrm{x}_1, \ldots, \mathrm{x}_n)$ with distribution $\mathbf{m}(d\mathrm{x})/\mathbf{m}(\mathscr{T})$ and by defining T_n to be the subtree of $\mathscr{T}$ spanned by the leaves $(\emptyset, \mathrm{x}_1, \ldots, \mathrm{x}_n)$. The tree T_n is a discrete rooted tree with edge-lengths, with n leaves. In most cases, the distribution of T_n is explicitly known. For the Brownian CRT, it was proven[9] by Aldous ([Ald93]) that T_n is a uniform binary tree with n labelled leaves, and that the edge-lengths of T_n are independent of the shape of T_n, distributed with density

$$f(e_1, \ldots, e_{2n-1}) = (e_1 + \cdots + e_{2n-1}) \exp\left(-(e_1 + \cdots + e_{2n-1})^2/2\right), \quad e_1, \ldots, e_{2n-1} > 0.$$

For the α-stable tree, the marginals were computed in [DL02] (Theorem 3.3.3).

In the unconditioned case, Le Gall ([Le 93b]) considered the Brownian excursion and showed an absolute continuity relation between the distributions of T_n under $\mathbb{N}$ and under $\mathbb{N}^{(1)}$: if F is a measurable functional on $\mathbb{T}_{\mathrm{cpct}}$, then

$$\mathbb{N}^{(1)}\left[F(\mathrm{T}_n)\right] = \mathbb{N}\left[\ell(\mathrm{T}_n)\, \mathrm{e}^{-\ell(\mathrm{T}_n)^2/2} F(\mathrm{T}_n)\right].$$

Also, in the unconditioned Lévy case, if a random number M of leaves are sampled, where M is a Poisson(λ)-distributed independent random variable, the marginal distribution of T_M was shown in [DL02] to be a Galton-Watson tree starting from one individual, with exponential edge-lengths (with parameter $\psi'(\psi^{-1}(\lambda))$), whose offspring distribution $\xi(\lambda)$ is given by its generating function

$$\mathbb{E}\left[r^{\xi(\lambda)}\right] = r + \frac{\psi((1-r)\psi^{-1}(\lambda))}{\psi^{-1}(\lambda)\psi'(\psi^{-1}(\lambda))}.$$

This observation, as well as the invariance of the family of distributions $(\xi(\lambda),\ \lambda > 0)$ under certain reduction operations, led to the definition of Lévy trees as an increasing limit of Galton-Watson trees with offspring distribution $\xi(\lambda)$, when $\lambda \to \infty$, in [DW07, DW12].

Finally, we should mention that the recent years have seen the beginning of a theory of random processes defined *on* Brownian or stable CRTs. For instance, authors were able to

[9] Actually, in Aldous's framework, this was the definition of the Brownian CRT.

compute classical analysis features of CRTs such as volume growth or heat kernel estimates ([Cro07, CH10]) as well as study the behavior of classical processes such as Brownian motion on the CRT ([AEW12, Cro09]), which is *not* the same object as Brownian motion indexed by the CRT (the Brownian snake). In the opinion of the author, this is a very rich and exciting subject which will see many developments in the years to come.

Spinal decompositions

A particularly convenient way of describing trees is by using so-called *spinal decompositions*. The general principle is always the same: given a rooted real tree $(\mathcal{T},\emptyset)$, and given a leaf x of the tree, consider the spine $[\![\emptyset,\mathrm{x}]\!]$ linking the root $\emptyset$ to x. The set $\mathcal{T}\setminus[\![\emptyset,\mathrm{x}]\!]$ is then a forest $\mathcal{F}^{\mathrm{x}}=(\mathcal{X}_i,\ i\in I)$ of disjoint subtrees of $\mathcal{T}$. For any connected component $\mathcal{X}_i$ of $\mathcal{T}\setminus[\![\emptyset,\mathrm{x}]\!]$, there is a unique point $s_i\in[\![\emptyset,\mathrm{x}]\!]$ (the Most Recent Common Ancestor, MRCA, of $\mathcal{X}_i$) such that

$$\bigcap_{x\in\mathcal{X}_i}[\![\emptyset,x]\!]=[\![\emptyset,s_i]\!].$$

For any $i\in I$, we will write $\mathcal{T}_i$ for the tree $\cup(\mathcal{X}_i\cup\{s_i\})$, reunion of all the trees having MRCA s_i, rooted at $s_i\in[\![\emptyset,\mathrm{x}]\!]$. Let $\mathcal{M}_{\mathrm{x}}=\sum_{i\in I}\delta_{(d(\emptyset,s_i),\mathcal{T}_i)}$ be the point measure with values in $\mathbf{R}_+\times\mathbb{T}$ describing the forest $\mathcal{F}^{\mathrm{x}}$ and the way it is attached to the spine. Spinal decompositions give the joint distribution of the spine $[\![\emptyset,\mathrm{x}]\!]$ and the measure $\mathcal{M}_{\mathrm{x}}$ for leaves selected in a particular way. Of course, keeping in mind that Lévy trees are constructed using excursions of real-valued random processes, most of these decompositions are actually classical path decompositions of said processes.

For instance, if the leaf x is selected according to mass measure $\mathbf{m}$, we easily see that this corresponds to the classical Bismut decomposition of Brownian motion ([Bis85]) with respect to a uniformly chosen point in the excursion interval $[0,\sigma]$. It should be stressed that the choice of a leaf according to mass measure introduces a mass-biasing of the Lévy tree.

Theorem 14 (Duquesne, Le Gall [DL05]). *Let ψ be a branching mechanism satisfying Assumptions 1 through 3. Then, if Φ is a measurable functional on $\mathbf{R}_+\times\mathbb{T}$,*

$$\mathbb{N}^{\psi}\left[\int_{\mathcal{T}}\mathbf{m}(dx)\exp\left(-\langle\mathcal{M}_{\mathrm{x}},\Phi\rangle\right)\right]=\int_0^{\infty}\mathrm{e}^{-\psi'(0)a}da\ \exp\left(-\int_0^a 2\beta\left(\mathbb{N}^{\psi}\left[1-\mathrm{e}^{-\Phi(t,\mathcal{T})}\right]\right)\right.$$
$$\left.+\int_0^{\infty}r\Pi(dr)\mathbb{P}^{\psi}_r\left[1-\mathrm{e}^{-\Phi(t,\mathcal{T})}\right]\ dt\right).$$

In other terms, on a spine $[\![\emptyset,s]\!]$ with length selected with density $\exp(-\psi'(0)a)da$ on $\mathbf{R}_+$, Lévy trees are grafted: trees distributed as $\mathbb{N}^{\psi}$ with intensity $2\beta\mathbf{1}_{[0,a]}(x)\ dx$ and trees distributed as $\mathbb{P}^{\psi}_r$ with intensity $\mathbf{1}_{[0,a]}(x)\ dx\otimes r\Pi(dr)\mathbf{1}_{(0,\infty)}(r)$. This decomposition can be extended to describe the distribution of the tree $\mathcal{T}$ conditionally on the subtree T_n spanned by the root and n uniform leaves ([Le 93b]).

When the tree is conditioned to have mass 1, the mass measure $\mathbf{m}$ is a probability measure—there is no more mass-biasing. However, the spinal decomposition is not as easy to describe in this case. It is known in the Brownian case that the length of the spine $[\![\emptyset,\mathrm{x}]\!]$ is Rayleigh-distributed. Furthermore, in the Brownian and stable cases, it is known ([HPW09]) that the spinal mass-distribution (the distribution of the vector $(\mathbf{m}(\mathscr{T}_i),\ i \in I)$, ranked in decreasing order) is the Poisson-Dirichlet distribution with parameters $(1-1/\alpha, 1-1/\alpha)$. One way of dealing with the conditioning is by disintegrating the unconditioned description of Theorem 14. This method will be used in chapter 4 for approximating quantities defined on $\mathscr{T}$ by their conditional expectations given T_n.

Another spinal decomposition arises when the leaf x is the highest leaf in the tree. This is analog to the classical Williams decomposition of the Brownian excursion with respect to its maximum ([Wil74]). By definition, the height of the highest leaf is distributed as the maximum of the height process on the excursion interval $[0,\sigma]$. This distribution is characterized by its "probability distribution function" $b(x) = \mathbb{N}^\psi[H(\mathscr{T}) \geq x]$, through the implicit equation:

$$\int_{b(x)}^{\infty} \frac{du}{\psi(u)} = x, \quad x > 0. \tag{5}$$

The spinal decomposition with respect to the spine $[\![\emptyset, \mathrm{x}]\!]$, where x is the highest leaf in the tree, is given by the following theorem:

Theorem 15 (Abraham, Delmas [AD09b]). *Let ψ be a subcritical or critical branching mechanism satisfying Assumptions 1 to 3. If $(\mathscr{T}, d, \emptyset, \mathbf{m})$ is a w-tree, let $\mathrm{x}_{\max} \in \mathscr{T}$ be such that $d(\emptyset, \mathrm{x}_{\max}) = \sup_{\mathrm{x}\in\mathscr{T}} d(\emptyset, \mathrm{x})$. Then, if Φ is a measurable functional on $\mathbf{R}_+ \times \mathbb{T}$,*

$$\begin{aligned}\mathbb{N}^\psi\left[\exp\left(-\langle \mathscr{M}_{\mathrm{x}_{\max}}, \Phi\rangle\right)\right] = \int_0^\infty dh\, \mu_{\max}(h) \\ \times \exp\Bigg(-\int_0^h ds\Big(2\beta\mathbb{N}^\psi[(1-\exp -\Phi(s,\mathscr{T}))\mathbf{1}_{\{H(\mathscr{T})\leq h-s\}}] \\ + \int_0^\infty r\mathrm{e}^{-rb(h-r)}\Pi(dr)\left[1-\mathrm{e}^{-r\mathbb{N}^\psi[(1-\exp -\Phi(s,\mathscr{T}))\mathbf{1}_{\{H(\mathscr{T})\leq h-s\}}]}\right]\Big)\Bigg),\end{aligned} \tag{6}$$

where $\mu_{\max}(h)$ is the density of $H(\mathscr{T})$ determined by (5).

In other words, the height of the spine is selected according to $\mu_{\max}(h)dh$, then Lévy trees are grafted on the spine, in such a way that at level $s \in [0,h]$, only Lévy trees restricted to $H(\mathscr{T}) \leq h-s$ are grafted. For an extension to non-homogeneous branching rates, see [DH11].

Tree-valued processes

The development of the topological theory of tree-spaces starting with [EPW05] made the definition and study of a number of tree-valued processes possible. We will review some of the existing processes, focusing mainly on the pruning process of Abraham-Delmas-Voisin, for which we will present a contribution by the author in the next section.

The definition of the root-growth with regrafting process of [EPW05] shows the strength of the tree-valued method: by defining the process directly on the general tree-space, the discrete and continuum trees are united in a same framework. The dynamic of this process can informally be described as follows: start with the trivial tree $\{\emptyset\}$ and with a set of cutting times $\tau_1 < \tau_2 < \ldots$ distributed as the jump times of a Poisson process with intensity $t\ dt$.

For $0 \le t < \tau_1$, define the tree T_t as just a spine $[\![\emptyset, t]\!]$ with length t. At time τ_1, select a vertex x_1 uniformly in the tree $[\![\emptyset, \tau_1]\!]$ and cut the tree at this vertex, giving two connected components $T^{(1)}_{\tau_1}$, containing the root, and $T^{(2)}_{\tau_1}$, not containing the root. Then, re-graft $T^{(2)}_{\tau_1}$ on the root. Continue the continuous growth of the root-edge in such a way that at time τ_2-, the tree T_{τ_2-} is a Y-shaped tree with the root-edge of length $\tau_2 - \tau_1$. Cut again at a uniform point and re-graft the bit not containing the root on the root. Continue growing the root-edge continuously between the cutting times and perform cutting with regrafting at the times τ_i. Then, as $t \to \infty$, the tree T_t converges a.s. to a Brownian CRT T_∞. Furthermore, this dynamic can be extended to infinite-length real trees (in which case cutting times are dense in $\mathbf{R}_+$). Then, by relating this growth procedure with the stick-breaking construction by Aldous ([Ald91a]), it can be shown that the distribution of the Brownian CRT is stationary for this dynamic, and, started from any rooted real tree T, the root-growth with re-grafting process will converge in distribution towards the Brownian CRT.

A somewhat related construction is the subtree-prune-and-regraft dynamic of [EW06]. In this case, at cutting times, the tree is cut at a uniform point in the skeleton (uniform meaning sampled according to length measure). Then, the bit not containing the root is grafted on a uniformly (according to mass measure) selected leaf. It is interesting to note that, in this paper, analytic methods such as Dirichlet forms were used on the tree-space in order to define and study this random process. This enabled to show, for instance, that the trivial tree is essentially polar for this dynamic, or that the distribution of the Brownian CRT is stationary.

The pruning procedure on Lévy trees was defined by Abraham, Delmas and Voisin ([ADV10]), in order to give a continuum analog to the pruning of Poisson Galton-Watson trees of [AP98b]. Furthermore, several results from the theory of branching processes ([AD09a, AD09b]) pointed to the fact that a CSBP with branching mechanism ψ could be constructed using another CSBP, with a "more subcritical" branching mechanism, with additional proportional immigration. Pruning procedures that could account for these observations had already been found in the purely Brownian case ([AS02, AP98a]) and in the case without Brownian part ([AD07]). We will present the pruning procedure on Lévy trees in some detail, as it will be central in this work. Note that we present only a particular case of the pruning procedure which leads to a tree-valued process. The general case can be found in [ADV10].

If ψ is a branching mechanism, satisfying Assumptions 1 and 2, we define, for $\theta \in \mathbf{R}$, the translated branching mechanism

$$\psi_\theta(u) = \psi(u+\theta) - \psi(\theta).$$

This transformation is known as the *Esscher transform*. Note that when ψ_θ is well-defined (it might not be, for instance $\psi(u) = u \log u$ for $\theta < 0$), ψ_θ is indeed a branching mechanism, and that Assumption 1 is still satisfied. However, it might very well be that ψ_θ is no longer conservative for negative θ. Therefore, we will introduce the set of admissible translations

$$\Theta^\psi = \left\{\theta \in \mathbf{R},\ \psi_\theta \text{ is conservative}\right\}.$$

To give a few examples, for a Brownian branching mechanism $\psi(u) = \beta u^2$, we see that $\psi_\theta(u) = \beta u^2 + 2\beta\theta u$, so that $\Theta^\psi = \mathbf{R}$. If $\psi(u) = u^\alpha$, the α-stable branching mechanism, then

$\Theta^\psi = [0,\infty)$. It is important to notice that, if ψ is for example critical ($\psi'(0) = 0$), then ψ_θ will be subcritical if $\theta > 0$ and supercritical if $\theta \in \Theta^\psi \cap (-\infty, 0)$. The main idea of [ADV10] and of [AD12a] is then to devise a cutting procedure on the Lévy tree with branching mechanism ψ so that the trimmed tree is distributed as a Lévy tree with branching mechanism ψ_θ with $\theta > 0$. So, let $\mathscr{T}$ be a ψ-Lévy tree, and, conditionally on $\mathscr{T}$, define

$$m^{(\mathrm{ske})}(dx, d\theta) = \sum_{i\in I^{\mathrm{ske}}} \delta_{(x_i,\theta_i)}(dx, d\theta)$$

a Poisson point measure on $\mathscr{T} \times \mathbf{R}_+$ with intensity $2\beta\ell^{\mathscr{T}}(dx)d\theta$. The atoms of this measure are distributed on the skeleton of $\mathscr{T}$. We also consider, conditionally on $\mathscr{T}$, a Poisson point measure

$$m^{(\mathrm{nod})}(dx, d\theta) = \sum_{i\in I^{\mathrm{nod}}} \delta_{(x_i,\theta_i)}(dx, d\theta)$$

on $\mathscr{T} \times \mathbf{R}_+$ with intensity

$$\sum_{y\in \mathrm{Br}_\infty(\mathscr{T})} \Delta_y \delta_y(dx)\, d\theta$$

where Δ_x is the width of the node x. Notice that the $\mathscr{T}$-components of the atoms of $m^{(\mathrm{nod})}$ are always infinitary nodes in $\mathscr{T}$. Moreover, if $\theta > 0$, a node $x \in \mathrm{Br}_\infty(\mathscr{T})$ is an atom of $m^{(\mathrm{nod})}(dx, [0,\theta])$ with probability $1 - \exp(-\theta\Delta_x)$. We then define the total measure of marks:

$$\mathscr{M}(dx, d\theta) = m^{(\mathrm{ske})}(dx, d\theta) + m^{(\mathrm{nod})}(dx, d\theta), \tag{7}$$

and consider the family of w-trees $(\mathscr{T}_\theta,\ \theta \ge 0)$, where the θ-pruned w-tree $\mathscr{T}_\theta$ is defined by:

$$\mathscr{T}_\theta = \{x \in \mathscr{T},\ \mathscr{M}([\![\emptyset, x[\![\times [0,\theta]) = 0\}.$$

In other words, $\mathscr{T}_\theta$ is the subtree of $\mathscr{T}$ containing all the points that have no mark of $\mathscr{M}(\cdot, [0,\theta])$ on the branch connecting them to the root. In particular, we have $\mathscr{T}_0 = \mathscr{T}$. The family of w-trees $(\mathscr{T}_\theta,\ \theta \ge 0)$ form a non-increasing process of real trees, in a sense that $\mathscr{T}_{\theta'} \supset \mathscr{T}_\theta$ for $0 \le \theta' \le \theta$. The following theorem is crucial in the understanding of the pruning process. If T is a tree, if the $(s_i, i \in I)$ are leaves of T, and if the $(T_i, i \in I)$ are trees, we write $T \circledast_{i\in I} (T_i, s_i)$ for the tree T with the T_i grafted on the s_i.

Theorem 16 (Abraham, Delmas, Voisin [ADV10]). *Let ψ be a branching mechanism satisfying Assumptions 1 to 3, let $\theta > 0$ and let $\mathscr{T}$ be distributed (under $\mathbb{P}^\psi_x$ or under $\mathbb{N}^\psi$) as a ψ-Lévy tree. Then, the pruned tree $\mathscr{T}_\theta$ is a Lévy tree with branching mechanism ψ_θ. Furthermore, conditionally on $\mathscr{T}_\theta$, the tree $\mathscr{T}$ can be written as*

$$\mathscr{T} = \mathscr{T}_\theta \circledast_{i\in I_\theta} (\mathscr{T}_i, s_i),$$

where $\sum_{i\in I_\theta} \delta_{(s_i,\mathscr{T}_i)}(ds, d\mathscr{T}')$ is a Poisson point measure on $\mathscr{T}_\theta \times \mathbb{T}$ with intensity

$$\mathbf{m}^{\mathscr{T}_\theta}(dx)\left(2\beta\theta\mathbb{N}^\psi[d\mathscr{T}'] + \int_{(0,+\infty)} \Pi(dr)\left(1 - \mathrm{e}^{-\theta r}\right)\mathbb{P}^\psi_r(d\mathscr{T}')\right),$$

and where $T \circledast (T', s)$ denotes the tree T with T' grafted on the leaf $s \in T$.

Hence, the process $(\mathcal{T}_\theta,\ \theta\geq 0)$ is a Lévy tree-valued process, such that, for any $\theta\geq 0$, $\mathcal{T}_\theta$ is distributed as a ψ_θ-Lévy tree. Another pruning process has been discovered by Curien and Haas ([CH12]): they show that the stable CRTs are nested, in the sense that if $1<\alpha<\alpha'\leq 2$, if $\mathcal{T}^\alpha$ is an α-stable CRT, then there exists a (rescaled) copy of an α'-stable CRT inside $\mathcal{T}^\alpha$. Their construction relies on clever pruning of nodes, and is not related to the general pruning of Lévy trees (note that if $\psi(u)=u^\alpha$ with $\alpha\in(1,2)$, then the translated branching mechanisms $(\psi_\theta(u),\ \theta\geq 0)$ never correspond to another stable branching mechanism).

The tree-valued pruning process has also been studied by Li ([Li12]) using stochastic differential equations. In this context, the tree-valued process $(\mathcal{T}_\theta,\ \theta\geq 0)$ is described using a path-valued process $(X_t(q),\ t\geq 0,\ q\in T)$ (where T is some interval depending on the branching mechanism), in which for any $q\in T$, $(X_t(q),\ t\geq 0)$ is a CSBP[10]. Then, if $Z_q(dt)=X_t(q)\ dt$ is the random measure defined by $X(q)$, the process $(Z_q,\ q\in T)$ is an inhomogeneous super-process with nonlocal branching.

Finally, it should be noted that not all continuum tree-valued processes are related to Lévy trees, unconditioned or conditioned. There is a growing literature on tree-valued processes in the context of population genetics, whether it be coalescent trees ([GPW08, PW06, PWW10]) or forward population models like the tree-valued Fleming-Viot process ([DGP11, GPW12]).

An alternative pathwise construction of the pruning process

In this paragraph, we shall summarize the results of Chapter 3, which contains the paper [ADH12c], submitted for publication.

The goal of this paper is to show how we can analyze the pathwise behavior of the pruning process by giving an alternative construction based on Poisson point measures on the tree-space. Note that the special Markov property of Theorem 16 only describes the finite-dimensional distributions of the pruning process, but does not give any information on its paths. We will define another Markov process having the same distribution as the pruning process $(\mathcal{T}_\theta,\theta\geq 0)$, whose paths will be easier to study. The main theorem is as follows:

Theorem 17 (Abraham, Delmas, H. [ADH12c]). *Let ψ be a branching mechanism satisfying Assumptions 1 to 3 and let $\theta>0$. Then there exists a $\mathbb{T}$-valued process $(\mathfrak{T}_q,\ q\in\Theta^\psi\cap(-\infty,\theta])$ that has the same distribution as the stopped pruning process $(\mathcal{T}_q,\ q\in\Theta^\psi\cap(-\infty,\theta])$, and such that its jump measure*

$$\sum_{i\in I}\delta_{(x_i,\Delta\mathfrak{T}_i,q_i)}(dx,d\mathfrak{T},dq)\mathbf{1}_{\mathfrak{T}_{q_i-}\neq\mathfrak{T}_{q_i}}(q)$$

has a backward predictable projection $\mathbf{m}^{\mathfrak{T}_q}(dx)\ \mathbf{N}^{\psi_q}[d\mathfrak{T}]\ \mathbf{1}_{\Theta^\psi\cap(-\infty,\theta]}(q)\ dq$, where $\mathbf{N}^{\psi_q}$ is the σ-finite measure on $\mathbb{T}$ defined by:

$$\mathbf{N}^{\psi_q}[d\mathfrak{T}]=2\beta\mathbb{N}^{\psi_q}[d\mathfrak{T}]+\int_{(0,\infty)}\Pi(dr)re^{-qr}\mathbb{P}_r^{\psi_q}(d\mathfrak{T})$$

In other words, at each jumping time q of the process, a tree $\Delta\mathfrak{T}_q$ distributed as $\mathbf{N}^{\psi_q}[d\mathcal{T}]$ is grafted on a uniformly selected leaf $x_q\in\mathfrak{T}_q$, so that $\mathfrak{T}_{q-}=\mathfrak{T}_q\circledast(\Delta\mathfrak{T}_q,x_q)$. We can write

[10] Actually, Li describes the process "backwards in time", so that for any $t\geq 0$, we have $X_t(q)\leq X_t(q')$ if $q\leq q'$.

down, at least formally, the infinitesimal generator $\mathscr{L}_\theta$ of the (inhomogeneous) growth process $(\widehat{\mathscr{T}_\theta} = \mathscr{T}_{-\theta},\ \theta \in -\Theta^\psi)$: if $\theta \in -\Theta^\psi$ and if F is a measurable, bounded functional on $\mathbb{T}$, for all w-tree t, we have

$$(\mathscr{L}_\theta F)(\mathrm{t}) = \int_{\mathrm{t}} \mathbf{m}^{\mathrm{t}}(ds) \int_{\mathbb{T}} \mathbf{N}^{\psi_\theta}[dT]\,(F(\mathrm{t} \circledast (T,s)) - F(\mathrm{t}))\,.$$

This theorem is proved by iterating the special Markov property of the pruning process (Theorem 16) recursively, which incidentally reveals an interesting generational decomposition of Lévy trees. We also give an application of this pathwise construction by studying special jump times of the pruning process. For instance, let

$$A_h = \sup\left\{\theta \in \Theta^\psi,\ H(\mathscr{T}_\theta) > h\right\}, \quad 0 < h \le \infty$$

be the ***exit time*** out of the domain $\{\mathscr{T},\ H(\mathscr{T}) \le h\}$. In the paper [AD12a], the time A_∞ was studied. In particular, the tree at time A_∞, which is a.s. finite, was related to a Lévy tree conditioned to survive, suitably pruned. We substantially extend these results, not only by considering all the cases $0 < h \le \infty$, but also by describing precisely the jump $\mathscr{T}_{A_h-} \rightsquigarrow \mathscr{T}_{A_h}$. Let us introduce the notation

$$b_h^\theta(s) = \mathbb{N}^{\psi_\theta}[H(\mathscr{T}) > h - s], \quad 0 \le s < h.$$

so that $\mathbb{N}^\psi[A_h > \theta] = \mathbb{N}^\psi[H(\mathscr{T}_\theta) > h] = b_h^\theta(0)$. The function $(b_h^\theta(s),\ 0 \le s < h)$ is uniquely determined by the implicit equation

$$\int_{b_h^\theta(s)}^\infty \frac{du}{\psi_\theta(u)} = h - s. \tag{8}$$

Notice that when $h = \infty$, we get $b_\infty^\theta(s) = \psi_\theta^{-1}(0)$ for all s. We manage to derive the joint distribution of $(\mathscr{T}_{A_h-}, \mathscr{T}_{A_h})$ under the form of a spinal decomposition with respect to the spine $[\![\emptyset, \mathrm{x}]\!]$, where x is the point on which the tree $\Delta\mathfrak{T}_{A_h}$ is grafted. Note that x is ***not*** the highest leaf in the tree $\mathscr{T}_{A_h}$. Contrasting with the Williams decomposition (Theorem 15), the spinal subtrees are no longer uniformly grafted along the spine, but with a density skewed by the potential $b_h^{A_h}$.

We will state the result for the purely quadratic case $\psi(u) = \beta u^2$; we refer to Theorem 3.33 in Chapter 3 for the complete description. In the quadratic case, equation (8) can be explicitly solved:

$$b_h^\theta(s) = \frac{2\theta}{\mathrm{e}^{2\beta\theta(h-s)} - 1}, \quad 0 \le s < h,$$

if $h < \infty$ and $b_\infty^\theta(s) = \psi_\theta^{-1}(0) = 2|\theta|$ for $\theta < 0$. Then we can prove that, conditionally on $\{A_h = \theta\}$, the height of x is distributed with the following density:

$$\mathbb{N}^\psi[d(\emptyset, \mathrm{x}) \in dx] \propto b_h^\theta(x) \exp\left(-|\theta| x - \int_0^x b_h^\theta(u)\, du\right) dx, \quad x \in [0, h).$$

Furthermore, conditionally on $\{A_h = \theta\}$ and $\{d(\emptyset, \mathrm{x}) = x\}$, the point process $\sum_{i\in I} \delta_{(d(\emptyset,s_i),\mathcal{T}_i)}$ recording the spinal decomposition with respect to $[\![\emptyset, \mathrm{x}]\!]$ is distributed as a Poisson point process on $[0,x)\times\mathbb{T}$ with intensity

$$2\beta\left(\theta + b_h^\theta(u)\right)\mathbf{1}_{[0,x)}(u)du\ \mathbb{N}^{\psi_\theta}[d\mathcal{T}, H(\mathcal{T}) < h-u].$$

In other words, seeing as how $b_h^\theta(u)$ grows with u, many trees are grafted closer to x, but they tend to be smaller than the trees grafted near the root, of which there are less. A particular case of these results is $h=\infty$, in which case we extend results from [AD12a]. In particular, conditionally on $A_\infty = \theta < 0$ and on $\mathcal{T}_{A_\infty} = \mathcal{T}$, the tree $\mathcal{T}_{A_\infty -}$ is distributed as $\mathcal{T}_{A_\infty} \circledast (T_\infty, s)$, where s is a uniform leaf of $\mathcal{T}_{A_\infty}$ and where T_∞ is distributed as a ψ_θ-Lévy tree restricted on having infinite mass. Also, in the quadratic case, we can derive the following explicit formula for the probability that the jump at A_h is infinite, that is, the probability that $A_h = A_\infty$: for $h>0$ and $\theta<0$,

$$\mathbb{N}^\psi[A_h = A_\infty | A_\infty = \theta] = \frac{\beta\theta h}{\sinh^2(\beta\theta h)} - \mathrm{cotanh}(\beta\theta h)$$

1.3 Fragmentations of continuum random trees

Self-similar fragmentations of continuum random trees

Self-similar fragmentation processes were introduced by Bertoin ([Ber01, Ber02]). They are Markov processes with values in the set $\mathcal{S}^\downarrow$ of nonincreasing nonnegative sequences ($s_1 \geq s_2 \geq \dots$) such that $\sum s_i \leq 1$. The s_i represent the "size" of fragments, whatever the underlying model. Self-similar fragmentations are characterized by a triple (α, c, ν), where $\alpha\in\mathbf{R}$ is the self-similarity index, $c\geq 0$ is the erosion rate and $\nu(d\mathbf{s})$ is a σ-finite measure on $\mathcal{S}^\downarrow$ called the *dislocation measure*. Informally, a fragment of size x will dislocate into fragments of size $x\mathbf{s}$ at a rate $x^\alpha\nu(d\mathbf{s})$, independently of all other fragments. Hence, if $\alpha<0$, the smaller fragments dislocate faster than the big ones, whereas if $\alpha>0$, the big fragments dislocate faster. If $\alpha=0$, the fragmentation rates are independent of fragment size. There can also be a continuous *erosion* of the fragments, but the fragmentations we consider are not affected by that phenomenon. We refer to the monograph [Ber06] for details on self-similar fragmentation processes.

In their paper [AP98a], Aldous and Pitman considered a specific fragmentation of the Brownian CRT. Given a rooted CRT $(\mathcal{T}, \emptyset)$, they consider a Poisson point measure $\mathcal{M}$ on $\mathcal{T}\times\mathbf{R}_+$ with intensity $\ell(ds)\otimes dt$. Then, if $t\geq 0$, they consider the random forest $\mathcal{F}_t = (\mathcal{T}_i(t),\ i\in I_t)$ of connected components of $\mathcal{T}$ defined by the marks of $\mathcal{M}(\cdot\times[0,t])$. This is very similar to the pruning procedure on Lévy trees, except that all connected components are kept, instead of just focusing on the one containing the root. It turns out that, in the language later framed by Bertoin, the mass-partition process $(\mathbf{m}(\mathcal{T}_i(t)),\ i\in I_t)^\downarrow$, ranked in decreasing order, is a self-similar fragmentation of index $1/2$, with no erosion and with a dislocation measure defined by

$$\nu_{AP}(\{s_1+s_2<1\}) = 0, \quad \nu_{AP}(s_1\in dx) = \frac{dx}{\sqrt{2\pi x^3(1-x)^3}}, \quad x\in[1/2,1).$$

Notice that the second condition implies that all the fragmentation events are *binary*: fragments always get shattered in exactly two pieces. In this very particular case described by Aldous and Pitman, the time-reversal of the fragmentation process is a coalescent process, the so-called *standard additive coalescent*. If m_t is, at time $t \geq 0$, the mass of the fragment containing the root, then the process $(m_t,\ t \geq 0)$ is equal in distribution to the process $(1/(1+\tau_t),\ t \geq 0)$, where τ is a stable subordinator with index $1/2$. This result can be recovered from the pruning of a Lévy tree with branching mechanism $\psi(u) = u^2/2$, conditioned on $\{\sigma = 1\}$.

The fragmentation properties of stable trees were studied by Miermont using two different fragmentation procedures. First, the *fragmentation at heights* ([Mie03]), in which, at time $t \geq 0$, all vertices s of the tree with $d(\emptyset, s) < t$ are discarded. The remaining connected components form a random forest $(\mathcal{T}_i(t),\ i \in I_t)$ such that the associated mass process $((\mathbf{m}(\mathcal{T}_i(t)),\ i \in I_t),\ t \geq 0)$ is a self-similar fragmentation process with index $1/\alpha - 1 \in (-1/2, 0)$, and no erosion. The dislocation measure of the fragmentation is equal to $D_\alpha \nu_\alpha$, where D_α is a deterministic constant, and where ν_α is defined by

$$\int_{\mathscr{S}^\downarrow} G(\mathbf{s})\ \nu_\alpha(d\mathbf{s}) = \mathbb{E}\left[S_1 G\left(S_1^{-1} \Delta S_{[0,1]}\right)\right],$$

for any measurable bounded functional G, where S is a stable subordinator with index $1/\alpha$. Note that since the index of the fragmentation is negative, loss of mass occurs even though there is no erosion and the dislocation measure is conservative.

A second way of fragmenting the α-stable tree, which is in some sense the analog of Aldous and Pitman's fragmentation is the *fragmentation at nodes* ([Mie04]), where the nodes of the tree are cut independently as time goes along: if Δ_x is the width of node x, then it gets cut at an exponential time with parameter Δ_x. When a cutting occurs, the tree gets fragmented into infinitely many bits, since there are infinitely many connected components attached to the node. In this case, the fragmentation is self-similar, with index $1/\alpha \in (1/2, 1)$ and no erosion. Surprisingly, the erosion measure is the same as for the fragmentation at heights, generalizing a result by Bertoin ([Ber02]) for the Brownian CRT. We refer to the original papers for a detailed discussion.

Self-similar fragmentations with negative index can also be described by means of a continuum random tree, the so-called *fragmentation tree*, discovered by Haas and Miermont ([HM04]). The general idea is to represent a self-similar fragmentation with negative index as the fragmentation at heights of a certain continuum random tree. More precisely, given a triple (α, c, ν) characterizing a self-similar fragmentation $(F(t), t \geq 0)$, the authors construct a rooted weighted continuum random tree $(\mathcal{T}_{(\alpha,c,\nu)}, \emptyset, \mu)$ such that its mass-fragmentation process at heights $((\mu(\mathcal{T}_i(t))^\downarrow,\ i \in I_t),\ t \geq 0)$, constructed as above, has the same distribution as F. The results of Bertoin and Miermont then show that the Brownian CRT and the α-stable CRT with $\alpha \in (1,2)$ are both examples of fragmentation trees, the Brownian tree being associated with self-similarity index $\alpha = -1/2$, erosion $c = 0$ and (binary) dislocation measure

$$\nu(s_1 \in dx) = \nu_{AP}(s_1 \in dx).$$

We already mentioned the fact that self-similar fragmentation trees are the scaling limits of a very general class of discrete trees, the Markov branching trees ([HMPW08, HM12]).

Edge-cutting procedures: discrete and continuum case

The edge-cutting problem, first studied by Meir and Moon ([MM70]), is as follows: given a discrete rooted tree $(T_n, \emptyset)$ with n unit-length edges, select one edge e uniformly at random. Remove e, as well as the connected component adjacent to e that doesn't contain the root. On the remaining tree, iterate this procedure. The (random) number of edge-removals needed to isolate the root will be noted $X(T_n)$. The question of the asymptotic behavior of $X(T_n)$ when $n \to \infty$ is not trivial even for very simple trees: for instance, if $T_n = [\![0, n]\!]$, the line-tree with n unit-length edges, we have $\mathbb{E}[X(T_n)] \sim \log n$, and by an application of the Lindeberg-Feller theorem, we get that $(\log n)^{-1/2}(X(T_n) - \log n)$ converges in distribution to a standard Gaussian r.v.

When T_n is randomly chosen, the asymptotic very much depends on the shape of the tree, the line-tree $[\![0, n]\!]$ and the star-shaped tree being extreme cases. The original paper by Meir and Moon examined the case where T_n is a Cayley tree (a uniform rooted labeled tree with $n+1$ vertices) and proved that

$$\mathbb{E}[X(T_n)] \sim \sqrt{\frac{\pi n}{2}}; \quad \mathrm{Var}(X(T_n)) \sim \left(2 - \frac{\pi}{2}\right) n.$$

Later, Panholzer ([Pan06]) and Janson ([Jan06]) proved that for any critical, finite-variance Galton-Watson tree conditioned on having n edges, we have a convergence in distribution

$$\lim_{n\to\infty} \frac{X(T_n)}{\sigma\sqrt{n}} = \mathscr{R}, \tag{9}$$

where σ^2 is the variance of the offspring distribution of T_n and where $\mathscr{R}$ has Rayleigh distribution $\mathbb{P}(\mathscr{R} \in dx) = x \exp(-x^2/2) \mathbf{1}_{(0,\infty)}(x)\, dx$. The proof of Janson relied strongly on Aldous's result (Theorem 1), and he actually proved that if μ_{T_n} denotes the conditional distribution of $X(T_n)/(\sigma\sqrt{n})$ given the rescaled tree $\sigma T_n/\sqrt{n}$, then we have

$$\lim_{n\to\infty} \mu_{T_n} = \mu_{2B} \tag{10}$$

where μ_{2B} is a random probability measure, specified by a moment problem conditionally on the Brownian excursion $2B$ arising as the limit of the rescaled contour process. Of course, when averaging over $2B$, one recovers the Rayleigh-distributed limit of (9). Since then, a number of different approaches have been used to describe cutting procedures on the Brownian CRT that could account for either the limiting distribution in (9) or the conditional distribution μ_{2B} in (10).

Using a reconstruction procedure on the CRT inspired by the classical Aldous-Broder algorithm, Addario-Berry, Broutin and Holmgren ([ABBH11]) extended the Aldous-Pitman fragmentation by discarding all fragmentation events that didn't affect the root-fragment. However, at each fragmentation time t a fragment $\mathscr{T}_t$ is separated from the root-fragment, it is re-attached to a continuously growing spine $[\![\emptyset, L(t)]\!]$, where the process $(L(t), t \geq 0)$ is a local time constructed from the fragmentation. When $t \to \infty$, the root-fragment reduces to $\{\emptyset\}$ and all the fragments will have been attached on the spine $[\![\emptyset, L(\infty)]\!]$. Then, $L(\infty)$ is

Rayleigh-distributed and this distribution is related to the edge-cutting in discrete (Cayley) trees, for which a similar reconstruction procedure can be defined.

Another approach was used by Bertoin and Miermont ([BM12]) in the case of critical, finite-variance Galton-Watson trees, keeping track of all the fragments in the edge-removal procedure in the discrete tree through the construction of another discrete tree $\mathrm{cut}(T_n)$. In the same spirit, one can construct a continuum tree $\mathrm{cut}(\mathcal{T})$ to keep track of the genealogy of the fragments in the Aldous-Pitman fragmentation. Then, on the one hand, it is proven that if $\mathcal{T}$ is a Brownian CRT, then $\mathrm{cut}(\mathcal{T})$ is also distributed as a Brownian CRT. On the other hand,

$$\lim_{n\to\infty}\left(\frac{\sigma T_n}{\sqrt{n}},\frac{\mathrm{cut}(T_n)}{\sigma\sqrt{n}}\right)=(\mathcal{T},\mathrm{cut}(\mathcal{T})),$$

in distribution for the Gromov-weak topology. This explains Janson's result, since $X(T_n)$ is exactly distributed as the height of a uniform leaf in $\mathrm{cut}(T_n)$, and since the height of a uniform leaf is continuous with respect to the Gromov-weak topology.

We shall use yet another way of explaining the Rayleigh distribution in the limit of (9), discovered by Abraham and Delmas ([AD11]). The idea is to look at the effect of the Aldous-Pitman fragmentation on the subtrees T_n of the Brownian CRT $\mathcal{T}$ spanned by n uniformly sampled leaves and the root. For any $s\in\mathcal{T}$, note $\theta(s)$ the time at which s gets separated from the root in the Aldous-Pitman fragmentation. We use the notation $\mathbb{P}^{(1)}_\infty$ for the distribution of $(\mathcal{T},\theta)$, reflecting the fact that $\theta(\emptyset)=\infty$. The process θ can also be defined such that $\theta(\emptyset)=q\in(0,\infty)$, for which we will use the notation $\mathbb{P}^{(1)}_q$.

The analog of $X(T_n)$ will be the quantity X_n^* of jumps of the process θ on T_n lying above the first branching point $s_{\emptyset,n}$. Indeed, after the first mark appears on the edge $[\![\emptyset,s_{\emptyset,n}]\!]$, the root will be separated from all other vertices in T_n, just as $X(T_n)$ was the number of cuts needed to isolate the root from all other vertices in Janson's edge-cutting process. The main result is then

Theorem 18 (Abraham, Delmas [AD11]). *When $n\to\infty$, we have $\mathbb{P}^{(1)}_\infty$-a.s.*

$$\lim_{n\to\infty}\frac{X_n^*}{\sqrt{2n}}=\Theta, \tag{11}$$

where we noted $\Theta=\int_{\mathcal{T}}\theta(s)\mathbf{m}(ds)$. Furthermore, the limiting r.v. Θ is Rayleigh-distributed under $\mathbb{P}^{(1)}_\infty$.

Note that the normalizing factor $\sqrt{2n}$ comes from the fact that T_n has n leaves, thus $2n-1$ edges, whereas the trees T_n from [Jan06] were selected among trees with n edges. Hence, the result is quite the same. It is also noteworthy that, conditionally on $\mathcal{T}=\mathcal{T}_{2B}$, the limiting r.v. actually has the distribution μ_{2B} predicted by Janson's result (10).

Finally, let us mention other results about the asymptotic behavior of $X(T_n)$ for different classes of trees. If T_n is a *random recursive tree* with n edges, it is proven in [DIMR08] that

$$\lim_{n\to\infty}\frac{X(T_n)-\frac{n}{\log n}-\frac{n\log\log n}{\log^2 n}}{\frac{n}{\log^2 n}}=Z,$$

in distribution, where Z is a 1-stable random variable[11]. A similar result was proven in [Hol11] for the much larger class of *split trees*, containing for instance binary search trees, quad trees, tries, digital search trees... What is important is that the random recursive trees and split trees belong to the class of so-called "$\log n$" trees, meaning that their height is asymptotic to $\log n$, whereas the conditioned, finite-variance Galton-Watson trees belong to the "$\sqrt{n}$" class.

Fluctuations for the record process on the Brownian tree

In this paragraph, we shall summarize the results of Chapter 4, which contains the paper [Hos12], submitted for publication.

The paper builds on the almost sure result of Abraham-Delmas (Theorem 18), and gives an asymptotic result for the fluctuations of the number of records $X_n^*/\sqrt{2n}$ around its asymptotic value $\Theta = \int_{\mathcal{T}} \theta(s)\ \mathbf{m}(ds)$.

Theorem 19 ([Hos12]). *Under $\mathbb{N}_\infty^{(1)}$, we have the following convergence in distribution:*

$$\lim_{n\to\infty} n^{1/4}\left(\frac{X_n^*}{\sqrt{2n}} - \Theta\right) = Z, \tag{12}$$

where Z is a random variable with characteristic function $\mathbb{E}_\infty^{(1)}\left[e^{itZ}\right] = \mathbb{E}_\infty^{(1)}\left[e^{-t^2\Theta/\sqrt{2}}\right]$.

In other words, Z has the distribution of $\sqrt{\Theta}G$, with G an independent standard normal r.v. To prove this theorem, the main idea is to introduce martingales indexed by the tree $\mathcal{T}$ and to use martingale limit theory to find a nontrivial limit in distribution. Indeed, recall the definition of the Aldous-Pitman fragmentation by means of a Poisson point process $\mathcal{M}$. If we thin this point process by keeping only the marks that led to an actual separation of a fragment from the root-fragment, *i.e.* the jumps of θ, we get a point process $\mathcal{M}_{\text{rec}}$ that is no longer Poisson, but it can be shown that it has intensity $\theta(s)\ell(ds)$ on $\mathcal{T}$. Hence, by using standard point process theory, when looking at a single branch $[\![x,y]\!] \subset \mathcal{T}$ (not containing the root), we get that the number $X([\![x,y]\!])$ of jumps on $[\![x,y]\!]$ satisfies:

$$\mathbb{E}_\infty^{(1)}\left[X([\![x,y]\!])\right] = \mathbb{E}_\infty^{(1)}\left[\int_{[\![x,y]\!]} \theta(s)\ \ell(ds)\right]$$

and that

$$\mathbb{E}_\infty^{(1)}\left[\left(X([\![x,y]\!]) - \int_{[\![x,y]\!]} \theta(s)\ \ell(ds)\right)^2\right] = \mathbb{E}_\infty^{(1)}\left[\int_{[\![x,y]\!]} \theta(s)\ \ell(ds)\right].$$

Since $X_n^* = X(\mathrm{T}_n^*)$, this enables to show that

$$\lim_{n\to\infty} n^{1/4}\left(\frac{X_n^*}{\sqrt{2n}} - \int_{\mathrm{T}_n^*} \theta(s)\ \frac{\ell(ds)}{\ell(T_n^*)}\right) = Z,$$

where, conditionally on Θ, Z is distributed as a centered Gaussian variable with variance Θ. Theorem 19 then follows from the fact that $n^{1/4}(\int_{\mathrm{T}_n^*} \theta(s)\ell(ds)/\ell(\mathrm{T}_n^*) - \Theta)$ converges to 0,

[11] Specified by its characteristic function $\mathbb{E}[\exp(itZ)] = \exp(it\log|t| - \pi|t|/2)$.

which we prove using a general disintegration theorem: if $\sum_{i\in I_n} \delta_{(\mathcal{T}_i, s_i)}$ is the point measure describing the decomposition of $\mathcal{T}$ with respect to T_n, and if $\mathscr{F}_n$ is the σ-algebra generated by T_n and by $(\theta(s),\ s \in \mathrm{T}_n)$, then

Theorem 20 ([Hos12]). *Let F be a nonnegative functional on $\mathbb{T} \times \mathcal{T}$. Then*

$$\mathbb{E}_\infty^{(1)}\left[\sum_{i\in I_n} F(\mathcal{T}_i, s_i)\middle|\mathscr{F}_n\right] = \int_0^1 \frac{\mathrm{e}^{-L_n^2 v/(2-2v)}}{\sqrt{2\pi}\, v^{3/2}(1-v)^{3/2}}\, dv \int_{\mathrm{T}_n} \ell(ds)\mathbb{E}_{\theta(s)}^{(v)}[F(\mathcal{T}, s)].$$

This proof of this theorem uses the same disintegration methods already used by Le Gall ([Le 93b]) to derive the distribution of T_n under $\mathbb{N}^{(1)}$ from its distribution under $\mathbb{N}$. In our context, this formula enables to estimate quite tightly the rate of convergence of $(\mathbb{E}_\infty^{(1)}[\Theta|\mathscr{F}_n] - \Theta)$ to 0, which entails Theorem 19.

CHAPTER 2

Gromov-Hausdorff-Prokhorov convergence of locally compact metric measure spaces

2.1 Introduction

In the present work, we aim to give a topological framework to certain classes of measured metric spaces. The methods go back to ideas from Gromov [Gro07], who first considered the so-called Gromov-Hausdorff metric in order to compare metric spaces who might not be subspaces of a common metric space. The classical theory of the Gromov-Hausdorff metric on the space of compact metric spaces, as well as its extension to locally compact spaces, is exposed in particular in Burago, Burago and Ivanov [BBI01].

Recently, the concept of Gromov-Hausdorff convergence has found striking applications in the field of probability theory, in the context of random graphs. Evans [Eva08] and Evans, Pitman and Winter [EPW05] considered the space of real trees, which is Polish when endowed with the Gromov-Hausdorff metric. This has given a framework to the theory of continuum random trees, which originated with Aldous [Ald91a]. In the monograph by Evans [Eva08], the author describes a topology on the space of compact real trees, equipped with a probability measure, using the Prokhorov metric to compare the measures, thus defining the so-called weighted Gromov-Hausdorff metric. Recently Greven, Pfaffelhuber and Winter [GPW08] take another approach by considering the space of complete, separable metric spaces, endowed with probability measures (metric measure spaces). In order to compare two such probability spaces, they consider embeddings of both these spaces into some common Polish metric space, and use the Prokhorov metric to compare the ensuing measures. This puts the emphasis on the probability measure carried by the space rather than its geometrical features. In his monograph, Villani [Vil09] gives an account of the theory of measured metric spaces and the different approaches to their topology. This theory, as well as the connection to mass transportation problems, was also developed in [Stu06a, Stu06b]. Miermont, in [Mie09], describes a combined approach, using both the Hausdorff metric and the Prokhorov metric to compare compact metric spaces equipped with probability measures. The metric he uses (called the Gromov-Hausdorff-Prokhorov metric) is not the same as Evans's, but they are shown to give rise to the same topology. This topology is stronger than the Gromov-Prokhorov topology, since it puts a stronger emphasis on the geometrical features of the spaces.

In the present paper, we describe several properties of the Gromov-Hausdorff-Prokhorov metric, d^c_{GHP}, on the set $\mathbb{K}$ of (isometry classes of) compact metric spaces, with a distinguished element called the root and endowed with a finite measure. Theorem 2.3 ensures that $(\mathbb{K}, d^c_{\mathrm{GHP}})$ is a Polish metric space. We extend those results by considering the Gromov-Hausdorff-Prokhorov metric, d_{GHP}, on the set $\mathbb{L}$ of (isometry classes of) rooted locally compact, complete length spaces, endowed with a boundedly finite measure. Theorem 2.7 ensures that $(\mathbb{L}, d_{\mathrm{GHP}})$ is also a Polish metric space. The proof of the completeness of $\mathbb{L}$ relies on a precompactness criterion given in Theorem 2.9. The methods used are similar to the methods used in [BBI01] to derive properties about the Gromov-Hausdorff topology of the set of locally compact complete length spaces. This work extends some of the results from [GPW08], which doesn't take into account the geometrical structure of the spaces, as well as the results from [Mie09], which consider only the compact case and probability measures. This comes at the price of having to restrict ourselves to the context of length spaces. In [Vil09] the Gromov-Hausdorff-Prokhorov topology is considered for general Polish spaces (instead of length spaces) but endowed with boundedly finite measures satisfying the doubling condition. We also mention the different approach of [ABBH11], using the ideas of correspondences between metric spaces and couplings of measures.

This work was developed for applications in the setting of weighted Lévy trees (which are elements of $\mathbb{L}$), see Abraham, Delmas and Hoscheit [ADH12c]. We give an hint of those applications by stating that the construction of a weighted tree coded in a continuous function with compact support is measurable with respect to the topology induced by d^c_{GHP} on $\mathbb{K}$ or by d_{GHP} on $\mathbb{L}$. This construction allows us to define random variables on $\mathbb{K}$ using continuous random processes on $\mathbf{R}$, in particular the Lévy trees of [DL05] that describe the genealogy of the so-called critical or sub-critical continuous state branching processes that become a.s. extinct. The measure $\mathbf{m}$ is then a "uniform" measure on the leaves of the tree which has finite mass. The construction can be generalized to super-critical continuous state branching processes which can live forever; in that case the corresponding genealogical tree is infinite and the measure $\mathbf{m}$ on the leaves is also infinite. This paper gives an appropriate framework to handle such tree-valued random variables and also tree-valued Markov processes as in [ADH12c].

The structure of the paper is as follow. Section 2.2 collects the main results of the paper. The application to real trees is given in Section 2.3. The proofs of the results in the compact case are given in Section 2.4. The proofs of the results in the locally compact case are given in Section 2.5.

2.2 Main results

Rooted weighted metric spaces

Let (X, d^X) be a Polish metric space. The diameter of $A \in \mathscr{B}(X)$ is given by:

$$\mathrm{diam}(A) = \sup\left\{d^X(x,y);\ x, y \in A\right\}.$$

For $A, B \in \mathscr{B}(X)$, we set:

$$d_{\mathrm{H}}^X(A,B) = \inf\{\varepsilon > 0;\ A \subset B^\varepsilon \text{ and } B \subset A^\varepsilon\},$$

the Hausdorff metric between A and B, where

$$A^\varepsilon = \left\{x \in X;\ \inf_{y \in A} d^X(x,y) < \varepsilon\right\} \tag{2.1}$$

is the ε-halo set of A. If X is compact, then the space of compact subsets of X, endowed with the Hausdorff metric, is compact, see theorem 7.3.8 in [BBI01]. To give pre-compactness criterion, we will need the notion of ε-nets.

Definition 2.1. *Let (X, d^X) be a metric space, and let $\varepsilon > 0$. A subset $A \subset X$ is called an ε-net of $B \subset X$ if:*

$$A \subset B \subset A^\varepsilon.$$

Notice that, for any $\varepsilon > 0$, compact metric spaces admit finite ε-nets and locally compact spaces admit boundedly finite ε-nets. Let $\mathscr{M}_f(X)$ denote the set of all finite Borel measures on X. If $\mu, \nu \in \mathscr{M}_f(X)$, we set:

$$d_{\mathrm{P}}^X(\mu,\nu) = \inf\{\varepsilon > 0;\ \mu(A) \le \nu(A^\varepsilon) + \varepsilon \text{ and } \nu(A) \le \mu(A^\varepsilon) + \varepsilon \quad \text{for any closed set } A\},$$

the Prokhorov metric between μ and ν. It is well known (see [DVJ03] Appendix A.2.5) that $(\mathscr{M}_f(X), d_{\mathrm{P}}^X)$ is a Polish metric space, and that the topology generated by d_{P}^X is exactly the topology of weak convergence (convergence against continuous bounded functionals).

The Prokhorov metric can be extended in the following way. Recall that a Borel measure is boundedly finite if the measure of any bounded Borel set is finite. Let $\mathscr{M}(X)$ denote the set of all boundedly finite Borel measures on X. Let $\emptyset$ be a distinguished element of X, which we will call the root. We will consider the closed ball of radius r centered at $\emptyset$:

$$X^{(r)} = \{x \in X; d^X(\emptyset, x) \le r\}, \tag{2.2}$$

and for $\mu \in \mathscr{M}(X)$ its restriction $\mu^{(r)}$ to $X^{(r)}$:

$$\mu^{(r)}(dx) = \mathbf{1}_{X^{(r)}}(x)\, \mu(dx). \tag{2.3}$$

If $\mu, \nu \in \mathscr{M}(X)$, we define a generalized Prokhorov metric between μ and ν:

$$d_{\mathrm{gP}}^X(\mu,\nu) = \int_0^\infty \mathrm{e}^{-r}\left(1 \wedge d_{\mathrm{P}}^X\left(\mu^{(r)}, \nu^{(r)}\right)\right)\, dr. \tag{2.4}$$

It is not difficult to check that d_{gP}^X is well defined (see Lemma 2.6 in a more general framework) and is a metric. Furthermore $(\mathscr{M}(X), d_{\mathrm{gP}}^X)$ is a Polish metric space, and the topology generated by d_{gP}^X is exactly the topology of vague convergence (convergence against continuous bounded functionals with bounded support), see [DVJ03] Appendix A.2.6.

When there is no ambiguity on the metric space (X, d^X), we may write d, d_{H}, and d_{P} instead of d^X, d_{H}^X and d_{P}^X. In the case where we consider different metrics on the same space, in order to stress that the metric is d^X, we will write $d_{\mathrm{H}}^{d^X}$ and $d_{\mathrm{P}}^{d^X}$ for the corresponding Hausdorff and Prokhorov metrics.

If $\Phi : X \to X'$ is a Borel map between two Polish metric spaces and if μ is a Borel measure on X, we will note $\Phi_*\mu$ the image measure on X' defined by $\Phi_*\mu(A) = \mu(\Phi^{-1}(A))$, for any Borel set $A \subset X$.

Definition 2.2.

- *A rooted weighted metric space $\mathscr{X} = (X, d, \emptyset, \mu)$ is a metric space (X, d) with a distinguished element $\emptyset \in X$, called the root, and a boundedly finite Borel measure μ.*
- *Two rooted weighted metric spaces $\mathscr{X} = (X, d, \emptyset, \mu)$ and $\mathscr{X}' = (X', d', \emptyset', \mu')$ are said to be GHP-isometric if there exists an isometric one-to-one map $\Phi : X \to X'$ such that $\Phi(\emptyset) = \emptyset'$ and $\Phi_*\mu = \mu'$. In that case, Φ is called a GHP-isometry.*

Notice that if (X, d) is compact, then a boundedly finite measure on X is finite and belongs to $\mathscr{M}_f(X)$. We will now use a procedure due to Gromov [Gro07] to compare any two compact rooted weighted metric spaces, even if they are not subspaces of the same Polish metric space.

Gromov-Hausdorff-Prokhorov metric for compact spaces

For convenience, we recall the Gromov-Hausdorff metric, see for example Definition 7.3.10 in [BBI01]. Let (X, d) and (X', d') be two compact metric spaces. The Gromov-Hausdorff metric between (X, d) and (X', d') is given by:

$$d_{\mathrm{GH}}^c\big((X,d),(X',d')\big) = \inf_{\varphi,\varphi',Z} d_{\mathrm{H}}^Z\big(\varphi(X),\varphi'(X')\big), \tag{2.5}$$

where the infimum is taken over all isometric embeddings $\varphi : X \hookrightarrow Z$ and $\varphi' : X' \hookrightarrow Z$ into some common Polish metric space (Z, d^Z). Note that Equation (2.5) does actually define a metric on the set of isometry classes of compact metric spaces.

Now, we introduce the Gromov-Hausdorff-Prokhorov metric for compact spaces. Let $\mathscr{X} = (X, d, \emptyset, \mu)$ and $\mathscr{X}' = (X', d', \emptyset', \mu')$ be two compact rooted weighted metric spaces, and define:

$$d_{\mathrm{GHP}}^c(\mathscr{X},\mathscr{X}') = \inf_{\Phi,\Phi',Z}\Big(d^Z(\Phi(\emptyset),\Phi'(\emptyset')) + d_{\mathrm{H}}^Z(\Phi(X),\Phi'(X')) + d_{\mathrm{P}}^Z(\Phi_*\mu,\Phi'_*\mu')\Big), \tag{2.6}$$

where the infimum is taken over all isometric embeddings $\Phi : X \hookrightarrow Z$ and $\Phi' : X' \hookrightarrow Z$ into some common Polish metric space (Z, d^Z).

Note that equation (2.6) does not actually define a metric, as $d_{\mathrm{GHP}}^c(\mathscr{X},\mathscr{X}') = 0$ if $\mathscr{X}$ and $\mathscr{X}'$ are GHP-isometric. Therefore, we will consider $\mathbb{K}$, the set of GHP-isometry classes of compact rooted weighted metric space and identify a compact rooted weighted metric space with its class in $\mathbb{K}$. Then the function d_{GHP}^c is finite on $\mathbb{K}^2$.

Theorem 2.3.

(i) The function d_{GHP}^c defines a metric on $\mathbb{K}$.

(ii) The space $(\mathbb{K}, d_{\mathrm{GHP}}^c)$ is a Polish metric space.

We will call d^c_{GHP} the Gromov-Hausdorff-Prokhorov metric. This extends the Gromov-Hausdorff metric on compact metric spaces, see [BBI01] section 7, as well as the Gromov-Hausdorff-Prokhorov metric on compact metric spaces endowed with a probability measure, see [Mie09]. See also [GPW08] for another approach on metric spaces endowed with a probability measure.

We end this Section by a pre-compactness criterion on $\mathbb{K}$.

Theorem 2.4. *Let $\mathscr{A}$ be a subset of $\mathbb{K}$, such that:*

(i) We have $\sup_{(X,d,\emptyset,\mu)\in\mathscr{A}} \mathrm{diam}(X) < +\infty$.

(ii) For every $\varepsilon > 0$, there exists a finite integer $N(\varepsilon) \geq 1$, such that for any $(X,d,\emptyset,\mu) \in \mathscr{A}$, there is an ε-net of X with cardinal less than $N(\varepsilon)$.

(iii) We have $\sup_{(X,d,\emptyset,\mu)\in\mathscr{A}} \mu(X) < +\infty$.

Then, $\mathscr{A}$ is relatively compact: every sequence in $\mathscr{A}$ admits a sub-sequence that converges in the d^c_{GHP} topology.

Notice that we could have defined a Gromov-Hausdorff-Prokhorov metric without reference to any root. However, the introduction of the root is necessary to define the Gromov-Hausdorff-Prokhorov metric for locally compact spaces, see next Section.

Gromov-Hausdorff-Prokhorov metric for locally compact spaces

To consider an extension to non compact weighted rooted metric spaces, we will consider complete and locally compact length spaces. We recall that a metric space (X,d) is a length space if for every $x,y \in X$, we have:

$$d(x,y) = \inf L(\gamma),$$

where the infimum is taken over all rectifiable curves $\gamma : [0,1] \to X$ such that $\gamma(0) = x$ and $\gamma(1) = y$, and where $L(\gamma)$ is the length of the rectifiable curve γ. We recall that (X,d) is a length space if is satisfies the mid-point condition (see Theorem 2.4.16 in [BBI01]): for all $\varepsilon > 0$, $x,y \in X$, there exists $z \in X$ such that:

$$|2d(x,z) - d(x,y)| + |2d(y,z) - d(x,y)| \leq \varepsilon.$$

Definition 2.5. *Let $\mathbb{L}$ be the set of GHP-isometry classes of rooted, weighted, complete and locally compact length spaces and identify a rooted, weighted, complete and locally compact length spaces with its class in $\mathbb{L}$.*

If $\mathscr{X} = (X,d,\emptyset,\mu) \in \mathbb{L}$, then for $r \geq 0$ we will consider its restriction to the closed ball of radius r centered at $\emptyset$, $\mathscr{X}^{(r)} = (X^{(r)}, d^{(r)}, \emptyset, \mu^{(r)})$, where $X^{(r)}$ is defined by (2.2), the metric $d^{(r)}$ is the restriction of d to $X^{(r)}$, and the measure $\mu^{(r)}$ is defined by (2.3). Recall that the Hopf-Rinow theorem implies that if (X,d) is a complete and locally compact length space, then every closed bounded subset of X is compact. In particular if $\mathscr{X}$ belongs to $\mathbb{L}$, then $\mathscr{X}^{(r)}$ belongs to $\mathbb{K}$ for all $r \geq 0$.

We state a regularity lemma of d^c_{GHP} with respect to the restriction operation.

Lemma 2.6. *Let $\mathscr{X}$ and $\mathscr{Y}$ be in $\mathbb{L}$. Then the function defined on $\mathbf{R}_+$ by $r \mapsto d^c_{\mathrm{GHP}}\left(\mathscr{X}^{(r)}, \mathscr{Y}^{(r)}\right)$ is càdlàg.*

This implies that the following function (inspired by (2.4)) is well defined on $\mathbb{L}^2$:

$$d_{\mathrm{GHP}}(\mathscr{X}, \mathscr{Y}) = \int_0^\infty \mathrm{e}^{-r} \left(1 \wedge d^c_{\mathrm{GHP}}\left(\mathscr{X}^{(r)}, \mathscr{Y}^{(r)}\right)\right) \, dr.$$

Theorem 2.7.

(i) The function d_{GHP} defines a metric on $\mathbb{L}$.

(ii) The space $(\mathbb{L}, d_{\mathrm{GHP}})$ is a Polish metric space.

The next result implies that d^c_{GHP} and d_{GHP} define the same topology on $\mathbb{K} \cap \mathbb{L}$.

Proposition 2.8. *Let $(\mathscr{X}_n, n \in \mathbf{N})$ and $\mathscr{X}$ be elements of $\mathbb{K} \cap \mathbb{L}$. Then the sequence $(\mathscr{X}_n, n \in \mathbf{N})$ converges to $\mathscr{X}$ in $(\mathbb{K}, d^c_{\mathrm{GHP}})$ if and only if it converges to $\mathscr{X}$ in $(\mathbb{L}, d_{\mathrm{GHP}})$.*

Finally, we give a pre-compactness criterion on $\mathbb{L}$ which is a generalization of the well-known compactness theorem for compact metric spaces, see for instance Theorem 7.4.15 in [BBI01].

Theorem 2.9. *Let $\mathscr{C}$ be a subset of $\mathbb{L}$, such that for every $r \geq 0$:*

(i) For every $\varepsilon > 0$, there exists a finite integer $N(r, \varepsilon) \geq 1$, such that for any $(X, d, \emptyset, \mu) \in \mathscr{C}$, there is an ε-net of $X^{(r)}$ with cardinal less than $N(r, \varepsilon)$.

(ii) We have $\sup_{(X,d,\emptyset,\mu)\in\mathscr{C}} \mu(X^{(r)}) < +\infty$.

Then, $\mathscr{C}$ is relatively compact: every sequence in $\mathscr{C}$ admits a sub-sequence that converges in the d_{GHP} topology.

2.3 Application to real trees coded by functions

A metric space (T, d) is a called real tree (or $\mathbf{R}$-tree) if the following properties are satisfied:

(i) For every $s, t \in T$, there is a unique isometric map $f_{s,t}$ from $[0, d(s,t)]$ to T such that $f_{s,t}(0) = s$ and $f_{s,t}(d(s,t)) = t$.

(ii) For every $s, t \in T$, if q is a continuous injective map from $[0,1]$ to T such that $q(0) = s$ and $q(1) = t$, then $q([0,1]) = f_{s,t}([0, d(s,t)])$.

Note that real trees are always length spaces and that complete real trees are the only complete connected spaces that satisfy the so-called four-point condition:

$$\forall x_1, x_2, x_3, x_4 \in X, \;\; d(x_1,x_2) + d(x_3,x_4) \leq (d(x_1,x_3) + d(x_2,x_4)) \vee (d(x_1,x_4) + d(x_2,x_3)). \quad (2.7)$$

We say that a real tree is rooted if there is a distinguished vertex $\emptyset$, which will be called the root of T.

Definition 2.10. *We denote by $\mathbb{T}$ the set of (GHP-isometry classes of) rooted, weighted, complete and locally compact real trees, in short w-trees.*

We deduce the following Corollary from Theorem 2.7 and the four-point condition characterization of real trees.

Corollary 2.11. *The set $\mathbb{T}$ is a closed subset of $\mathbb{L}$ and $(\mathbb{T}, d_{\mathrm{GHP}})$ is a Polish metric space.*

Let f be a continuous non-negative function defined on $[0, +\infty)$, such that $f(0) = 0$, with compact support. We set:

$$\sigma^f = \sup \{t; f(t) > 0\},$$

with the convention $\sup \emptyset = 0$. Let d^f be the non-negative function defined by:

$$d^f(s,t) = f(s) + f(t) - 2 \inf_{u \in [s \wedge t, s \vee t]} f(u).$$

It can be easily checked that d^f is a semi-metric on $[0, \sigma^f]$. One can define the equivalence relation associated with d^f by $s \sim t$ if and only if $d^f(s,t) = 0$. Moreover, when we consider the quotient space

$$T^f = \left[0, \sigma^f\right]_{/\sim}$$

and, noting again d^f the induced metric on T^f and rooting T^f at $\emptyset^f$, the equivalence class of 0, it can be checked that the space $(T^f, d^f, \emptyset^f)$ is a rooted compact real tree. We denote by p^f the canonical projection from $[0, \sigma^f]$ onto T^f, which is extended by $p^f(t) = \emptyset^f$ for $t \geq \sigma^f$. Notice that p^f is continuous. We define the Borel measure $\mathbf{m}^f$ on T^f as the image measure of the Lebesgue measure on $[0, \sigma^f]$ by p^f. We consider the (compact) w-tree $\mathscr{T}^f = (T^f, d^f, \emptyset^f, \mathbf{m}^f)$.

We have the following elementary result (see Lemma 2.3 of [DL05] when dealing with the Gromov-Hausdorff metric instead of the Gromov-Hausdorff-Prokhorov metric). For a proof, see [ADH12c].

Proposition 2.12. *Let f, g be two compactly supported, non-negative continuous functions with $f(0) = g(0) = 0$. Then, we have:*

$$d^c_{\mathrm{GHP}}(\mathscr{T}^f, \mathscr{T}^g) \leq 6\|f - g\|_\infty + \left|\sigma^f - \sigma^g\right|. \tag{2.8}$$

This result and Proposition 2.8 ensure that the map $f \mapsto \mathscr{T}^f$ (defined on the space of continuous functions with compact support which vanish at 0, with the uniform topology) taking values in $(\mathbb{T} \cap \mathbb{K}, d^c_{\mathrm{GHP}})$ or $(\mathbb{T}, d_{\mathrm{GHP}})$ is measurable.

2.4 Gromov-Hausdorff-Prokhorov metric for compact metric spaces

Proof of (i) of Theorem 2.3

In this Section, we will prove that d^c_{GHP} defines a metric on $\mathbb{K}$.

First, we will prove the following technical lemma, which is a generalization of Remark 7.3.12 in [BBI01]. Let $\mathscr{X} = (X, d^X, \emptyset^X, \mu^X)$ and $\mathscr{Y} = (Y, d^Y, \emptyset^Y, \mu^Y)$ be two elements of $\mathbb{K}$. We will use the notation $X \sqcup Y$ for the disjoint union of the sets X and Y. We will abuse notations and note $X, \mu^X, \emptyset^X$ and $Y, \mu^Y, \emptyset^Y$ the images of $X, \mu^X, \emptyset^X$ and of $Y, \mu^Y, \emptyset^Y$ respectively by the canonical embeddings $X \hookrightarrow X \sqcup Y$ and $Y \hookrightarrow X \sqcup Y$.

Lemma 2.13. *Let $\mathscr{X} = (X, d^X, \emptyset^X, \mu^X)$ and $\mathscr{Y} = (Y, d^Y, \emptyset^Y, \mu^Y)$ be two elements of $\mathbb{K}$. Then, we have:*

$$d^c_{\mathrm{GHP}}(\mathscr{X}, \mathscr{Y}) = \inf_d \left\{ d(\emptyset^X, \emptyset^Y) + d^d_{\mathrm{H}}(X, Y) + d^d_{\mathrm{P}}(\mu^X, \mu^Y) \right\}, \tag{2.9}$$

where the infimum is taken over all metrics d on $X \sqcup Y$ such that the canonical embeddings $X \hookrightarrow X \sqcup Y$ and $Y \hookrightarrow X \sqcup Y$ are isometries.

Proof. We only have to show that:

$$\inf_d \left\{ d(\emptyset^X, \emptyset^Y) + d^d_{\mathrm{H}}(X, Y) + d^d_{\mathrm{P}}(\mu^X, \mu^Y) \right\} \leq d^c_{\mathrm{GHP}}(\mathscr{X}, \mathscr{Y}), \tag{2.10}$$

since the other inequality is obvious. Let (Z, d^Z) be a Polish space and Φ^X and Φ^Y be isometric embeddings of X and Y in Z. Let $\delta > 0$. We define the following function on $(X \sqcup Y)^2$:

$$d(x, y) = \begin{cases} d^Z(\Phi^X(x), \Phi^Y(y)) + \delta & \text{if } x \in X,\ y \in Y, \\ d^X(x, y) & \text{if } x, y \in X, \\ d^Y(x, y) & \text{if } x, y \in Y. \end{cases} \tag{2.11}$$

It is obvious that d is a metric on $X \sqcup Y$, and that the canonical embeddings of X and Y in $X \sqcup Y$ are isometric. Furthermore, by definition, we have

$$d\left(\emptyset^X, \emptyset^Y\right) = d^Z\left(\Phi^X(\emptyset^X), \Phi^Y(\emptyset^Y)\right) + \delta.$$

Concerning the Hausdorff distance between X and Y, we get that:

$$d^d_{\mathrm{H}}(X, Y) \leq d^Z_{\mathrm{H}}(\Phi^X(X), \Phi^Y(Y)) + \delta.$$

Finally, let us compute the Prokhorov distance between μ^X and μ^Y. Let $\varepsilon > 0$ be such that $d^Z_{\mathrm{P}}(\Phi^X_* \mu^X, \Phi^Y_* \mu^Y) < \varepsilon$. Let A be a closed subset of $X \sqcup Y$. By definition, we have:

$$\begin{aligned} \mu^X(A) = \mu^X(A \cap X) &= \Phi^X_* \mu^X\left(\Phi^X(A \cap X)\right) \\ &< \Phi^Y_* \mu^Y\left(\left\{z \in Z,\ d^Z(z, \Phi^X(A \cap X)) < \varepsilon\right\}\right) + \varepsilon \\ &= \Phi^Y_* \mu^Y\left(\left\{z \in \Phi^Y(Y),\ d^Z(z, \Phi^X(A \cap X)) < \varepsilon\right\}\right) + \varepsilon \\ &\leq \mu^Y(\{y \in Y,\ d(y, A \cap X) < \varepsilon + \delta\}) + \varepsilon \\ &\leq \mu^Y(\{y \in X \sqcup Y,\ d(y, A) < \varepsilon + \delta\}) + \varepsilon. \end{aligned}$$

The symmetric result holds for (X, Y) replaced by (Y, X) and therefore we get that actually $d^d_{\mathrm{P}}(\mu^X, \mu^Y) < \varepsilon + \delta$. This implies:

$$d^d_{\mathrm{P}}\left(\mu^X, \mu^Y\right) \leq d^Z_{\mathrm{H}}\left(\Phi^X_* \mu^X, \Phi^Y_* \mu^Y\right) + \delta.$$

Eventually, we get:

$$d\left(\varnothing^X, \varnothing^Y\right) + d_{\mathrm{H}}^d(X,Y) + d_{\mathrm{P}}^d\left(\mu^X,\mu^Y\right) \\ \leq d^Z\left(\Phi^X(\varnothing^X), \Phi^Y(\varnothing^Y)\right) + d_{\mathrm{H}}^Z\left(\Phi^X(X), \Phi^Y(Y)\right) + d_{\mathrm{H}}^Z\left(\Phi_*^X\mu^X, \Phi_*^Y\mu^Y\right) + 3\delta.$$

Thanks to (2.6) and since $\delta > 0$ is arbitrary, we get (2.10). □

We now prove that d_{GHP}^c does indeed satisfy all the axioms of a metric (as is done in [BBI01] for the Gromov-Hausdorff metric and in [Mie09] in the case of probability measures on compact metric spaces). The symmetry and positiveness of d_{GHP}^c being obvious, let us prove the triangular inequality and positive definiteness.

Lemma 2.14. *The function d_{GHP}^c satisfies the triangular identity on $\mathbb{K}$.*

Proof. Let $\mathcal{X}_1, \mathcal{X}_2$ and $\mathcal{X}_3$ be elements of $\mathbb{K}$. Let us assume that $d_{\mathrm{GHP}}^c(\mathcal{X}_i, \mathcal{X}_2) < r_i$ for $i \in \{1,3\}$. With obvious notations, for $i \in \{1,3\}$, we consider, as in Lemma 2.13, metrics d_i on $X_i \sqcup X_2$. Let us then consider $Z = X_1 \sqcup X_2 \sqcup X_3$, on which we define:

$$d(x,y) = \begin{cases} d_i(x,y) & \text{if } x,y \in (X_i \sqcup X_2)^2 \text{ for } i \in \{1,3\}, \\ \inf_{z \in X_2}\{d_1(x,z) + d_3(z,y)\} & \text{if } x \in X_1, y \in X_3. \end{cases} \tag{2.12}$$

The function d is in fact a metric on Z, and the canonical embeddings are isometries, since they are for d_1 and d_3. By definition, we have:

$$d_{\mathrm{H}}^d(X_1,X_3) = \left(\sup_{x_1 \in X_1} \inf_{x_3 \in X_3} d(x_1,x_3)\right) \vee \left(\sup_{x_3 \in X_3} \inf_{x_1 \in X_1} d(x_1,x_3)\right).$$

We notice that:

$$\begin{aligned} \sup_{x_1 \in X_1} \inf_{x_3 \in X_3} d(x_1,x_3) &= \sup_{x_1 \in X_1} \inf_{x_2 \in X_2,\ x_3 \in X_3} d_1(x_1,x_2) + d_3(x_2,x_3) \\ &\leq d_{\mathrm{H}}^{d_1}(X_1,X_2) + \inf_{x_2 \in X_2,\ x_3 \in X_3} d_3(x_2,x_3) \\ &\leq d_{\mathrm{H}}^{d_1}(X_1,X_2) + d_{\mathrm{H}}^{d_3}(X_2,X_3). \end{aligned}$$

Thus, we deduce that $d_{\mathrm{H}}^d(X_1,X_3) \leq d_{\mathrm{H}}^{d_1}(X_1,X_2) + d_{\mathrm{H}}^{d_3}(X_2,X_3)$.

As far as the Prokhorov distance is concerned, let $\varepsilon_i,\ i \in \{1,3\}$ be such that $d_{\mathrm{P}}^{d_i}(\mu_i,\mu_2) < \varepsilon_i$. Then, if $A \subset Z$ is closed, we have:

$$\begin{aligned} \mu_1(A) = \mu_1(A \cap X_1) &< \mu_2(\{x \in X_1 \sqcup X_2,\ d_1(x, A \cap X_1) < \varepsilon_1\}) + \varepsilon_1 \\ &\leq \mu_2(A^{\varepsilon_1} \cap X_2) + \varepsilon_1 \\ &< \mu_3(\{x \in X_3 \sqcup X_2,\ d_3(x, A^{\varepsilon_1} \cap X_2) < \varepsilon_3\}) + \varepsilon_1 + \varepsilon_3 \\ &\leq \mu_3(A^{\varepsilon_1+\varepsilon_3}) + \varepsilon_1 + \varepsilon_3, \end{aligned}$$

where $A^\varepsilon = \{z \in Z,\ d(z, A) < \varepsilon\}$, for $\varepsilon = \varepsilon_1$ and $\varepsilon = \varepsilon_1 + \varepsilon_3$. A similar result holds with (μ_1, μ_3) replaced by (μ_3, μ_1). We deduce that $d_{\mathrm{P}}^d(\mu_1, \mu_3) < \varepsilon_1 + \varepsilon_3$, which implies that

$$d_{\mathrm{P}}^d(\mu_1, \mu_3) \le d_{\mathrm{P}}^{d_1}(\mu_1, \mu_2) + d_{\mathrm{P}}^{d_3}(\mu_2, \mu_3).$$

By summing up all the results, we get:

$$d(\varnothing_1, \varnothing_3) + d_{\mathrm{H}}^d(X_1, X_3) + d_{\mathrm{P}}^d(\mu_1, \mu_3) \le \sum_{i \in \{1,3\}} d^{d_i}(\varnothing_i, \varnothing_2) + d_{\mathrm{H}}^{d_i}(X_i, X_2) + d_{\mathrm{P}}^{d_i}(\mu_i, \mu_2).$$

Then use the definition (2.6) and Lemma 2.13 to get the triangular inequality:

$$d_{\mathrm{GHP}}^c(\mathscr{X}_1, \mathscr{X}_3) \le d_{\mathrm{GHP}}^c(\mathscr{X}_1, \mathscr{X}_2) + d_{\mathrm{GHP}}^c(\mathscr{X}_2, \mathscr{X}_3).$$

□

This proves that d_{GHP}^c is a semi-metric on $\mathbb{K}$. We then prove the positive definiteness.

Lemma 2.15. *Let $\mathscr{X}, \mathscr{Y}$ be two elements of $\mathbb{K}$ such that $d_{\mathrm{GHP}}^c(\mathscr{X}, \mathscr{Y}) = 0$. Then $\mathscr{X} = \mathscr{Y}$ (as GHP-isometry classes of rooted weighted compact metric spaces).*

Proof. Let $\mathscr{X} = (X, d^X, \varnothing^X, \mu^X)$ and $\mathscr{Y} = (Y, d^Y, \varnothing^Y, \mu^Y)$ in $\mathbb{K}$ such that $d_{\mathrm{GHP}}^c(\mathscr{X}, \mathscr{Y}) = 0$. According to Lemma 2.13, we can find a sequence of metrics $(d^n, n \ge 1)$ on $X \sqcup Y$, such that

$$d^n(\varnothing^X, \varnothing^Y) + d_{\mathrm{H}}^n(X, Y) + d_{\mathrm{P}}^n(\mu^X, \mu^Y) < \varepsilon_n, \tag{2.13}$$

for some positive sequence $(\varepsilon_n, n \ge 1)$ decreasing to 0, where d_{H}^n and d_{P}^n stand for $d_{\mathrm{H}}^{d^n}$ and $d_{\mathrm{P}}^{d^n}$. For any $k \ge 1$, let S_k be a finite $(1/k)$-net of X, containing the root. Since X is compact, we get by Definition 2.1 that S_k is in fact an $(\frac{1}{k} - \delta)$-net of X for some $\delta > 0$. Let $N_k + 1$ be the cardinal of S_k. We will write:

$$S_k = \{x_{0,k} = \varnothing^X, x_{1,k}, \dots, x_{N_k,k}\}.$$

Let $(V_{i,k}, 0 \le i \le N_k)$ be Borel subsets of X with diameter less than $1/k$, that is:

$$\sup_{x, x' \in V_{i,k}} d^X(x, x') < 1/k,$$

such that $\cup_{0 \le i \le N_k} V_{i,k} = X$ and for all $0 \le i, i' \le N_k$, we have $V_{i,k} \cap V_{i',k} = \emptyset$ and $x_{i,k} \in V_{i,k}$ if $V_{i,k} \neq \emptyset$. We set:

$$\mu_k^X(dx) = \sum_{i=0}^{N_k} \mu^X(V_{i,k}) \delta_{x_{i,k}}(dx),$$

where $\delta_{x'}(dx)$ is the Dirac measure at x'. Notice that:

$$d_{\mathrm{H}}^X(X, S_k) \le \frac{1}{k} \quad \text{and} \quad d_{\mathrm{P}}^X(\mu_k^X, \mu^X) \le \frac{1}{k} \cdot$$

We set $y_{0,k} = y_{0,k}^n = \varnothing^Y$. By (2.13), we get that for any $k \ge 1, 0 \le i \le N_k$, there exists $y_{i,k}^n \in Y$ such that $d^n(x_{i,k}, y_{i,k}^n) < \varepsilon_n$. Since Y is compact, the sequence $(y_{i,k}^n, n \ge 1)$ is relatively compact, hence admits a converging sub-sequence. Using a diagonal argument, and without loss of generality (by considering the sequence instead of the sub-sequence), we may assume that for $k \geq 1, 0 \leq i \leq N_k$, the sequence $(y^n_{i,k}, n \geq 1)$ converges to some $y_{i,k} \in Y$.

For any $y \in Y$, let $x \in X$ such that $d^n(x,y) < \varepsilon_n$ and i,k such that $d^X(x, x_{i,k}) < \frac{1}{k} - \delta$. Then, we get:

$$d^Y\left(y, y^n_{i,k}\right) = d^n\left(y, y^n_{i,k}\right) \leq d^n(y,x) + d^X\left(x, x_{i,k}\right) + d^n\left(x_{i,k}, y^n_{i,k}\right) \leq \frac{1}{k} - \delta + 2\varepsilon_n.$$

Thus, the set $\{y^n_{i,k}, 0 \leq i \leq N_k\}$ is a $(2\varepsilon_n + 1/k - \delta)$-net of Y, and the set $S^Y_k = \{y_{i,k}, 0 \leq i \leq N_k\}$ is an $1/k$-net of Y.

If $k, k' \geq 1$ and $0 \leq i \leq N_k, 0 \leq i' \leq N_{k'}$, then we have:

$$\begin{aligned} d^Y\left(y_{i,k}, y_{i',k'}\right) &\leq d^Y\left(y^n_{i,k}, y_{i,k}\right) + d^Y\left(y^n_{i,k}, y^n_{i',k'}\right) + d^Y\left(y^n_{i',k'}, y_{i',k'}\right) \\ &\leq d^Y\left(y^n_{i,k}, y_{i,k}\right) + d^Y\left(y^n_{i',k'}, y_{i',k'}\right) + 2\varepsilon_n + d^X\left(x_{i,k}, x_{i',k'}\right), \end{aligned}$$

and, since the terms $d(y^n_{i,k}, y_{i,k})$ and $d(y^n_{i',k'}, y_{i',k'})$ can be made arbitrarily small, we deduce:

$$d\left(y_{i,k}, y_{i',k'}\right) \leq d\left(x_{i,k}, x_{i',k'}\right).$$

The reverse inequality is proven using similar arguments, so that the above inequality is in fact an equality. Therefore the map defined by $\Phi(x_{i,k}) = (y_{i,k})$ from $\cup_{k\geq 1} S_k$ onto $\cup_{k\geq 1} S^Y_k$ is a root-preserving isometry. By density, this map can be extended uniquely to an isometric one-to-one root preserving embedding from X to Y which we still denote by Φ. Hence the metric spaces X and Y are root-preserving isometric.

As far as the measures are concerned, we set:

$$\mu^{Y,n}_k = \sum_{i=0}^{N_k} \mu^X(V_{i,k})\delta_{y^n_{i,k}} \quad \text{and} \quad \mu^Y_k = \sum_{i=0}^{N_k} \mu^X(V_{i,k})\delta_{y_{i,k}}.$$

By construction, we have $d^n_{\mathrm{P}}(\mu^{Y,n}_k, \mu^X_k) \leq \varepsilon_n$. We get:

$$\begin{aligned} d^Y_{\mathrm{P}}\left(\mu^Y_k, \mu^Y\right) = d^n_{\mathrm{P}}\left(\mu^Y_k, \mu^Y\right) &\leq d^Y_{\mathrm{P}}\left(\mu^Y_k, \mu^{Y,n}_k\right) + d^n_{\mathrm{P}}\left(\mu^{Y,n}_k, \mu^X_k\right) + d^X_{\mathrm{P}}\left(\mu^X_k, \mu^X\right) + d^n_{\mathrm{P}}(\mu^X, \mu^Y) \\ &< d^Y_{\mathrm{P}}\left(\mu^Y_k, \mu^{Y,n}_k\right) + \varepsilon_n + \frac{1}{k} + \varepsilon_n. \end{aligned}$$

Furthermore, as n goes to infinity, we have that $d^Y_{\mathrm{P}}(\mu^Y_k, \mu^{Y,n}_k)$ converges to 0, since the $y^n_{i,k}$ converge towards the $y_{i,k}$. Thus, we actually have:

$$d^Y_{\mathrm{P}}\left(\mu^Y_k, \mu^Y\right) \leq 1/k.$$

This implies that $(\mu^Y_k, k \geq 1)$ converges weakly to μ^Y. Since by definition $\mu^Y_k = \Phi_* \mu^X_k$ and since Φ is continuous, by passing to the limit, we get $\mu^Y = \Phi_* \mu^X$. This gives that $\mathscr{X}$ and $\mathscr{Y}$ are GHP-isometric. □

This proves that the function d^c_{GHP} defines a metric on $\mathbb{K}$.

Proof of Theorem 2.4 and of (ii) of Theorem 2.3

The proof of Theorem 2.4 is very close to the proof of Theorem 7.4.15 in [BBI01], where only the Gromov-Hausdorff metric is involved. It is in fact a simplified version of the proof of Theorem 2.9, and is thus left to the reader.

We are left with the proof of (ii) of Theorem 2.3. It is in fact enough to check that if $(\mathscr{X}_n, n \in \mathbf{N})$ is a Cauchy sequence, then it is relatively compact.

First notice that if (Z, d^Z) is a Polish metric space, then for any closed subsets A, B, we have $d_{\mathrm{H}}^Z(A,B) \geq |\mathrm{diam}(A) - \mathrm{diam}(B)|$, and for $\mu, \nu \in \mathscr{M}_f(Z)$, $d_{\mathrm{P}}^Z(\mu,\nu) \geq \left|\mu(Z) - \nu(Z)\right|$. This implies that for any $\mathscr{X} = (X, d^X, \emptyset^X, \mu), \mathscr{Y} = (Y, d^Y, \emptyset^Y, \nu) \in \mathbb{K}$:

$$d_{\mathrm{GHP}}^c(\mathscr{X},\mathscr{Y}) \geq |\mathrm{diam}(X) - \mathrm{diam}(Y)| + \left|\mu(X) - \nu(Y)\right|. \tag{2.14}$$

Furthermore, using the definition of the Gromov-Hausdorff metric (2.5), we clearly have:

$$d_{\mathrm{GHP}}^c(\mathscr{X},\mathscr{Y}) \geq d_{\mathrm{GH}}^c((X,d^X),(Y,d^Y)). \tag{2.15}$$

We deduce that if $\mathscr{A} = (\mathscr{X}_n, n \in \mathbf{N})$ is a Cauchy sequence, then (2.14) implies that conditions (i) and (iii) of Theorem 2.4 are fulfilled. Furthermore, thanks to (2.15), the sequence $((X_n, d^{X_n}), n \in \mathbf{N})$ is a Cauchy sequence for the Gromov-Hausdorff metric. Then point (2) of Proposition 7.4.11 in [BBI01] readily implies condition (ii) of Theorem 2.4.

2.5 Extension to locally compact length spaces

First results

First, let us state two elementary lemmas. Let $(X, d, \emptyset)$ be a rooted metric space. Recall notation (2.2). We set:

$$\partial_r X = \{x \in X;\ d(\emptyset^x, x) = r\}.$$

Lemma 2.16. *Let $(X,d,\emptyset)$ be a complete rooted length space and $r, \varepsilon > 0$. Then we have, for all $\delta > 0$:*

$$X^{(r+\varepsilon)} \subset \left(X^{(r)}\right)^{\varepsilon+\delta}.$$

Proof. Let $x \in X^{(r+\varepsilon)} \setminus X^{(r)}$ and $\delta > 0$. There exists a rectifiable curve γ defined on $[0,1]$ with values in X such that $\gamma(0) = \emptyset$ and $\gamma(1) = x$, such that $L(\gamma) < d(\emptyset, x) + \delta \leq r + \varepsilon + \delta$. There exists $t \in (0,1)$ such that $\gamma(t) \in \partial_r X$. We can bound $d(\gamma(t), x)$ by the length of the fragment of γ joining $\gamma(t)$ and x, that is the length of γ minus the length of the fragment of γ joining $\emptyset$ to $\gamma(t)$. The latter being equal to or larger than $d(\emptyset^X, \gamma(t)) = r$, we get:

$$d(\gamma(t), x) \leq L(\gamma) - r < \varepsilon + \delta.$$

Since $\gamma(t) \in X^{(r)}$, we get $x \in \left(X^{(r)}\right)^{\varepsilon+\delta}$. This ends the proof. □

Lemma 2.17. *Let $\mathscr{X} = (X, d, \emptyset, \mu) \in \mathbb{L}$. For all $\varepsilon > 0$ and $r > 0$, we have:*

$$d_{\mathrm{GHP}}^c\left(\mathscr{X}^{(r)}, \mathscr{X}^{(r+\varepsilon)}\right) \leq \varepsilon + \mu\left(X^{(r+\varepsilon)} \setminus X^{(r)}\right).$$

Proof. The identity map is an obvious embedding $X^{(r)} \hookrightarrow X^{(r+\varepsilon)}$ which is root-preserving. Then, we have:

$$d^c_{\mathrm{GHP}}(\mathscr{X}^{(r)}, \mathscr{X}^{(r+\varepsilon)}) \leq d_{\mathrm{H}}(X^{(r)}, X^{(r+\varepsilon)}) + d_{\mathrm{P}}(\mu^{(r)}, \mu^{(r+\varepsilon)}).$$

Thanks to Lemma 2.16, we have $d_{\mathrm{H}}(X^{(r)}, X^{(r+\varepsilon)}) \leq \varepsilon$.

Let $A \subset X$ be closed. We have obviously $\mu^{(r)}(A) \leq \mu^{(r+\varepsilon)}(A)$. On the other hand, we have:

$$\mu^{(r+\varepsilon)}(A) \leq \mu^{(r)}(A) + \mu(A \cap (X^{(r+\varepsilon)} \setminus X^{(r)})) \leq \mu^{(r)}(A) + \mu(X^{(r+\varepsilon)} \setminus X^{(r)}).$$

This proves that $d_{\mathrm{P}}(\mu^{(r)}, \mu^{(r+\varepsilon)}) \leq \mu(X^{(r+\varepsilon)} \setminus X^{(r)})$, which ends the proof. □

It is then straightforward to prove Lemma 2.6.

Proof of Lemma 2.6. Let $\mathscr{X} = (X, d^X, \emptyset^X, \mu^X)$ and $\mathscr{Y} = (Y, d^Y, \emptyset^Y, \mu^Y)$ be two elements of $\mathbb{L}$. Using the triangular inequality twice and Lemma 2.17, we get for $r > 0$ and $\varepsilon > 0$:

$$\begin{aligned}
\left| d^c_{\mathrm{GHP}}(\mathscr{X}^{(r)}, \mathscr{Y}^{(r)}) - d^c_{\mathrm{GHP}}(\mathscr{X}^{(r+\varepsilon)}, \mathscr{Y}^{(r+\varepsilon)}) \right| &\leq d^c_{\mathrm{GHP}}\left(\mathscr{X}^{(r)}, \mathscr{X}^{(r+\varepsilon)}\right) + d^c_{\mathrm{GHP}}\left(\mathscr{Y}^{(r)}, \mathscr{Y}^{(r+\varepsilon)}\right) \\
&\leq 2\varepsilon + \mu^X\left(X^{(r+\varepsilon)} \setminus X^{(r)}\right) + \mu^Y\left(Y^{(r+\varepsilon)} \setminus Y^{(r)}\right).
\end{aligned}$$

As ε goes down to 0, the expression above converges to 0, so that we get right-continuity of the function $r \mapsto d^c_{\mathrm{GHP}}(\mathscr{X}^{(r)}, \mathscr{Y}^{(r)})$.

We write $\mathscr{X}^{(r-)}$ for the compact metric space $X^{(r)}$ rooted at $\emptyset^X$ along with the induced metric and the restriction of μ to the *open* ball $\{x \in X;\ d^X(\emptyset^X, x) < r\}$. We define $\mathscr{Y}^{(r-)}$ similarly. Similar arguments as above yield for $r > \varepsilon > 0$:

$$\begin{aligned}
&\left| d^c_{\mathrm{GHP}}(\mathscr{X}^{(r-)}, \mathscr{Y}^{(r-)}) - d^c_{\mathrm{GHP}}(\mathscr{X}^{(r-\varepsilon)}, \mathscr{Y}^{(r-\varepsilon)}) \right| \\
&\qquad \leq d^c_{\mathrm{GHP}}(\mathscr{X}^{(r-)}, \mathscr{X}^{(r-\varepsilon)}) + d^c_{\mathrm{GHP}}(\mathscr{Y}^{(r)}, \mathscr{Y}^{(r-\varepsilon)}) \\
&\qquad \leq 2\varepsilon + \mu^X\left(\{x \in X,\ r - \varepsilon < d^X(\emptyset^X, x) < r\}\right) + \mu^Y\left(\{y \in Y,\ r - \varepsilon < d^Y(\emptyset^Y, y) < r\}\right).
\end{aligned}$$

As ε goes down to 0, the expression above also converges to 0, which shows the existence of left limits for the function $r \mapsto d^c_{\mathrm{GHP}}(\mathscr{X}^{(r)}, \mathscr{Y}^{(r)})$. □

The next result corresponds to (i) in Theorem 2.7.

Proposition 2.18. *The function d_{GHP} is a metric on $\mathbb{L}$.*

Proof. The symmetry and positivity of d_{GHP} are obvious. The triangle inequality is not difficult either, since d^c_{GHP} satisfies the triangle inequality and the map $x \mapsto 1 \wedge x$ is non-decreasing and sub-additive.

We need to check that d_{GHP} is definite positive. To that effect, let $\mathscr{X} = (X, d^X, \emptyset^X, \mu)$ and $\mathscr{Y} = (Y, d^Y, \emptyset^Y, \nu)$ be two elements of $\mathbb{L}$ such that $d_{\mathrm{GHP}}(\mathscr{X}, \mathscr{Y}) = 0$. We want to prove that $\mathscr{X}$ and $\mathscr{Y}$ are GHP-isometric. We follow the spirit of the proof of Lemma 2.15.

By definition, we get that for almost every $r > 0$, $d^c_{\mathrm{GHP}}(\mathscr{X}^{(r)}, \mathscr{Y}^{(r)}) = 0$. Let $(r_n,\ n \geq 1)$ be a sequence such that $r_n \uparrow \infty$ and such that for $n \geq 1$, $d^c_{\mathrm{GHP}}(\mathscr{X}^{(r_n)}, \mathscr{Y}^{(r_n)}) = 0$. Since d^c_{GHP} is a metric on $\mathbb{K}$, there exists a GHP-isometry $\Phi^n : X^{(r_n)} \to Y^{(r_n)}$ for every $n \geq 1$. Since all the $X^{(r)}$

are compact, we may consider, for $n \geq 1$ and for $k \geq 1$, a finite $1/k$-net of $X^{(r_n)}$ containing the root:

$$S_k^n = \left\{ x_{0,k}^n = \emptyset^X, x_{1,k}^n, \ldots, x_{N_k^n,k}^n \right\}.$$

Then, if $k \geq 1$, $n \geq 1$, $0 \leq i \leq N_k^n$, the sequence $(\Phi^j(x_{i,k}^n),\ j \geq n)$ is bounded since the Φ^j are isometries. Using a diagonal procedure, we may assume without loss of generality, that for every $k \geq 1$, $n \geq 1$, $0 \leq i \leq N_k^n$, the sequence $(\Phi^j(x_{i,k}^n),\ j \geq n)$ converges to some limit $y_{i,k}^n \in Y$. We define the map Φ on $S := \bigcup_{n\geq 1,\ k\geq 1} S_k^n$ taking values in Y by:

$$\Phi\left(x_{i,k}^n\right) = y_{i,k}^n.$$

Notice that Φ is an isometry and root preserving as $\Phi(\emptyset^X) = \emptyset^Y$ (see the proof of Lemma 2.15). The set $\Phi(S_k^n)$ is obviously a $2/k$-net of $Y^{(r_n)}$, and thus $\Phi(S)$ is a dense subset of Y. Therefore the map Φ can be uniquely extended into a one-to-one root preserving isometry from X to Y, which we will still denote by Φ. It remains to prove that Φ is a GHP-isometry, that is, such that $\nu = \Phi_* \mu$.

For $n \geq 1$, $k \geq 1$, let $(V_{i,k}^n,\ 0 \leq i \leq N_k^n)$ be Borel subsets of $X^{(r_n)}$ with diameter less than $1/k$, such that $\bigcup_{0\leq i\leq N_k} V_{i,k}^n = X^{(r_n)}$ and for all $0 \leq i, i' \leq N_k$, we have $V_{i,k}^n \cap V_{i',k}^n = \emptyset$ and $x_{i,k}^n \in V_{i,k}^n$ if $V_{i,k}^n \neq \emptyset$. We then define the following measures:

$$\mu_k^n = \sum_{i=0}^{N_k^n} \mu\left(V_{i,k}^n\right) \delta_{x_{i,k}^n} \quad \text{and} \quad \nu_k^n = \sum_{i=0}^{N_k^n} \mu\left(V_{i,k}^n\right) \delta_{y_{i,k}^n}.$$

Let $A \subset X$ be closed. We obviously have $\mu_k^n(A) \leq \mu^{(r_n)}(A^{1/k})$ and $\mu^{(r_n)}(A) \leq \mu_k^n(A^{1/k})$ that is:

$$d_{\mathrm{P}}^X(\mu_k^n, \mu^{(r_n)}) \leq \frac{1}{k}. \tag{2.16}$$

For any $n \geq 1$, $k \geq 1$, we have by construction $\nu_k^n = \Phi_* \mu_k^n$ and $\nu^{(r_n)} = \Phi_*^j \mu^{(r_n)}$ for any $j \geq n \geq 1$. We can then write, for $j \geq n$:

$$\begin{aligned}
d_{\mathrm{P}}^Y\left(\nu_k^n, \nu^{(r_n)}\right) &= d_{\mathrm{P}}^Y\left(\Phi_* \mu_k^n, \Phi_*^j \mu^{(r_n)}\right) \\
&\leq d_{\mathrm{P}}^Y\left(\Phi_* \mu_k^n, \Phi_*^j \mu_k^n\right) + d_{\mathrm{P}}^Y\left(\Phi_*^j \mu_k^n, \Phi_*^j \mu^{(r_n)}\right) \\
&\leq d_{\mathrm{P}}^Y\left(\Phi_* \mu_k^n, \Phi_*^j \mu_k^n\right) + \frac{1}{k},
\end{aligned}$$

where for the last inequality we used $d_{\mathrm{P}}^Y(\Phi_*^j \mu_k^n, \Phi_*^j \mu^{(r_n)}) = d_{\mathrm{P}}^X(\mu_k^n, \mu^{(r_n)})$ and (2.16). Since the two measures $\Phi_* \mu_k^n$ and $\Phi_*^j \mu_k^n$ have the same masses distributed on a finite number of atoms, and the atoms $\Phi^j(x_{i,k}^n)$ of $\Phi_*^j \mu_k^n$ converge towards the atoms $y_{i,k}^n$ of $\Phi_* \mu_k^n$, we deduce that:

$$\lim_{j\to+\infty} d_{\mathrm{P}}^Y\left(\Phi_* \mu_k^n, \Phi_*^j \mu_k^n\right) = 0.$$

Hence, $(\nu_k^n, k \geq 1)$ converges weakly towards $\nu^{(r_n)}$. According to (2.16), the sequence $(\mu_k^n, k \geq 1)$ converges weakly to $\mu^{(r_n)}$. Since we have $\nu_k^n = \Phi_* \mu_k^n$ and Φ is continuous, we get $\nu^{(r_n)} = \Phi_* \mu^{(r_n)}$ for any $n \geq 1$, and thus $\nu = \Phi_* \mu$. This ends the proof. □

We are now ready to prove Proposition 2.8. Notice that we will not use (ii) of Theorem 2.7 in this Section as it is not yet proved.

Proof of Proposition 2.8. By construction, the convergence in $\mathbb{K}\cap\mathbb{L}$ for the d_{GHP} metric implies the convergence for the d^c_{GHP} metric. We only have to prove that the converse is also true.

Let $\mathscr{X}=(X,d^X,\emptyset,\mu)$ and $\mathscr{X}_n=(X_n,d^{X_n},\emptyset_n,\mu_n)$ be elements of $\mathbb{K}\cap\mathbb{L}$ and $(\varepsilon_n,n\in\mathbf{N})$ be a positive sequence converging towards 0 such that, for all $n\in\mathbf{N}$:

$$d^c_{\mathrm{GHP}}(\mathscr{X}_n,\mathscr{X})<\varepsilon_n.$$

Using Lemma 2.13, we consider a metric d^n on the disjoint union $X_n\sqcup X$, such that we have for $n\in\mathbf{N}$, and writing d^n_{H} and d^n_{P} respectively for $d^{d^n}_{\mathrm{H}}$ and $d^{d^n}_{\mathrm{P}}$:

$$d^n(\emptyset_n,\emptyset)+d^n_{\mathrm{H}}(X_n,X)+d^n_{\mathrm{P}}(\mu_n,\mu)<\varepsilon_n.$$

If $x_n\in X_n^{(r)}$, by definition of the Hausdorff metric, there exists $x\in X$ such that $d^n(x_n,x)\le d^n_{\mathrm{H}}(X_n,X)$. Then, we have:

$$d^n(\emptyset,x)\le d^n(\emptyset,\emptyset_n)+d^n(\emptyset_n,x_n)+d^n(x_n,x)\le d^n(\emptyset_n,\emptyset)+r+d^n_{\mathrm{H}}(X_n,X)<r+\varepsilon_n.$$

We get that x belongs to $X^{(r+\varepsilon'_n)}$ for some $\varepsilon'_n<\varepsilon_n$ and thus, according to Lemma 2.16, it belongs to $(X^{(r)})^{\varepsilon_n}$, since X is a complete length space. Therefore we have $X_n^{(r)}\subset(X^{(r)})^{\varepsilon_n}$. Similar arguments yield $X^{(r)}\subset(X_n^{(r)})^{\varepsilon_n}$. We deduce that:

$$d^n_{\mathrm{H}}\left(X_n^{(r)},X^{(r)}\right)\le\varepsilon_n. \tag{2.17}$$

If $A\subset X_n\sqcup X$ is closed, we may compute:

$$\begin{aligned}\mu_n^{(r)}(A)=\mu_n\left(A\cap X_n^{(r)}\right)&\le\mu\left(A^{\varepsilon_n}\cap(X_n^{(r)})^{\varepsilon_n}\right)+\varepsilon_n\\&\le\mu^{(r)}(A^{\varepsilon_n})+\mu\left(\left(X_n^{(r)}\right)^{\varepsilon_n}\setminus X^{(r)}\right)+\varepsilon_n\\&\le\mu^{(r)}(A^{\varepsilon_n})+\mu\left(X^{(r+2\varepsilon_n)}\setminus X^{(r)}\right)+\varepsilon_n,\end{aligned}$$

since $(X_n^{(r)})^{\varepsilon_n}\subset(X^{(r)})^{2\varepsilon_n}\subset X^{(r+2\varepsilon_n)}$. Similarly, we also have:

$$\begin{aligned}\mu^{(r)}(A)&\le\mu\left(A\cap X^{(r-2\varepsilon_n)}\right)+\mu\left(X^{(r)}\setminus X^{(r-2\varepsilon_n)}\right)\\&\le\mu_n\left(A^{\varepsilon_n}\cap\left(X^{(r-2\varepsilon_n)}\right)^{\varepsilon_n}\right)+\mu\left(X^{(r)}\setminus X^{(r-2\varepsilon_n)}\right)+\varepsilon_n\\&\le\mu_n^{(r)}(A^{\varepsilon_n})+\mu\left(X^{(r)}\setminus X^{(r-2\varepsilon_n)}\right)+\varepsilon_n,\end{aligned}$$

since $(X_n^{(r-2\varepsilon_n)})^{\varepsilon_n}\subset X^{(r)}$. Hence, we finally deduce:

$$d^n_{\mathrm{P}}\left(\mu_n^{(r)},\mu^{(r)}\right)\le\varepsilon_n+\mu\left(X^{(r+2\varepsilon_n)}\setminus X^{(r-2\varepsilon_n)}\right).$$

This and (2.17) yields:

$$d^c_{\mathrm{GHP}}(\mathscr{X}_n^{(r)},\mathscr{X}^{(r)})\le 3d^c_{\mathrm{GHP}}(\mathscr{X}_n,\mathscr{X})+\mu\left(X^{(r+2\varepsilon_n)}\setminus X^{(r-2\varepsilon_n)}\right).$$

Therefore, if $\mu(\partial_r X)=0$, we have $\lim_{n\to+\infty}d^c_{\mathrm{GHP}}(\mathscr{X}_n^{(r)},\mathscr{X}^{(r)})=0$. Since μ is by definition a finite measure, the set $\{r>0,\ \mu(\partial_r X)\neq 0\}$ is at most countable. By dominated convergence, we get $\lim_{n\to+\infty}d_{\mathrm{GHP}}(\mathscr{X}_n,\mathscr{X})=0$. $\square$

In order to prove Theorem 2.9 on the pre-compactness criterion, we will approximate the elements of a sequence in $\mathscr{C}$ by nets of small radius. The following lemma guarantees that we can construct such nets in a consistent way. We use the convention that $X^{(r)} = \emptyset$ if $r < 0$. In the sequel, if $r > 0$ and $k \geq 0$, we will often use the notation $A_{r,k}(X)$ for the annulus $X^{(r)} \setminus X^{(r-2^{-k})}$.

Lemma 2.19. *If $\mathscr{X} = (X, \emptyset, d, \mu) \in \mathbb{L}$ satisfies condition (i) of Theorem 2.9, then for any $k, \ell \in \mathbf{N}$, there exists a 2^{-k}-net of the annulus $A_{\ell 2^{-k},k}(X) = X^{(\ell 2^{-k})} \setminus X^{((\ell-1)2^{-k})}$ with at most $N(\ell 2^{-k}, 2^{-k-1})$ elements.*

Proof. Let S' be a finite 2^{-k-1}-net of $X^{(\ell 2^{-k})}$ of cardinal at most $N(\ell 2^{-k}, 2^{-k-1})$. Let S'' be the set of elements x in $S' \cap A_{(\ell-1)2^{-k},k+1}(X)$ such that there exists at least one element, say y_x, in $A_{\ell 2^{-k},k}(X)$ at distance at most 2^{-k-1} of x. The set $\left(S' \cap A_{\ell 2^{-k},k}\right) \bigcup \{y_x,\ x \in S''\}$ is obviously a 2^{-k}-net of $A_{\ell 2^{-k},k}(X)$, and its cardinal is bounded by $N(\ell 2^{-k}, 2^{-k-1})$. □

Proof of Theorem 2.9

Notice that we will not use (ii) of Theorem 2.7 in this Section as it is not yet proved.

The proof will be divided in several parts. The idea, as in [BBI01], is to construct an abstract limit space, along with a measure, and to check that we can get a convergence (up to extraction). Let $(\mathscr{X}_n, n \in \mathbf{N})$ be a sequence in $\mathscr{C}$, with $\mathscr{X}_n = (X_n, d^{X_n}, \emptyset_n, \mu_n)$. For $\ell, k \in \mathbf{N}$, we will write ℓ_k for $\ell 2^{-k}$.

Construction of the limit space.

Let $\ell, k \in \mathbf{N}$. Recall that, by Lemma 2.19, we can consider $\mathfrak{A}^n_{\ell_k,k}$ a 2^{-k-1}-net of the annulus $A_{\ell_k,k}(X_n)$ with at most $N(\ell_k, 2^{-k-2})$ elements. In order to have a finer sequence of nets, we will consider:

$$S^n_{\ell_k,k} = \bigcup_{0 \leq k' \leq k} \left(A_{\ell_k,k}(X_n) \cap \mathfrak{A}^n_{\lceil \ell_k 2^{k'} \rceil 2^{-k'},k'} \right).$$

By construction $S^n_{\ell_k,k}$ is a 2^{-k-1}-net of $A_{\ell_k,k}(X_n)$ with cardinal at most:

$$\bar{N}\left(\ell_k, 2^{-k-2}\right) = \sum_{k'=0}^{k} N\left(\lceil \ell_k 2^{k'} \rceil 2^{-k'}, 2^{-k'-2} \right).$$

Let $U_{\ell_k,k} = \{(k, \ell, i); 0 \leq i \leq \bar{N}(\ell_k, 2^{-k-2})\}$ and $U = \bigcup_{k \in \mathbf{N}, \ell \in \mathbf{N}} U_{\ell_k,k}$. We number the elements of $S^n_{\ell_k,k}$ in such a way that:

$$S^n_{\ell_k,k} \cup \{\emptyset_n\} = \{x^n_u,\ u = (k, \ell, i), u \in U_{\ell_k,k}\}, \tag{2.18}$$

where $(x^n_u, u \in U)$ is some sequence in X_n and $x^n_{(k,\ell,0)} = \emptyset_n$. Notice that $S^n_{\ell_k,k}$ is empty for ℓ_k large if X_n is bounded. For $u, u' \in U$, we set:

$$d^n_{u,u'} = d^{X_n}\left(x^n_u, x^n_{u'}\right).$$

Notice that the sequence $(d^n_{u,u'}, n \in \mathbf{N})$ is bounded. Thus, without loss of generality (by considering the sequence instead of the sub-sequence), we may assume that for all $u, u' \in U$, the sequence $(d^n_{u,u'}, n \geq 1)$ converges in $\mathbf{R}$ to some limit $d_{u,u'}$. We then consider an abstract space, $X' = \{x_u, u \in U\}$. On this space, the function d defined by $(x_u, x_{u'}) \mapsto d_{u,u'}$ is a semi-metric. We then consider the quotient space $X'/\sim$, where $x_u \sim x_{u'}$ if $d_{u,u'} = 0$. We will denote by x_u the equivalent class containing x_u. Notice that $d_{u,u'} = 0$ for any $u = (k, \ell, 0)$ and $u' = (k', \ell', 0)$ elements of U and let $\emptyset$ denote their equivalence class. Finally, we let X be the completion of $X'/\sim$ with respect to the metric d, so that $(X, d, \emptyset)$ is a rooted complete metric space.

Approximation by nets

We set:

$$U^+_{\ell_k,k} = \bigcup_{0\leq j\leq \ell} U_{j2^{-k},k}, \quad S^{n,+}_{\ell_k,k} = \bigcup_{0\leq j\leq \ell} S^n_{j2^{-k},k} = \left\{x^n_u, u \in U^+_{\ell_k,k}\right\} \quad \text{and} \quad S^+_{\ell_k,k} = \left\{x_u, u \in U^+_{\ell_k,k}\right\}.$$

By construction $S^{n,+}_{\ell_k,k}$ is a 2^{-k-1}-net of $X_n^{(\ell_k)}$ and $S^{n,+}_{\ell_k,k} \subset S^{n,+}_{\ell'_{k'},k'}$ as well as $S^+_{\ell_k,k} \subset S^+_{\ell'_{k'},k'}$ for any $k \leq k'$ and $\ell_k \leq \ell'_{k'}$.

Remark 1. *We also have that for $v \in U \backslash U^+_{\ell_k,k}$, either $x^n_v = \emptyset_n$ or $d^{X_n}(\emptyset_n, x^n_v) > \ell_k$ and either $x_v = \emptyset$ or $d(\emptyset, x_v) \geq \ell_k$. Notice that the former inequality is strict but the latter is large.*

A correspondence R between two sets A and B is a subset of $A \times B$ such that the projection of R on A (resp. B) is A (resp. B). It is clear that the set defined by:

$$\mathfrak{R}^{n,+}_{\ell_k,k} = \left\{(x^n_u, x_u), u \in U^+_{\ell_k,k}\right\} \tag{2.19}$$

is a correspondence between $S^{n,+}_{\ell_k,k}$ and $S^+_{\ell_k,k}$. The distorsion $\delta_n(\ell_k, k)$ of this correspondence is defined by:

$$\delta_n(\ell_k, k) = \sup\left\{\left|d^{X_n}\left(x^n_u, x^n_{u'}\right) - d(x_u, x_{u'})\right|;\ u, u' \in U^+_{\ell_k,k}\right\}. \tag{2.20}$$

Notice that for $k \leq k'$ and $\ell_k \leq \ell'_{k'}$, we have:

$$\delta_n(\ell_k, k) \leq \delta_n\left(\ell'_{k'}, k'\right). \tag{2.21}$$

Since $U^+_{\ell_k,k}$ is finite, for all $\ell, k \in \mathbf{N}$, we have by construction $\lim_{n\to+\infty} \delta_n(\ell_k, k) = 0$.

Lemma 2.20. *The set $S^+_{\ell_k,k}$ is a 2^{-k}-net of $X^{(\ell_k)}$.*

Proof. Let $x \in X^{(\ell_k)}$. There exists $v = (k', \ell', j) \in U$ such that $d(x, x_v) < 2^{-k-3}$. Notice that $d(\emptyset, x_v) < \ell_k + 2^{-k-3}$. We may choose n large enough, so that $\delta_n(\ell_k \vee \ell'_{k'}, k \vee k') < 2^{-k-3}$. As $x^n_v \in S^{n,+}_{\ell_k\vee\ell'_{k'},k\vee k'}$, we have $|d^{X_n}(\emptyset_n, x^n_v) - d(\emptyset, x_v)| < 2^{-k-3}$ and thus $d^{X_n}(\emptyset_n, x^n_v) < \ell_k + 2^{-k-2}$. Thanks to Lemma 2.16 and since X_n is a length space, we get that x^n_v belongs to $(X_n^{(\ell_k)})^{2^{-k-2}}$.

As $S^{n,+}_{\ell_k,k}$ is a 2^{-k-1}-net of $X_n^{(\ell_k)}$, there exists $u \in U^+_{\ell_k,k}$ such that $d^{X_n}(x_u^n, x_v^n) < 2^{-k-1} + 2^{-k-2}$. Furthermore, we have that x_u^n and x_v^n belongs to $S^{n,+}_{\ell_k \vee \ell'_{k'}, k\vee k'}$. We deduce that:

$$d(x, x_u) \le d(x, x_v) + d(x_v, x_u) \le 2^{-k-3} + \delta_n\left(\ell_k \vee \ell'_{k'}, k \vee k'\right) + d^{X_n}\left(x_u^n, x_v^n\right) < 2^{-k}.$$

This gives the result. □

We give an immediate consequence of this approximation by nets.

Lemma 2.21. *The metric space (X, d) is a length space.*

Proof. The proof of this lemma is inspired by the proof of Theorem 7.3.25 in [BBI01]. We will check that (X, d) satisfies the mid-point condition.

Let $k \in \mathbf{N}$ and $x, x' \in X$. According to Lemma 2.20, there exists $\ell \in \mathbf{N}$ large enough and $u, u' \in U^+_{\ell_k,k}$ such that $d(x, x_u) < 2^{-k}$ and $d(x', x_{u'}) < 2^{-k}$. For n large enough, we get that $\delta_n(\ell_k, k) < 2^{-k}$. Since (X_n, d^{X_n}) is a length space, there exists $z \in X_n$ such that:

$$\left|2d^{X_n}\left(z, x_u^n\right) - d^{X_n}\left(x_u^n, x_{u'}^n\right)\right| + \left|2d^{X_n}\left(z, x_{u'}^n\right) - d^{X_n}\left(x_u^n, x_{u'}^n\right)\right| \le 2^{-k}.$$

There exists $u'' \in U^+_{\ell_k,k}$ such that $d^{X_n}(x^n_{u''}, z) \le 2^{-k}$. Then, we deduce that:

$$\begin{aligned}
&|2d(x_{u''}, x) - d(x, x')| + |2d(x_{u''}, x') - d(x, x')| \\
&\quad\le 4d(x, x_u) + 4d(x', x_{u'}) + |2d(x_{u''}, x_u) - d(x_u, x_{u'})| + |2d(x_{u''}, x_{u'}) - d(x_u, x_{u'})| \\
&\quad\le 8 \cdot 2^{-k} + 6\delta_n(\ell_k, k) + \left|2d^{X_n}\left(x^n_{u''}, x^n_u\right) - d^{X_n}\left(x^n_u, x^n_{u'}\right)\right| + \left|2d^{X_n}\left(x^n_{u''}, x^n_{u'}\right) - d^{X_n}\left(x^n_u, x^n_{u'}\right)\right| \\
&\quad\le 19 \cdot 2^{-k}.
\end{aligned}$$

Since k is arbitrary, we get that (X, d) satisfies the mid-point condition and thus is a length space. □

Approximation of the measures

Let $(V_u^n, u \in U_{\ell_k,k})$ be Borel subsets of $A_{\ell_k,k}(X_n)$ with diameter less than 2^{-k} such that $\bigcup_{u\in U_{\ell_k,k}} V_u^n = A_{\ell_k,k}(X_n)$ and for all $u, u' \in U_{\ell_k,k}$, we have $V_u^n \bigcap V_{u'}^n = \emptyset$ and $x_u^n \in V_u^n$ as soon as $V_u^n \neq \emptyset$. We set $U_{\infty,k} = \bigcup_{\ell\in\mathbf{N}} U_{\ell_k,k}$ and we consider the following approximation of the measure μ_n:

$$\mu_{n,k} = \sum_{u\in U_{\infty,k}} \mu_n(V_u^n)\delta_{x_u^n}.$$

Notice that $\mu_{n,k}^{(\ell_k)} = \sum_{u\in U_{\ell_k,k}} \mu_n(V_u^n)\delta_{x_u^n}$. The measures $\mu_{n,k}$ are boundedly finite Borel measures on X_n. It is clear that the sequence $(\mu_{n,k}, k \in \mathbf{N})$ converges vaguely towards μ_n as k goes to infinity, since we have for any $r \in \mathbf{N}$, $d_{\mathrm{P}}^{d^{X_n}}(\mu_{n,k}^{(r)}, \mu_n^{(r)}) \le 2^{-k}$. On the limit space X, we define:

$$\nu_{n,k} = \sum_{u\in U_{\infty,k}} \mu_n(V_u^n)\delta_{x_u} \quad \text{and} \quad \nu_{n,k}^{\{\ell_k\}} = \sum_{u\in U_{\ell_k,k}} \mu_n(V_u^n)\delta_{x_u}.$$

Notice that $\nu_{n,k}^{\{\ell_k\}} \leq \nu_{n,k}^{(\ell_k)}$ but they may be distinct as $\nu_{n,k}^{(\ell_k)}$ may have some atoms on $\partial_{\ell_k} X$ which are in $S^+_{(\ell+1)_k,k}$ but not in $S^+_{\ell_k,k}$, as indicated in Remark 1.

Let us show that the sequence $(\nu_{n,k}, k \in \mathbf{N})$ converges, up to an extraction, towards a boundedly finite measure ν on X. For $m \in 2^{-k}\mathbf{N}$, we have:

$$\begin{aligned}\nu_{n,k}(X^{(m)}) = \sum_{u \in U_{\infty,k}} \mu_n(V_u^n) \mathbf{1}_{\{d(x_u,\varnothing) \leq m\}} &\leq \sum_{u \in U_{\infty,k}} \mu_n(V_u^n) \mathbf{1}_{\{d^{X_n}(x_u^n, \varnothing_n) \leq m + \delta_n(m,k)\}} \\ &\leq \mu_n\left(X_n^{(m+\delta_n(m,k)+2^{-k})}\right),\end{aligned} \tag{2.22}$$

where for the first inequality we used (2.20). Recall that for all $\ell, k \in \mathbf{N}$, the sequence $\delta_n(\ell_k, k)$ converges to 0 as $n \to \infty$. We define $\eta_k = \delta_{n_k}(k,k)$. Using a diagonal argument, there exists a sub-sequence $(n_k, k \in \mathbf{N})$ such that:

$$\eta_k \leq 2^{-k}. \tag{2.23}$$

By (2.21), we have $\delta_{n_k}(m,k) \leq \eta_k$ for $k \geq m$. Thanks to property (ii) of Theorem 2.9, we get that $\mu_{n_k}(X_{n_k})^{(m+\delta_{n_k}(m,k)+2^{-k})}$ is uniformly bounded in $k \in \mathbf{N}$ for m fixed. From the classical pre-compactness criterion for vague convergence of boundedly finite measures on a Polish metric space (see Appendix 2.6 of [DVJ03]), we deduce that there exists an extraction of the sub-sequence $(n_k, k \in \mathbf{N})$, which we still note $(n_k, k \in \mathbf{N})$, such that $(\nu_{n_k,k}, k \in \mathbf{N})$ converges vaguely towards some boundedly finite measure ν on X. This implies the weak convergence of the finite measures $(\nu_{n_k,k}^{(r)}, k \in \mathbf{N})$ towards $\nu^{(r)}$ as soon as $\nu(\partial_r X) = 0$. Since ν is boundedly finite, the set

$$A_\nu = \{r \geq 0;\ \nu(\partial_r X) > 0\} \tag{2.24}$$

is at most countable. Thus, we have $\lim_{n \to +\infty} d_{\mathrm{P}}(\nu_{n_k,k}^{(r)}, \nu^{(r)}) = 0$ for almost every $r > 0$.

Convergence in the d_{GHP} metric.

We set $\mathscr{X} = (X, d, \varnothing, \nu)$. Notice that $\mathscr{X} \in \mathbb{L}$ thanks to Lemma 2.21. We will prove that $d_{\mathrm{GHP}}(\mathscr{X}_{n_k}, \mathscr{X})$ converges to 0.

Let $r > 0$. For any $k \in \mathbf{N}$, set $\ell = \lceil 2^k r \rceil$ and recall $\ell_k = 2^{-k}\lceil 2^k r \rceil$. We set:

$$\mathscr{Y}_k^n = \left(S^{n,+}_{\ell_k,k}, d^{X_n}, \varnothing_n, \mu_{n,k}^{(\ell_k)}\right), \quad \mathscr{Z}_k^n = \left(S^+_{\ell_k,k}, d, \varnothing, \nu_{n,k}^{\{\ell_k\}}\right) \quad \text{and} \quad \mathscr{W}_k^n = \left(X^{(\ell_k)}, d, \varnothing, \nu_{n,k}^{\{\ell_k\}}\right).$$

The triangular inequalities give:

$$d^c_{\mathrm{GHP}}(\mathscr{X}_n^{(r)}, \mathscr{X}^{(r)}) \leq B_n^1 + B_n^2 + B_n^3 + B_n^4 + B_n^5 + B_n^6, \tag{2.25}$$

with:

$$\begin{aligned} B_n^1 &= d^c_{\mathrm{GHP}}\left(\mathscr{X}_n^{(r)}, \mathscr{X}_n^{(\ell_k)}\right), \quad B_n^2 = d^c_{\mathrm{GHP}}\left(\mathscr{X}_n^{(\ell_k)}, \mathscr{Y}_k^n\right), \quad B_n^3 = d^c_{\mathrm{GHP}}\left(\mathscr{Y}_k^n, \mathscr{Z}_k^n\right), \\ B_n^4 &= d^c_{\mathrm{GHP}}\left(\mathscr{Z}_k^n, \mathscr{W}_k^n\right), \quad B_n^5 = d^c_{\mathrm{GHP}}\left(\mathscr{W}_k^n, \mathscr{X}^{(\ell_k)}\right), \quad B_n^6 = d^c_{\mathrm{GHP}}\left(\mathscr{X}^{(\ell_k)}, \mathscr{X}^{(r)}\right). \end{aligned}$$

Lemma 2.17 implies that:

$$B_n^1 = d^c_{\mathrm{GHP}}\left(\mathscr{X}_n^{(r)}, \mathscr{X}_n^{(\ell_k)}\right) \leq 2^{-k} + \mu_n\left(X_n^{(\ell_k)} \setminus X_n^{(r)}\right). \tag{2.26}$$

As $S^{n,+}_{\ell_k,k}$ is a 2^{-k-1}-net of $X_n^{\ell_k}$ and by definition of $\mu_{n,k}$, we clearly have:

$$d_{\mathrm{H}}^{d^{X_n}}\left(X_n^{(\ell_k)}, S^{n,+}_{\ell_k,k}\right) \le 2^{-k-1} \quad \text{and} \quad d_{\mathrm{P}}^{d^{X_n}}\left(\mu_n^{(\ell_k)}, \mu_{n,k}\mathbf{1}_{S^{n,+}_{\ell_k,k}}\right) \le 2^{-k}.$$

By considering the identity map from $S^{n,+}_{\ell_k,k}$ to $X^{(\ell_k)}$, we deduce that:

$$B_n^2 = d^c_{\mathrm{GHP}}\left(\mathscr{X}_n^{(\ell_k)}, \mathscr{Y}_k^n\right) \le 2^{-k+1}. \tag{2.27}$$

Recall the correspondence (2.19). It is easy to check that the function defined on the set $\left(S^{n,+}_{\ell_k,k} \sqcup S^+_{\ell_k,k}\right)^2$ by:

$$d_n(y,z) = \begin{cases} d^{X_n}(y,z) & \text{if } y,z \in S^{n,+}_{\ell_k,k},\\ d(y,z) & \text{if } y,z \in S^{+}_{\ell_k,k},\\ \inf\left\{d^{X_n}(y,y') + d(z,z') + \frac{1}{2}\delta_n(\ell_k,k);\ (y',z') \in \mathfrak{R}^{n,+}_{\ell_k,k}\right\} & \text{if } y \in S^{n,+}_{\ell_k,k}, z \in S^{+}_{\ell_k,k} \end{cases} \tag{2.28}$$

is a metric. For this particular metric, we easily have $d_n(\emptyset_n, \emptyset) \le \frac{1}{2}\delta_n(\ell_k,k)$ as well as:

$$d_{\mathrm{H}}^{d_n}\left(S^{n,+}_{\ell_k,k}, S^{+}_{\ell_k,k}\right) \le \frac{1}{2}\delta_n(\ell_k,k) \quad \text{and} \quad d_{\mathrm{P}}^{d_n}\left(\mu_{n,k}^{(\ell_k)}, \nu_{n,k}^{\{\ell_k\}}\right) \le \frac{1}{2}\delta_n(\ell_k,k).$$

We deduce that:

$$B_n^3 = d^c_{\mathrm{GHP}}\left(\mathscr{Y}_k^n, \mathscr{Z}_k^n\right) \le \frac{3}{2}\delta_n(\ell_k,k). \tag{2.29}$$

As $S^+_{\ell_k,k}$ is a 2^{-k}-net of X^{ℓ_k}, thanks to Lemma 2.20, we get:

$$B_n^4 = d^c_{\mathrm{GHP}}\left(\mathscr{Z}_k^n, \mathscr{W}_k^n\right) \le 2^{-k}. \tag{2.30}$$

Concerning B_n^5, we only need to bound the Prokhorov distance between $\nu_{n,k}^{\{\ell_k\}}$ and $\nu_{n,k}^{(\ell_k)}$. Recall that $\nu_{n,k}^{\{\ell_k\}} \le \nu_{n,k}^{(\ell_k)}$ and that $\nu_{n,k}^{(\ell_k)}$ may differ only on $\partial_{\ell_k}X$. For A closed, we have:

$$\nu_{n,k}^{\{\ell_k\}}(A) \le \nu_{n,k}^{(\ell_k)}(A) \quad \text{and} \quad \nu_{n,k}^{(\ell_k)}(A) \le \nu_{n,k}^{\{\ell_k\}}(A) + \nu_{n,k}(\partial_{\ell_k}X).$$

Recall (2.24). Let $\rho(r) \ge r+3$ such that $\rho(r) \not\in A_\nu$ and:

$$\varepsilon_{n,k} = 2d_{\mathrm{P}}\left(\nu_{n,k}^{(\rho(r))}, \nu^{(\rho(r))}\right). \tag{2.31}$$

As $\ell_k \le r + 2^{-k}$, we have:

$$\nu_{n,k}(\partial_{\ell_k}X) \le \nu((\partial_{\ell_k}X)^{\varepsilon_{n,k}}) + \varepsilon_{n,k} \le \nu\left(X^{(r+2^{-k}+\varepsilon_{n,k})} \setminus X^{(r-2\varepsilon_{n,k})}\right) + \varepsilon_{n,k}.$$

We deduce that:

$$B_n^5 = d^c_{\mathrm{GHP}}\left(\mathscr{W}_k^n, \mathscr{X}^{(\ell_k)}\right) \le \nu\left(X^{(r+2^{-k}+\varepsilon_{n,k})} \setminus X^{(r-2\varepsilon_{n,k})}\right) + \varepsilon_{n,k}. \tag{2.32}$$

Lemma 2.17 and the fact that X is a length space gives:

$$B_n^6 = d_{\mathrm{GHP}}^c\left(\mathscr{X}^{(\ell_k)}, \mathscr{X}^{(r)}\right) \le 2^{-k} + \nu\left(X^{(\ell_k)} \setminus X^{(r)}\right). \tag{2.33}$$

Putting (2.26), (2.27), (2.29), (2.30), (2.32), (2.33) in (2.25), we get:

$$\begin{aligned} d_{\mathrm{GHP}}^c(\mathscr{X}_n^{(r)}, \mathscr{X}^{(r)}) \le 5\cdot 2^{-k} + \mu_n\left(X_n^{(\ell_k)} \setminus X_n^{(r)}\right) \\ + \frac{3}{2}\delta_n(\ell_k, k) + \nu\left(X^{(r+2^{-k}+\varepsilon_{n,k})} \setminus X^{(r-2\varepsilon_{n,k})}\right) + \varepsilon_{n,k} + \nu\left(X^{(\ell_k)} \setminus X^{(r)}\right). \end{aligned} \tag{2.34}$$

We give a more precise upper bound for $\mu_n(X_n^{(\ell_k)} \setminus X_n^{(r)})$. Using arguments similar to those used to get (2.22), we have:

$$\begin{aligned} \mu_n\left(X_n^{(\ell_k)} \setminus X_n^{(r)}\right) &\le \mu_n\left(X_n^{(\ell_k)}\right) - \mu_n\left(X_n^{(\ell_k - 2^{-k})}\right) \\ &\le \nu_{n,k}\left(X^{(\ell_k + \delta_n(\ell_k,k) + 2^{-k})}\right) - \nu_{n,k}\left(X^{(\ell_k - \delta_n(\ell_k,k) - 4\cdot 2^{-k})}\right). \end{aligned}$$

For $k \ge r+1$, we have $\delta_n(\ell_k, k) \le \delta_n(k,k)$ thanks to (2.21). Then using the sub-sequence $(n_k, k \in \mathbf{N})$ defined at the end of Section 2.5 with (2.23), we get that:

$$\begin{aligned} \mu_{n_k}\left(X_{n_k}^{(\ell_k)} \setminus X_{n_k}^{(r)}\right) &\le \nu_{n_k,k}\left(X^{(\ell_k + 2\cdot 2^{-k})}\right) - \nu_{n_k,k}\left(X^{(\ell_k - 5\cdot 2^{-k})}\right) \\ &\le \nu\left(X^{(\ell_k + 2\cdot 2^{-k} + \varepsilon_{n_k,k})}\right) - \nu\left(X^{(\ell_k - 5\cdot 2^{-k} - \varepsilon_{n_k,k})}\right) + 2\varepsilon_{n_k,k}. \end{aligned}$$

Notice that the sub-sequence $(n_k, k \in \mathbf{N})$ does not depend on r: it is the same for all $r \ge 0$. Using (2.34), we get for $k \ge r+1$:

$$d_{\mathrm{GHP}}^c(\mathscr{X}_{n_k}^{(r)}, \mathscr{X}^{(r)}) \le 5\cdot 2^{-k} + \frac{3}{2}\eta_k + 2\nu\left(X^{(\ell_k + 2^{-k} + \varepsilon_{n,k})} \setminus X^{(\ell_k - 5\cdot 2^{-k} - 2\varepsilon_{n,k})}\right) + 3\varepsilon_{n_k,k}.$$

As $\lim_{k\to+\infty} \ell_k = r$ and $\lim_{k\to+\infty} \varepsilon_{n_k,k} = 0$, we get using (2.23), that for $r \notin A_\nu$:

$$\lim_{k\to+\infty} d_{\mathrm{GHP}}^c(\mathscr{X}_{n_k}^{(r)}, \mathscr{X}^{(r)}) = 0.$$

By dominated convergence, we get that $\lim_{k\to+\infty} d_{\mathrm{GHP}}(\mathscr{X}_{n_k}, \mathscr{X}) = 0$. Thus we have a converging sub-sequence in $\mathscr{C}$.

Proof of (ii) of Theorem 2.7

We need to prove that the metric space $(\mathbb{L}, d_{\mathrm{GHP}})$ is separable and complete.

Lemma 2.22. *The metric space* $(\mathbb{L}, d_{\mathrm{GHP}})$ *is separable.*

Proof. We can notice that the set $\mathbb{K} \cap \mathbb{L}$ is dense in $(\mathbb{L}, d_{\mathrm{GHP}})$, since for $\mathscr{X} \in \mathbb{L}$, for all $r > 0$ we have $\mathscr{X}^{(r)} \in \mathbb{K}$ and $d_{\mathrm{GHP}}(\mathscr{X}^{(r)}, \mathscr{X}) \le \mathrm{e}^{-r}$. Every element of $\mathbb{K}$ can be approximated in the d_{GHP}^c topology by a sequence of metric spaces with finite cardinal, rational edge-lengths and rational weights. Hence, $(\mathbb{K} \cap \mathbb{L}, d_{\mathrm{GHP}}^c)$ is separable, being a subspace of a separable metric space. According to Proposition 2.8, $(\mathbb{K} \cap \mathbb{L}, d_{\mathrm{GHP}})$ is also separable. As $\mathbb{K} \cap \mathbb{L}$ is dense in $(\mathbb{L}, d_{\mathrm{GHP}})$, we deduce that $(\mathbb{L}, d_{\mathrm{GHP}})$ is separable. $\square$

Lemma 2.23. *The metric space* $(\mathbb{L}, d_{\mathrm{GHP}})$ *is complete.*

Proof. Let $(\mathscr{X}_n, n \in \mathbf{N})$, with $\mathscr{X}_n = (X_n, d^{X_n}, \emptyset_n, \mu_n)$, be a Cauchy sequence in $(\mathbb{L}, d_{\mathrm{GHP}})$. It is enough to prove that it is relatively compact. Thus, we need to prove it satisfies condition (i) and (ii) of Theorem 2.9.

Assume there exists $r_0 \in \mathbf{R}_+$ such that $\sup_{n \in \mathbf{N}} \mu_n(X_n^{(r_0)}) = +\infty$. By considering a subsequence, we may assume that $\lim_{n \to +\infty} \mu_n(X_n^{(r_0)}) = +\infty$. This implies that for any $r \geq r_0$, $\lim_{n \to +\infty} \mu_n(X_n^{(r)}) = +\infty$. Thus, we have for any $m \in \mathbf{N}$:

$$\lim_{n \to +\infty} \int_0^{+\infty} \mathrm{e}^{-r} \left(1 \wedge |\mu_n(X_n^{(r)}) - \mu_m(X_m^{(r)})|\right) dr \geq \mathrm{e}^{-r_0}.$$

Then use (2.14) to get that $(\mathscr{X}_n, n \in \mathbf{N})$ is not a Cauchy sequence. Thus, if $(\mathscr{X}_n, n \in \mathbf{N})$ is a Cauchy sequence, then (ii) of Theorem 2.9 is satisfied.

Let $g_{n,m}(r) = d^c_{\mathrm{GH}}((X_n^{(r)}, d^{X_n^{(r)}}), (X_m^{(r)}, d^{X_m^{(r)}}))$. On the one hand, use (2.15) to get:

$$\lim_{\min(n,m) \to +\infty} \int_0^{+\infty} \mathrm{e}^{-r} \left(1 \wedge g_{n,m}(r)\right) dr = 0. \tag{2.35}$$

On the other hand, using (2.15) and Lemma 2.17, and arguing as in the proof of Lemma 2.6, we get that for any $r, \varepsilon \geq 0$:

$$|g_{n,m}(r) - g_{n,m}(r + \varepsilon)| \leq 2\varepsilon.$$

This implies the functions $g_{n,m}$ are 2-Lipschitz. Thus, we deduce from (2.35), that for all $r \geq 0$, $\lim_{\min(n,m) \to +\infty} g_{n,m}(r) = 0$. Thus the sequence $((X_n^{(r)}, d^{X_n^{(r)}}), n \in \mathbf{N})$ is a Cauchy sequence for the Gromov-Hausdorff metric. Then point (2) of Proposition 7.4.11 in [BBI01] readily implies condition (i) of Theorem 2.9. $\square$

CHAPTER 3

Pathwise construction of the Lévy tree-valued pruning process

3.1 Introduction

Lévy trees arise as a natural generalization to the continuum trees defined by Aldous [Ald91a]. They are located at the intersection of several important fields: combinatorics of large discrete trees, Lévy processes and branching processes. Consider a branching mechanism ψ, that is a function of the form

$$\psi(\lambda) = \alpha\lambda + \beta\lambda^2 + \int_{(0,+\infty)} \left(\mathrm{e}^{-\lambda x} - 1 + \lambda x \mathbf{1}_{\{x<1\}}\right) \Pi(dx) \tag{3.1}$$

with $\alpha \in \mathbf{R}$, $\beta \geq 0$, Π a Lévy measure such that $\int_{(0,+\infty)} 1 \wedge x^2 \, \Pi(dx) < +\infty$. In the (sub)critical case $\psi'(0) \geq 0$, Le Gall and Le Jan [LL98b] defined a ***continuum*** tree structure, which can be described by a tree $\mathscr{T}$, for the genealogy of a population whose size is given by a CSBP with branching mechanism ψ. We shall consider the distribution $\mathbb{P}_r^\psi(d\mathscr{T})$ of this Lévy tree when the CSBP starts at mass $r > 0$, or its excursion measure $\mathbb{N}^\psi[d\mathscr{T}]$, when the CSBP is distributed under its canonical measure. The ψ-Lévy tree possesses several striking features as pointed out in the works of Duquesne and Le Gall [DL02, DL05]. For instance, the branching nodes can only be of degree 3 (binary branching) if $\beta > 0$ or of infinite degree (when removing the branching point, the tree is separated in infinitely many connected components) if $\Pi \neq 0$. Furthermore, there exists a mass measure $\mathbf{m}^{\mathscr{T}}$ on the leaves of $\mathscr{T}$, whose total mass corresponds to the total population size $\sigma = \mathbf{m}^{\mathscr{T}}(\mathscr{T})$ of the CSBP. We shall also consider the extinction time of the CSBP which corresponds to the height $H_{\max}(\mathscr{T})$ of the tree $\mathscr{T}$. The results can be extended to the super-critical case, using a Girsanov transformation given by Abraham and Delmas [AD12a].

In [AD12a], a decreasing continuum tree-valued process is defined using the so-called pruning procedure of Lévy trees introduced in Abraham, Delmas and Voisin [ADV10]. By marking a ψ-Lévy tree with two different kinds of marks (the first ones lying on the skeleton of the tree, the other ones on the nodes of infinite degree), one can prune the tree by throwing away all the points having a mark on their ancestral line connecting them to the root. The main result of [ADV10] is that the remaining tree is still a Lévy tree, with branching mechanism related to ψ. The idea of [AD12a] is to consider a particular pruning with an intensity depending on a parameter θ, so that the corresponding branching mechanism ψ_θ is ψ shifted by θ:

$$\psi_\theta(\lambda) = \psi(\theta + \lambda) - \psi(\theta).$$

Letting θ vary enables to define a decreasing tree-valued Markov process $(\mathcal{T}_\theta, \theta \in \Theta^\psi)$, with $\Theta^\psi \subset \mathbf{R}$ the set of θ for which ψ_θ is well-defined, and such that $\mathcal{T}_\theta$ is distributed according to $\mathbb{N}^{\psi_\theta}$. If we write $\sigma_\theta = \mathbf{m}^{\mathcal{T}_\theta}(\mathcal{T}_\theta)$ for the total mass of $\mathcal{T}_\theta$, then the process $(\sigma_\theta, \theta \in \Theta^\psi)$ is a pure-jump process. The case $\Pi = 0$ was studied by Aldous and Pitman [AP98a]. The time-reversed tree-valued process is also a Markov process which defines a growing tree process. Let us mention that the same kind of ideas have been used by Aldous and Pitman [AP98b] and by Abraham, Delmas and He [ADH12a] in the framework of Galton-Watson trees to define growing discrete tree-valued Markov processes.

In the discrete framework of [ADH12a], it is possible to define the infinitesimal transition rates of the growing tree process. In [EW06], Evans and Winter define another continuum tree-valued process using a prune and re-graft procedure. This process is reversible with respect to the law of Aldous's continuum random tree and its infinitesimal transitions are described using the theory of Dirichlet forms.

In this paper, we describe the infinitesimal behavior of the growing continuum tree-valued process, that is of $(\mathcal{T}_\theta, \theta \in \Theta^\psi)$ seen backwards in time. The Special Markov Property in [ADV10] describes only two-dimensional distributions and hence the transition probabilities but, since the space of real trees is not locally compact, we cannot use the theory of infinitesimal generators to describe its infinitesimal transitions. Dirichlet forms cannot be used either since the process is not symmetric (it is increasing). However, it is a pure-jump process and our first main result shows that the infinitesimal transitions of the process can be described using a random point process of trees which are grafted one by one on the leaves of the growing tree. More precisely, let $\{\theta_j; j \in J\}$ be the set of jumping times of the mass process $(\sigma_\theta, \theta \in \Theta^\psi)$. Then, informally, at time θ_j, a tree $\mathcal{T}^j$ distributed according to $\mathbf{N}^{\psi_{\theta_j}}[\mathcal{T} \in \bullet]$, with:

$$\mathbf{N}^{\psi_\theta}[\mathcal{T} \in \bullet] = 2\beta \mathbb{N}^{\psi_\theta}[\mathcal{T} \in \bullet] + \int_{(0,+\infty)} \Pi(dr) r e^{-\theta r} \mathbb{P}_r^{\psi_\theta}(\mathcal{T} \in \bullet),$$

is grafted at x_j, a leaf of $\mathcal{T}_{\theta_j}$ chosen at random (according to the mass measure $\mathbf{m}^{\mathcal{T}_{\theta_j}}$). We also prove that the random point measure

$$\mathcal{N} = \sum_{j \in J} \delta_{(x_j, \mathcal{T}^j, \theta_j)}$$

has predictable compensator:

$$\mathbf{m}^{\mathcal{T}_\theta}(dx) \mathbf{N}^{\psi_\theta}[d\mathcal{T}]\, \mathbf{1}_{\{\theta \in \Theta^\psi\}}\, d\theta$$

with respect to the backwards in time natural filtration of the process. See Corollary 3.28 for a precise statement.

Notice that the precise statement relies on the introduction of the set of locally compact weighted real trees endowed with a Gromov-Hausdorff-Prohorov distance. Therefore, we will assume that Lévy trees are locally compact, which corresponds to the Grey condition: $\int^{+\infty} \frac{du}{\psi(u)} < \infty$. In the (sub)critical case this implies that the corresponding height process of the Lévy tree is continuous and that the tree is compact. However, the tree-valued process is defined in [ADV10] without this assumption and we conjecture that the jump representation of the tree-valued Markov process holds without this assumption.

The representation using the random point measure allows to describe the ascension time or explosion time (when it is defined):

$$A = \inf\{\theta \in \Theta^{\psi},\ \sigma_\theta < \infty\}$$

as $\inf\{\theta_j, \mathbf{m}^{\mathcal{T}^j}(\mathcal{T}^j) < \infty\}$, the first time (backwards in time) at which a tree with infinite mass is grafted. This representation is also used in Abraham and Delmas [AD11, AD12b] respectively on the asymptotics of the records on discrete subtrees of the continuum random tree and on the study of the record process in general Lévy trees.

This structure, somewhat similar to the Poissonian structure of the jumps of a Lévy process (although in our case the structure is neither homogeneous nor independent), enables us to study the exit time of first passage of the growing tree-valued process above a given height:

$$A_h = \sup\{\theta \in \Theta^{\psi},\ H_{\max}(\mathcal{T}_\theta) > h\}.$$

We give the joint distribution of the ascension time and the exit time (A, A_h), see Proposition 3.31. In particular, A_h goes to A as h goes to infinity: for h very large, with high probability the process up to A will not have crossed height h, so that the first jump to cross height h will correspond to the grafting time of the first infinite tree, which happens at ascension time A.

We also give in Theorem 3.33 the joint distribution of $(\mathcal{T}_{A_h-}, \mathcal{T}_{A_h})$ the tree just after and just before the jumping time A_h. And we give a spinal decomposition of $\mathcal{T}_{A_h}$ along the ancestral branch of the leaf on which the overshooting tree is grafted, which is similar to the classical Bismut decomposition of Lévy trees. Conditionally on this ancestral branch, the overshooting tree is then distributed as a regular Lévy tree, conditioned on being high enough to perform the overshooting. This generalizes results in [AD12a] about the ascension time of the tree-valued process. Notice that this approach could easily be generalized to study spatial exit times of growing families of super-Brownian motions.

All the results of this paper are stated in terms of real trees and not in terms of the height process or the exploration process that encode the tree as in [ADV10]. For this purpose, we define in Section 3.2 the state space of rooted real trees with a mass measure (called here weighted trees or w-trees) endowed with the so-called Gromov-Hausdorff-Prohorov metric defined in Abraham, Delmas and Hoscheit [ADH12b] which is a slight generalization of the Gromov-Hausdorff metric on the space of metric spaces, and also a generalization of the Gromov-Prohorov topology of [GPW08] on the space of compact metric spaces endowed with a probability measure.

The paper is organized as follows. In Section 3.2, we introduce all the material for our study: the state space of weighted real trees and the metric on it, see Section 3.2 ; the definition of sub(critical) Lévy trees via the height process ; the extension of the definition to super-critical Lévy trees ; the pruning procedure of Lévy trees. In Section 3.3, we recall the definition of the growing tree-valued process by the pruning procedure as in [ADV10] in the setting of real trees and give another construction using the grafting of trees given by random point processes. We prove in Theorem 3.26 that the two definitions agree and then give in Corollary 3.28 the random Point measure description. Section 3.4 is devoted to the application of this construction on the distribution of the tree at the times it overshoots a given height and just before, see Theorem 3.33.

3.2 The pruning of Lévy trees

Real trees

The first definitions of continuum random trees go back to Aldous [Ald91a]. Later, Evans, Pitman and Winter [EPW05] used the framework of real trees, previously used in the context of geometric group theory, to describe continuum trees. We refer to [Eva08, Le 06] for a general presentation of random real trees. Informally, real trees are metric spaces without loops, locally isometric to the real line.

More precisely, a metric space (T,d) is a *real tree* (or $\mathbf{R}$-*tree*) if the following properties are satisfied:

1. For every $s,t\in T$, there is a unique isometric map $f_{s,t}$ from $[0,d(s,t)]$ to T such that $f_{s,t}(0)=s$ and $f_{s,t}(d(s,t))=t$.
2. For every $s,t\in T$, if q is a continuous injective map from $[0,1]$ to T such that $q(0)=s$ and $q(1)=t$, then $q([0,1])=f_{s,t}([0,d(s,t)])$.

We say that a real tree is *rooted* if there is a distinguished vertex $\emptyset$, which will be called the *root* of T. Such a real tree is noted $(T,d,\emptyset)$. If $s,t\in T$, we will note $[\![s,t]\!]$ the range of the isometric map $f_{s,t}$ described above. We will also note $[\![s,t[\![$ for the set $[\![s,t]\!]\setminus\{t\}$. We give some vocabulary on real trees, which will be used constantly when dealing with Lévy trees. Let T be a real tree. If $x\in T$, we shall call *degree* of x, and note by $n(x)$, the number of connected components of the set $T\setminus\{x\}$. In a general tree, this number can be infinite, and this will actually be the case with Lévy trees. The set of *leaves* is defined as:

$$\mathrm{Lf}(T)=\{x\in T\backslash\{\emptyset\},\ n(x)=1\}.$$

If $n(x)\geq 3$, we say that x is a *branching point.* The set of branching points will be noted $\mathrm{Br}(T)$. Among those, there is the set of *infinite* branching points, defined by

$$\mathrm{Br}_\infty(T)=\{x\in\mathrm{Br}(T),\ n(x)=\infty\}.$$

Finally, the *skeleton* of a real tree, noted $\mathrm{Sk}(T)$, is the set of points in the tree that aren't leaves. It should be noted, following Evans, Pitman and Winter [EPW05], that the trace of the Borel σ-field of T on $\mathrm{Sk}(T)$ is generated by the sets $[\![s,s']\!]$, $s,s'\in\mathrm{Sk}(T)$. Hence, it is possible to define a σ-finite Borel measure l^T on T, such that

$$l^T(\mathrm{Lf}(T))=0\quad\text{and}\quad l^T([\![s,s']\!])=d(s,s').$$

This measure will be called *length measure* on T. If x, y are two points in a rooted real tree $(T, d, \emptyset)$, then there is a unique point $z \in T$, called the Most Recent Common Ancestor (MRCA) of x and y such that $[\![\emptyset, x]\!] \cap [\![\emptyset, y]\!] = [\![\emptyset, z]\!]$. This vocabulary is an illustration of the genealogical vision of real trees, in which the root is seen as the ancestor of the population represented by the tree. Similarly, if $x \in T$, we shall call *height* of x, and note by H_x the distance $d(\emptyset, x)$ to the root. The function $x \mapsto H_x$ is continuous on T, and we define the height of T:

$$H_{\max}(T) = \sup_{x \in T} H_x.$$

Gromov-Prohorov metric

Rooted weighted metric spaces

This section is inspired by [DW07], but for the fact that we include measures on the trees, in the spirit of [Mie09]. The detailed proofs of the results stated in this Section are in [ADH12b].

Let (X, d^X) be a Polish metric space. For $A, B \in \mathscr{B}(X)$, we set:

$$d_{\mathrm{H}}^X(A, B) = \inf\{\varepsilon > 0,\ A \subset B^\varepsilon \text{ and } B \subset A^\varepsilon\},$$

the Hausdorff distance between A and B, where $A^\varepsilon = \{x \in X, \inf_{y \in A} d^X(x, y) < \varepsilon\}$ is the ε-halo set of A. If X is compact, then the space of compact subsets of X, endowed with the Hausdorff distance, is compact, see theorem 7.3.8 in [BBI01].

Recall that a Borel measure is locally finite if the measure of any bounded Borel set is finite. We will use the notation $\mathscr{M}_f(X)$ for the space of all finite Borel measures on X. If $\mu, \nu \in \mathscr{M}_f(X)$, we set:

$$d_{\mathrm{P}}^X(\mu, \nu) = \inf\{\varepsilon > 0,\ \mu(A) \leq \nu(A^\varepsilon) + \varepsilon \text{ and } \nu(A) \leq \mu(A^\varepsilon) + \varepsilon \ \text{ for all closed set } A\},$$

the Prohorov distance between μ and ν. It is well known that $(\mathscr{M}_f(X), d_{\mathrm{P}}^X)$ is a Polish metric space, and that the topology generated by d_P^X is exactly the topology of weak convergence (convergence against continuous bounded functionals).

If $\Phi : X \to X'$ is a Borel map between two Polish metric spaces and if μ is a Borel measure on X, we will note $\Phi_* \mu$ the image measure on X' defined by $\Phi_* \mu(A) = \mu(\Phi^{-1}(A))$, for any Borel set $A \subset X$.

Definition 3.1.

- *A* rooted weighted metric space $\mathscr{X} = (X, d^X, \emptyset^X, \mu^X)$ *is a metric space* (X, d^X) *with a distinguished element* $\emptyset^X \in X$ *and a locally finite Borel measure* μ^X.
- *Two rooted weighted metric spaces* $\mathscr{X} = (X, d^X, \emptyset^X, \mu^X)$ *and* $\mathscr{X}' = (X', d^{X'}, \emptyset^{X'}, \mu^{X'})$ *are said* GHP-isometric *if there exists an isometric bijection* $\Phi : X \to X'$ *such that* $\Phi(\emptyset^X) = \emptyset^{X'}$ *and* $\Phi_* \mu^X = \mu^{X'}$.

Notice that if (X, d^X) is compact, then a locally finite measure on X is finite and belongs to $\mathscr{M}_f(X)$. We will now use a procedure due to Gromov [Gro07] to compare any two compact rooted weighted metric spaces, even if they are not subspaces of the same Polish metric space.

Gromov-Hausdorff-Prohorov distance for compact metric spaces

Let $\mathscr{X} = (X, d, \emptyset, \mu)$ and $\mathscr{X}' = (X', d', \emptyset', \mu')$ be two compact rooted weighted metric spaces, and define:

$$d^c_{\mathrm{GHP}}(\mathscr{X}, \mathscr{X}') = \inf_{\Phi, \Phi', Z} \left(d^Z_{\mathrm{H}}(\Phi(X), \Phi'(X')) + d^Z(\Phi(\emptyset), \Phi'(\emptyset')) + d^Z_{\mathrm{P}}(\Phi_* \mu, \Phi'_* \mu') \right), \tag{3.2}$$

where the infimum is taken over all isometric embeddings $\Phi : X \hookrightarrow Z$ and $\Phi' : X' \hookrightarrow Z$ into some common Polish metric space (Z, d^Z).

Note that equation (3.2) does not actually define a distance function, as $d^c_{\mathrm{GHP}}(\mathscr{X}, \mathscr{X}') = 0$ if $\mathscr{X}$ and $\mathscr{X}'$ are GHP-isometric. Therefore, we shall consider $\mathbb{K}$, the set of GHP-isometry classes of compact rooted weighted metric space and identify a compact rooted weighted metric space with its class in $\mathbb{K}$. Then the function d^c_{GHP} is finite on $\mathbb{K}^2$.

Theorem 3.2. *The function d^c_{GHP} defines a metric on $\mathbb{K}$ and the space $(\mathbb{K}, d^c_{GHP})$ is a Polish metric space.*

We shall call d^c_{GHP} the *Gromov-Hausdorff-Prohorov* metric. This extends the Gromov-Hausdorff metric on compact metric spaces, see [BBI01] section 7, as well as the Gromov-Hausdorff-Prohorov metric on compact metric spaces endowed with a probability measure, see [Mie09]. See also [GPW08] for an other approach on metric spaces endowed with a probability measure.

Gromov-Hausdorff-Prohorov distance

However, the definition of Gromov-Hausdorff-Prohorov distance on compact metric space is not yet general enough, as we want to deal with unbounded trees with σ-finite measures. To consider such an extension, we shall consider complete and locally compact length spaces.

We recall that a metric space (X, d) is a *length space* if for every $x, y \in X$, we have:

$$d(x, y) = \inf L(\gamma),$$

where the infimum is taken over all rectifiable curves $\gamma : [0, 1] \to X$ such that $\gamma(0) = x$ and $\gamma(1) = y$, and where $L(\gamma)$ is the length of the rectifiable curve γ.

Definition 3.3. *Let $\mathbb{L}$ be the set of GHP-isometry classes of rooted weighted complete and locally compact length spaces and identify a rooted weighted complete and locally compact length spaces with its class in $\mathbb{L}$.*

If $\mathscr{X} = (X, d, \emptyset, \mu) \in \mathbb{L}$, then for $r \geq 0$ we will consider its restriction to the ball of radius r centered at $\emptyset$, $\mathscr{X}^{(r)} = (X^{(r)}, d^{(r)}, \emptyset, \mu^{(r)})$, where

$$X^{(r)} = \{x \in X; d(\emptyset, x) \leq r\},$$

the metric $d^{(r)}$ is the restriction of d to $X^{(r)}$, and the measure $\mu^{(r)}(dx) = \mathbf{1}_{X^{(r)}}(x)\, \mu(dx)$ is the restriction of μ to $X^{(r)}$. Recall the Hopf-Rinow theorem implies that if (X, d) is a complete and locally compact length space, then every closed bounded subset of X is compact. In particular if $\mathscr{X}$ belongs to $\mathbb{L}$, then $\mathscr{X}^{(r)}$ belongs to $\mathbb{K}$ for all $r \geq 0$.

We state a regularity Lemma of d^c_{GHP} with respect to the restriction operation.

Lemma 3.4. *Let $\mathscr{X}$ and $\mathscr{Y}$ belong to $\mathbb{L}$. Then the function defined on $\mathbf{R}_+$:*

$$r \mapsto d^c_{GHP}\big(\mathscr{X}^{(r)}, \mathscr{Y}^{(r)}\big)$$

is càdlàg.

This implies that the following function is well defined on $\mathbb{L}^2$:

$$d_{\text{GHP}}(\mathscr{X}, \mathscr{Y}) = \int_0^\infty \mathrm{e}^{-r} \left(1 \wedge d^c_{\text{GHP}}\left(\mathscr{X}^{(r)}, \mathscr{Y}^{(r)}\right)\right) \, dr.$$

Theorem 3.5. *The function d_{GHP} defines a metric on $\mathbb{L}$ and the space $(\mathbb{L}, d_{GHP})$ is a Polish metric space.*

The next result implies that d^c_{GHP} and d_{GHP} defines the same topology on $\mathbb{K} \cap \mathbb{L}$.

Theorem 3.6. *Let $(\mathscr{X}_n, n \in \mathbb{N})$ and $\mathscr{X}$ be elements of $\mathbb{K} \cap \mathbb{L}$. Then the sequence $(\mathscr{X}_n, n \in \mathbb{N})$ converges to $\mathscr{X}$ in $(\mathbb{K}, d^c_{GHP})$ if and only if it converges to $\mathscr{X}$ in $(\mathbb{L}, d_{GHP})$.*

The space of w-trees

Note that real trees are always length spaces and that complete real trees are the only complete connected spaces that satisfy the so-called *four-point condition*:

$$\forall x_1, x_2, x_3, x_4 \in X,\ d(x_1, x_2) + d(x_3, x_4) \leq (d(x_1, x_3) + d(x_2, x_4)) \vee (d(x_1, x_4) + d(x_2, x_3)). \quad (3.3)$$

Definition 3.7. *We denote by $\mathbb{T}$ be the set of (GHP-isometry classes of) complete locally compact rooted real trees endowed with a locally finite Borel measure, in short* w-trees.

We deduce the following Corollary from Theorem 3.5 and the four-point condition characterization of real trees.

Corollary 3.8. *The set $\mathbb{T}$ is a closed subset of $\mathbb{L}$ and $(\mathbb{T}, d_{GHP})$ is a Polish metric space.*

Height erasing

We define the restriction operators on the space of w-trees. Let $a \geq 0$. If $(T, d, \emptyset, \mathbf{m})$ is a w-tree, let

$$\pi_a(T) = \{x \in T,\ d(\emptyset, x) \leq a\} \quad (3.4)$$

and $(\pi_a(T), d^{\pi_a(T)}, \emptyset, \mathbf{m}^{\pi_a(T)})$ be the w-tree constituted of the points of T having height lower than a, where $d^{\pi_a(T)}$ and $\mathbf{m}^{\pi_a(T)}$ are the restrictions of d and $\mathbf{m}$ to $\pi_a(T)$. When there is no confusion, we will also write $\pi_a(T)$ for $(\pi_a(T), d^{\pi_a(T)}, \emptyset, \mathbf{m}^{\pi_a(T)})$. We will also write $T(a) = \{x \in T,\ d(\emptyset, x) = a\}$ for the level set at height a. We say a w-tree T is bounded if $\pi_a(T) = T$ for some finite a. Notice that a tree T is bounded if and only if $H_{\max}(T)$ is finite.

Grafting procedure

We will define in this section a procedure by which we add (graft) w-trees on an existing w-tree. More precisely, let $T \in \mathbb{T}$ and let $((T_i, x_i), i \in I)$ be a finite or countable family of elements of $\mathbb{T} \times T$. We define the real tree obtained by grafting the trees T_i on T at point x_i. We set $\tilde{T} = T \sqcup \left(\bigsqcup_{i \in I} T_i \backslash \{\emptyset^{T_i}\} \right)$ where the symbol $\sqcup$ means that we choose for the sets T and $(T_i)_{i \in I}$ representatives of isometry classes in $\mathbb{T}$ which are disjoint subsets of some common set and that we perform the disjoint union of all these sets. We set $\emptyset^{\tilde{T}} = \emptyset^T$. The set $\tilde{T}$ is endowed with the following metric $d^{\tilde{T}}$: if $s, t \in \tilde{T}$,

$$d^{\tilde{T}}(s,t) = \begin{cases} d^T(s,t) & \text{if } s,t \in T, \\ d^T(s,x_i) + d^{T_i}\left(\emptyset^{T_i}, t\right) & \text{if } s \in T,\ t \in T_i \backslash \{\emptyset^{T_i}\}, \\ d^{T_i}(s,t) & \text{if } s,t \in T_i \backslash \{\emptyset^{T_i}\}, \\ d^T(x_i,x_j) + d^{T_j}\left(\emptyset^{T_j}, s\right) + d^{T_i}\left(\emptyset^{T_i}, t\right) & \text{if } i \neq j \text{ and } s \in T_j \backslash \{\emptyset^{T_j}\},\ t \in T_i \backslash \{\emptyset^{T_i}\}. \end{cases}$$

We define the mass measure on $\tilde{T}$ by:

$$\mathbf{m}^{\tilde{T}} = \mathbf{m}^T + \sum_{i \in I} \mathbf{1}_{T_i \backslash \{\emptyset^{T_i}\}} \mathbf{m}^{T_i} + \mathbf{m}^{T_i}\left(\{\emptyset^{T_i}\}\right) \delta_{x_i},$$

where δ_x is the Dirac mass at point x. It is clear that the metric space $(\tilde{T}, d^{\tilde{T}}, \emptyset^{\tilde{T}})$ is still a rooted complete real tree. However, it is not always true that $\tilde{T}$ remains locally compact (it still remains a length space anyway), or, for that matter, that $\mathbf{m}^{\tilde{T}}$ defines a locally finite measure (on $\tilde{T}$). So, we will have to check that $(\tilde{T}, d^{\tilde{T}}, \emptyset^{\tilde{T}}, \mathbf{m}^{\tilde{T}})$ is a w-tree in the particular cases we will consider.

We will use the following notation:

$$\left(\tilde{T}, d^{\tilde{T}}, \emptyset^{\tilde{T}}, \mathbf{m}^{\tilde{T}}\right) = T \circledast_{i \in I} (T_i, x_i) \tag{3.5}$$

and write $\tilde{T}$ instead of $(\tilde{T}, d^{\tilde{T}}, \emptyset^{\tilde{T}}, \mathbf{m}^{\tilde{T}})$ when there is no confusion.

Real trees coded by functions

Lévy trees are natural generalizations of Aldous's Brownian tree, where the underlying process coding for the tree (reflected Brownian motion in Aldous's case) is replaced by a certain functional of a Lévy process, the ***height process***. Le Gall and Le Jan [LL98b] and Duquesne and Le Gall [DL05] showed how to generate random real trees using the excursions of a Lévy process. We shall briefly recall this construction, in order to introduce the pruning procedure on Lévy trees. Let us first work in a deterministic setting.

Let f be a continuous non-negative function defined on $[0, +\infty)$, such that $f(0) = 0$, with compact support. We set:

$$\sigma^f = \sup\{t; f(t) > 0\},$$

with the convention $\sup \emptyset = 0$. Let d^f be the non-negative function defined by:

$$d^f(s,t) = f(s) + f(t) - 2 \inf_{u \in [s \wedge t, s \vee t]} f(u).$$

It can be easily checked that d^f is a semi-metric on $[0,\sigma^f]$. One can define the equivalence relation associated to d^f by $s \sim t$ if and only if $d^f(s,t)=0$. Moreover, when we consider the quotient space

$$T^f = \left[0,\sigma^f\right]_{/\sim}$$

and, noting again d^f the induced metric on T^f and rooting T^f at $\emptyset^f$, the equivalence class of 0, it can be checked that the space $(T^f, d^f, \emptyset^f)$ is a compact rooted real tree. We denote by p^f the canonical projection from $[0,\sigma^f]$ onto T^f, which is extended by $p^f(t)=\emptyset^f$ for $t \geq \sigma^f$. Notice that p^f is continuous. We define $\mathbf{m}^f$, the mass measure on T^f as the image measure on T^f of the Lebesgue measure on $[0,\sigma^f]$ by p^f. We consider the (compact) w-tree $(T^f, d^f, \emptyset^f, \mathbf{m}^f)$, which we shall denote T^f.

It should be noted that, if $x \in T^f$ is an equivalence class, the common value of f on all the points in this equivalence class is exactly $d^f(\emptyset, x) = H_x$. Notice that, in this setting, $H_{\max}(T^f) = \|f\|_\infty$ where $\|f\|_\infty$ stands for the uniform norm of f.

We have the following elementary result (see Lemma 2.3 of [DL05] when dealing with the Gromov-Hausdorff metric instead of the Gromov-Hausdorff-Prohorov metric).

Proposition 3.9. *Let f,g be two compactly supported, non-negative continuous functions with $f(0)=g(0)=0$. Then:*

$$d^c_{GHP}\left(T^f, T^g\right) \leq 6\|f-g\|_\infty + \left|\sigma^f - \sigma^g\right|. \tag{3.6}$$

Proof. The Gromov-Hausdorff distance can be evaluated using correspondences, see [BBI01], section 7.3. A correspondence between two metric spaces (E_1,d_1) and (E_2,d_2) is a subset $\mathfrak{R}$ of $E_1 \times E_2$ such that for $\delta \in \{1,2\}$ the projection of $\mathfrak{R}$ on E_δ is onto: $\{x_\delta; (x_1,x_2) \in \mathfrak{R}\} = E_\delta$. The distortion of $\mathfrak{R}$ is defined by:

$$\mathrm{dis}(\mathfrak{R}) = \sup\left\{|d_1(x_1,y_1) - d_2(x_2,y_2)|;\ (x_1,y_1) \in \mathfrak{R}, (x_2,y_2) \in \mathfrak{R}\right\}.$$

Let $Z = E_1 \sqcup E_2$ by the disjoint union of E_1 and E_2 and consider the function d^Z defined on Z^2 by: $d^Z = d_\delta$ on E_δ^2 for $\delta \in \{1,2\}$ and for $x_1 \in E_1$, $x_2 \in E_2$:

$$d^Z(x_1,x_2) = \inf\left\{d_1(x_1,y_1) + \frac{1}{2}\mathrm{dis}(\mathfrak{R}) + d_2(y_2,x_2);\ (y_1,y_2) \in \mathfrak{R}\right\}.$$

Then if $\mathrm{dis}(\mathfrak{R}) > 0$, the function d^Z is a metric on Z. And we have:

$$d_H^Z(E_1,E_2) \leq \frac{1}{2}\mathrm{dis}(\mathfrak{R}).$$

Let f,g be compactly supported, non-negative continuous functions with $f(0)=g(0)=0$. Following [DL05], we consider the following correspondence between $\mathcal{T}^f$ and $\mathcal{T}^g$:

$$\mathfrak{R} = \left\{(x^f, x^g);\ x^f = p^f(t) \text{ and } x^g = p^g(t) \text{ for some } t \geq 0\right\},$$

and we have $\mathrm{dis}(\mathfrak{R}) \leq 4\|f-g\|_\infty$ according to the proof of Lemma 2.3 in [DL05]. Notice $(\emptyset^f, \emptyset^g) \in \mathfrak{R}$. Thus, with the notation above and $E_1 = T^f$, $E_2 = T^g$, we get:

$$d_H^Z(T^f, T^g) \leq 2\|f-g\|_\infty \quad \text{and} \quad d^Z(\emptyset^f, \emptyset^g) \leq 2\|f-g\|_\infty.$$

Then, we consider the Prohorov distance between $\mathbf{m}^f$ and $\mathbf{m}^g$. Let A^f be a Borel set of T^f. We set $I = \{t \in [0, \sigma^f];\ p^f(t) \in A\}$. By definition of $\mathbf{m}^f$, we have $\mathbf{m}^f(A^f) = \mathrm{Leb}(I)$. We set $A^g = p^g(I \cap [0, \sigma^g])$ so that $\mathbf{m}^g(A^g) = \mathrm{Leb}(I \cap [0, \sigma^g]) \geq \mathrm{Leb}(I) - |\sigma^f - \sigma^g|$. By construction, we also have that for any $x^g \in A^g$, there exists $t \in I$ such that $p^g(t) = x^g$ and such that $d^Z(x^g, x^f) = \frac{1}{2}\mathrm{dis}(\mathfrak{R})$, with $x^f = p^f(t) \in A^f$. This implies that $A^g \subset (A^f)^r$ for any $r > \frac{1}{2}\mathrm{dis}(\mathfrak{R})$. We deduce that:

$$\mathbf{m}^f\left(A^f\right) \leq \mathbf{m}^g\left(A^g\right) + \left|\sigma^f - \sigma^g\right| \leq \mathbf{m}^g\left((A^f)^r\right) + \left|\sigma^f - \sigma^g\right|.$$

The same is true with f and g replaced by g and f. We deduce that:

$$d_P^Z\left(\mathbf{m}^f, \mathbf{m}^g\right) \leq \frac{1}{2}\mathrm{dis}(\mathfrak{R}) + \left|\sigma^f - \sigma^g\right| \leq 2\|f - g\|_\infty + \left|\sigma^f - \sigma^g\right|.$$

We get:

$$d_H^Z\left(T^f, T^g\right) + d^Z\left(\emptyset^f, \emptyset^g\right) + d_P^Z\left(\mathbf{m}^f, \mathbf{m}^g\right) \leq 6\|f - g\|_\infty + \left|\sigma^f - \sigma^g\right|.$$

This gives the result. □

Remark 2. *We could define the correspondence for more general functions f: lower semi-continuous functions that satisfy the intermediate values property (see [DL02]). In that case, the associated real tree is not even locally compact (hence not necessarily proper). But the measurability of the mapping $f \mapsto T^f$ is not clear in this general setting, that is why we only consider continuous function f here and thus will assume the Grey condition (see next Section) for Lévy trees.*

Branching mechanisms

Let Π be a σ-finite measure on $(0, +\infty)$ such that we have $\int (1 \wedge x^2)\Pi(dx) < \infty$. We set:

$$\Pi_\theta(dr) = \mathrm{e}^{-\theta r}\ \Pi(dr). \tag{3.7}$$

Let Θ' be the set of $\theta \in \mathbf{R}$ such that $\int_{(1,+\infty)} \Pi_\theta(dr) < +\infty$. If $\Pi = 0$, then $\Theta' = \mathbf{R}$. We also set $\theta_\infty = \inf \Theta'$. It is obvious that $[0, +\infty) \subset \Theta'$, $\theta_\infty \leq 0$ and either $\Theta' = [\theta_\infty, +\infty)$ or $\Theta' = (\theta_\infty, +\infty)$.

Let $\alpha \in \mathbf{R}$ and $\beta \geq 0$. We consider the branching mechanism ψ associated with (α, β, Π):

$$\psi(\lambda) = \alpha\lambda + \beta\lambda^2 + \int_{(0,+\infty)} (\mathrm{e}^{-\lambda r} - 1 + \lambda r \mathbf{1}_{\{r<1\}})\Pi(dr), \quad \lambda \in \Theta'. \tag{3.8}$$

Notice that the function ψ is smooth and convex over $(\theta_\infty, +\infty)$. We say that ψ is conservative if for all $\varepsilon > 0$:

$$\int_{(0,\varepsilon]} \frac{du}{|\psi(u)|} = +\infty.$$

A sufficient condition for ψ to be conservative is to have $\psi'(0+) > -\infty$. This last condition is actually equivalent to $\int_{(1,\infty)} r\Pi(dr) < \infty$. We will always make the following assumption.

Assumption 1. *The function ψ is conservative and we have $\beta > 0$ or $\int_{(0,1)} \ell\Pi(d\ell) = +\infty$.*

The branching mechanism is said to be sub-critical (resp. critical, super-critical) if $\psi'(0+) > 0$ (resp. $\psi'(0+) = 0$, $\psi'(0+) < 0$). We say that ψ is (sub)critical if it is critical or sub-critical.

We introduce the following branching mechanisms ψ_θ for $\theta \in \Theta'$:

$$\psi_\theta(\lambda) = \psi(\lambda+\theta) - \psi(\theta), \quad \lambda+\theta \in \Theta'. \tag{3.9}$$

Let Θ^ψ be the set of $\theta \in \Theta'$ such that ψ_θ is conservative. Obviously, we have:

$$[0,+\infty) \subset \Theta^\psi \subset \Theta' \subset \Theta^\psi \cup \{\theta_\infty\}.$$

If $\theta \in \Theta^\psi$, we set:

$$\bar{\theta} = \max\left\{q \in \Theta^\psi; \psi(q) = \psi(\theta)\right\}. \tag{3.10}$$

We can give an alternative definition of $\bar{\theta}$ if Assumption 1 holds. Let θ^* be the unique positive root of ψ' if it exists. Notice that $\theta^* = 0$ if ψ is critical and that θ^* exists and is positive if ψ is super-critical. If θ^* exists, then the branching mechanism ψ_{θ^*} is critical. We set Θ^ψ_* for $[\theta^*, +\infty)$ if θ^* exists and $\Theta^\psi_* = \Theta^\psi$ otherwise. The function ψ is a one-to-one mapping from Θ^ψ_* onto $\psi(\Theta^\psi_*)$. We write ψ^{-1} for the inverse of the previous mapping. The set $\{q \in \Theta^\psi; \psi(q) = \psi(\theta)\}$ has at most two elements and we have:

$$\bar{\theta} = \psi^{-1} \circ \psi(\theta).$$

In particular, if ψ_θ is (sub)critical we have $\bar{\theta} = \theta$ and if ψ_θ is super-critical then we have $\theta < \theta^* < \bar{\theta}$.

We will later on consider the following assumption.

Assumption 2. *(Grey condition) The branching mechanism is such that:*

$$\int^{+\infty} \frac{du}{\psi(u)} < \infty.$$

Let us remark that Assumption 2 implies that $\beta > 0$ or $\int_{(0,1)} r\Pi(dr) = +\infty$.

Connections with branching processes

Let ψ be a branching mechanism satisfying Assumption 1. A continuous state branching process (CSBP) with branching mechanism ψ and initial mass $x > 0$ is the càdlàg $\mathbf{R}_+$-valued Markov process $(Z_a,\ a \geq 0)$ whose distribution is characterized by $Z_0 = x$ and:

$$\mathbb{E}[\exp(-\lambda Z_{a+a'})|Z_a] = \exp(-Z_a u(a', \lambda)), \quad \lambda \geq 0,$$

where $(u(a,\lambda), a \geq 0, \lambda > 0)$ is the unique non-negative solution to the integral equation:

$$\int_{u(a,\lambda)}^{\lambda} \frac{dr}{\psi(r)} = a\ ; \quad u(0,\lambda) = \lambda. \tag{3.11}$$

The distribution of the CSBP started at mass x will be noted $\mathbf{P}^\psi_x$. For a detailed presentation of CSBPs, we refer to the monographs [Kyp06],[Lam08] or [Li11].

In this context, the conservative assumption is equivalent to the CSBP not blowing up in finite time, and Assumption 2 is equivalent to the strong extinction time, $\inf\{a; Z_a = 0\}$, being a.s. finite. If Assumption 2 holds, then for all $h > 0$, $\mathbf{P}_x^\psi(Z_h > 0) = \exp(-xb(h))$, where $b(h) = \lim_{\lambda\to+\infty} u(h,\lambda)$. In particular $b(h)$ is such that

$$\int_{b(h)}^{\infty} \frac{dr}{\psi(r)} = h. \tag{3.12}$$

Let us now describe a Girsanov transform for CSBPs introduced in [AD12a] related to the shift of the branching mechanism ψ defined by (3.9). Recall notation Θ^ψ and θ_∞ from the previous Section. For $\theta \in \Theta^\psi$, we consider the process $M^{\psi,\theta} = (M_a^{\psi,\theta}, a \geq 0)$ defined by:

$$M_a^{\psi,\theta} = \exp\left(\theta x - \theta Z_a - \psi(\theta)\int_0^a Z_s ds\right). \tag{3.13}$$

Theorem 3.10 (Girsanov transformation for CSBPs, [AD12a]). *Let ψ be a branching mechanism satisfying Assumption 1. Let $(Z_a, a \geq 0)$ be a CSBP with branching mechanism ψ and let $\mathscr{F} = (\mathscr{F}_a, a \geq 0)$ be its natural filtration. Let $\theta \in \Theta^\psi$ such that either $\theta \geq 0$ or $\theta < 0$ and $\int_{(1,+\infty)} r\Pi_\theta(dr) < +\infty$. Then we have the following:*

1. *The process $M^{\psi,\theta}$ is a $\mathscr{F}$-martingale under $\mathbf{P}_x^\psi$.*
2. *Let $a, x \geq 0$. On $\mathscr{F}_a$, the probability measure $\mathbf{P}_x^{\psi_\theta}$ is absolutely continuous w.r.t. $\mathbf{P}_x^\psi$, and*

$$\frac{d\mathbf{P}^{\psi_\theta}_{x\,|\mathscr{F}_a}}{d\mathbf{P}^{\psi}_{x\,|\mathscr{F}_a}} = M_a^{\psi,\theta}.$$

The height process

Let $(X_t, t \geq 0)$ be a Lévy process with Laplace exponent ψ satisfying Assumption 1. This assumption implies that a.s. the paths of X have infinite total variation over any non-trivial interval. The distribution of the Lévy process will be noted $\mathbb{P}^\psi(dX)$. It is a probability measure on the Skorokhod space of real-valued càdlàg processes. For the remainder of this section, **we will assume that ψ is (sub)critical**.

For $t \geq 0$, let us write $\hat{X}^{(t)}$ for the time-returned process:

$$\hat{X}_s^{(t)} = X_t - X_{(t-s)_-}, \quad 0 \leq s < t$$

and $\hat{X}_t^{(t)} = X_t$. Then $(\hat{X}_s^{(t)}, 0 \leq s \leq t)$ has same distribution as the process $(X_s, 0 \leq s \leq t)$. We will also write $\hat{S}_s^{(t)} = \sup_{[0,s]} \hat{X}_r^{(t)}$ for the supremum process of $\hat{X}^{(t)}$.

Proposition 3.11 (The height process, [DL02]). *Let ψ be a (sub)critical branching mechanism satisfying Assumption 1. There exists a lower semi-continuous process $H = (H_t, t \geq 0)$ taking values in $[0, +\infty]$, with the intermediate values property, which is a local time at 0, at time t, of the process $\hat{X}^{(t)} - \hat{S}^{(t)}$, such that the following convergence holds in probability:*

$$H_t = \lim_{\varepsilon\downarrow 0} \frac{1}{\varepsilon}\int_0^t \mathbf{1}_{\{I_s^t \leq X_s \leq I_s^t + \varepsilon\}} ds$$

where $I_s^t = \inf_{s\leq r\leq t} X_r$. Furthermore, if Assumption 2 holds, then the process H admits a continuous modification.

From now on, we always assume that Assumptions 1 and 2 hold, and we always work with this continuous version of H. The process H is called the height process.

For $x > 0$, we consider the stopping time $\tau_x = \inf\{t \geq 0,\ I_t \leq -x\}$, where $I_t = I_0^t$ is the infimum process of X. We denote by $\mathbb{P}_x^\psi(dH)$ the distribution of the stopped height process $(H_{t\wedge\tau_x}, t \geq 0)$ under $\mathbb{P}^\psi$, defined on the space $\mathscr{C}_+([0,+\infty))$ of non-negative continuous functions on $[0,+\infty)$. The (sub)criticality of the branching mechanism entails $\tau_x < \infty$ $\mathbb{P}^\psi$-a.s., so that under $\mathbb{P}_x^\psi(dH)$, the height process has a.s. compact support.

The excursion measure

The height process is not a Markov process, but it has the same zero sets as $X - I$ (see [DL02], Paragraph 1.3.1), so we can develop an excursion theory based on the latter. By standard fluctuation theory, it is easy to see that 0 is a regular point for $X - I$ and that $-I$ is a local time of $X - I$ at 0. We denote by $\mathbb{N}^\psi$ the associated excursion measure. As such, $\mathbb{N}^\psi$ is a σ-finite measure. Under $\mathbb{P}_x^\psi$ or $\mathbb{N}^\psi$, we set:

$$\sigma(H) = \int_0^\infty \mathbf{1}_{\{H_t \neq 0\}} dt.$$

When there is no risk of confusion, we will write σ for $\sigma(H)$. Notice that, under $\mathbb{P}_x^\psi$, $\sigma = \tau_x$ and that under $\mathbb{N}^\psi$, σ represents the lifetime of the excursion. Abusing notations, we will write $\mathbb{P}_x^\psi(dH)$ and $\mathbb{N}^\psi[dH]$ for the distribution of H under $\mathbb{P}_x^\psi$ or $\mathbb{N}^\psi$. Let us also recall the Poissonian decomposition of the measure $\mathbb{P}_x^\psi$. Under $\mathbb{P}_x^\psi$, let $(a_j, b_j)_{j\in J}$ be the excursion intervals of $X - I$ away from 0. Those are also the excursion intervals of the height process away from 0. For $j \in J$, we shall denote by $H^{(j)} : [0,\infty) \to \mathbf{R}_+$ the corresponding excursion, that is

$$H_t^{(j)} = H_{(a_j+t)\wedge b_j}, \quad t \geq 0.$$

Proposition 3.12 ([DL05]). *Let ψ be a (sub)critical branching mechanism satisfying Assumption 1. Under $\mathbb{P}_x^\psi$, the random point measure $\mathcal{N} = \sum_{j\in J} \delta_{H^{(j)}}(dH)$ is a Poisson point measure with intensity $x\mathbb{N}^\psi[dH]$.*

Local times of the height process

Proposition 3.13 ([DL02], Formula (36)). *Let ψ be a (sub)critical branching mechanism satisfying Assumption 1. Under $\mathbb{N}^\psi$, there exists a jointly measurable process $(L_s^a, a \geq 0, s \geq 0)$ which is continuous and non-decreasing in the variable s such that,*

$$L_s^0 = 0, \quad s \geq 0$$

and for every $t \geq 0$, for every $\delta > 0$ and every $a > 0$

$$\lim_{\varepsilon\to 0} \mathbb{N}^\psi \left[\mathbf{1}_{\{\sup H > \delta\}} \sup_{0\leq s\leq t\wedge\sigma} \left| \varepsilon^{-1} \int_0^s \mathbf{1}_{\{a < H_r \leq a+\varepsilon\}}\, dr - L_s^a \right| \right] = 0.$$

Moreover, by Lemma 3.3. of [DL05], the process $(L_\sigma^a,\ a \geq 0)$ has a càdlàg modification under $\mathbb{N}^\psi$ with no fixed discontinuities.

(Sub)critical Lévy trees

Let ψ be a (sub)critical branching mechanism satisfying Assumptions 1 and 2. Let H be the height process defined under $\mathbb{P}_x^\psi$ or $\mathbb{N}^\psi$. We consider the so-called Lévy tree $\mathcal{T}^H$ which is the random w-tree coded by the function H, see Section 3.2. Notice that we are indeed within the framework of proper real trees, since Assumption 2 entails compactness of $\mathcal{T}^H$. The measurability of the random variable $\mathcal{T}^H$ taking values in $\mathbb{T}$ follows from Proposition 3.9 and Theorem 3.6. When there is no confusion, we shall write $\mathcal{T}$ for $\mathcal{T}^H$. Abusing notations, we will write $\mathbb{P}_x^\psi(d\mathcal{T})$ and $\mathbb{N}^\psi[d\mathcal{T}]$ for the distribution on $\mathbb{T}$ of $\mathcal{T} = \mathcal{T}^H$ under $\mathbb{P}_x^\psi(dH)$ or $\mathbb{N}^\psi[dH]$. By construction, under $\mathbb{P}_x^\psi$ or under $\mathbb{N}^\psi$, we have that the total mass of the mass measure on $\mathcal{T}$ is given by:

$$\mathbf{m}^{\mathcal{T}}(\mathcal{T}) = \sigma. \tag{3.14}$$

Proposition 3.12 enables us to view the measure $\mathbb{N}^\psi[d\mathcal{T}]$ as describing a single Lévy tree. Thus, we will mostly work under this excursion measure, which is the distribution of the (isometry class of the) w-tree $\mathcal{T}$ described by the height process under $\mathbb{N}^\psi$. In order to state the branching property of a Lévy tree, we must first define a local time at level a on the tree. Let $(\mathcal{T}^{i,\circ}, i \in I)$ be the trees that were cut off by cutting at level a, namely the connected components of the set $\mathcal{T} \setminus \pi_a(\mathcal{T})$. If $i \in I$, then all the points in $\mathcal{T}^{i,\circ}$ have the same MRCA x_i in $\mathcal{T}$ which is precisely the point where the tree was cut off. We consider the compact tree $\mathcal{T}^i = \mathcal{T}^{i,\circ} \cup \{x_i\}$ with the root x_i, the metric $d^{\mathcal{T}^i}$, which is the metric $d^{\mathcal{T}}$ restricted to $\mathcal{T}^i$, and the mass measure $\mathbf{m}^{\mathcal{T}^i}$, which is the mass measure $\mathbf{m}^{\mathcal{T}}$ restricted to $\mathcal{T}^i$. Then $(\mathcal{T}^i, d^{\mathcal{T}^i}, x_i, \mathbf{m}^{\mathcal{T}^i})$ is a w-tree. Let

$$\mathcal{N}_a^{\mathcal{T}}(dx, d\mathcal{T}') = \sum_{i\in I} \delta_{(x_i,\mathcal{T}^i)}(dx, d\mathcal{T}') \tag{3.15}$$

be the point measure on $\mathcal{T}(a) \times \mathbb{T}$ taking account of the cutting points as well as the trees cut away. The following theorem gives the structure of the decomposition we just described. From excursion theory, we deduce that $b(h) = \mathbb{N}^\psi[H_{\max}(\mathcal{T}) > h]$, where $b(h)$ solves (3.12). An easy extension of [DL05] from real trees to w-trees gives the following result.

Theorem 3.14 ([DL05]). *Let ψ be a (sub)critical branching mechanism satisfying Assumptions 1 and 2. There exists a $\mathcal{T}$-measure valued process $(\ell^a, a \geq 0)$ càdlàg for the weak topology on finite measure on $\mathcal{T}$ such that $\mathbb{N}^\psi$-a.e.:*

$$\mathbf{m}^{\mathcal{T}}(dx) = \int_0^\infty \ell^a(dx)\,da, \tag{3.16}$$

$\ell^0 = 0$, $\inf\{a > 0; \ell^a = 0\} = \sup\{a \geq 0; \ell^a \neq 0\} = H_{max}(\mathcal{T})$ *and for every fixed* $a \geq 0$, $\mathbb{N}^\psi$*-a.e.:*

- *ℓ^a is supported on $\mathcal{T}(a)$,*
- *We have for every bounded continuous function ϕ on $\mathcal{T}$:*

$$\begin{aligned}
\langle \ell^a, \phi\rangle &= \lim_{\varepsilon\downarrow 0} \frac{1}{b(\varepsilon)} \int \phi(x) \mathbf{1}_{\{h(\mathcal{T}')\geq\varepsilon\}} \mathcal{N}_a^{\mathcal{T}}(dx, d\mathcal{T}') &(3.17)\\
&= \lim_{\varepsilon\downarrow 0} \frac{1}{b(\varepsilon)} \int \phi(x) \mathbf{1}_{\{h(\mathcal{T}')\geq\varepsilon\}} \mathcal{N}_{a-\varepsilon}^{\mathcal{T}}(dx, d\mathcal{T}'), \text{ if } a > 0. &(3.18)
\end{aligned}$$

Furthermore, we have the branching property: for every $a>0$, the conditional distribution of the point measure $\mathcal{N}_a^{\mathcal{T}}(dx, d\mathcal{T}')$ under $\mathbb{N}^\psi[d\mathcal{T}|H_{max}(\mathcal{T})>a]$, given $\pi_a(\mathcal{T})$, is that of a Poisson point measure on $\mathcal{T}(a)\times\mathbb{T}$ with intensity $\ell^a(dx)\mathbb{N}^\psi[d\mathcal{T}']$.

The measure ℓ^a will be called the *local time* measure of $\mathcal{T}$ at level a. In the case of Lévy trees, it can also be defined as the image of the measure $d_s L_s^a(H)$ by the canonical projection p^H (see [DL02]), so the above statement is in fact the translation of the excursion theory of the height process in terms of real trees. This definition shows that the local time is a function of the tree $\mathcal{T}$ and does not depend on the choice of the coding height function. It should be noted that Equation (3.18) implies that ℓ^a is measurable with respect to the σ-algebra generated by $\pi_a(\mathcal{T})$.

The next theorem, also from [DL05], relates the discontinuities of the process $(\ell^a, a\geq 0)$ to the infinite nodes in the tree. Recall $\mathrm{Br}_\infty(\mathcal{T})$ denotes the set of infinite nodes in the Lévy tree $\mathcal{T}$.

Theorem 3.15 ([DL05]). *Let ψ be a (sub)critical branching mechanism satisfying Assumptions 1 and 2. The set $\{d(\emptyset,x),\ x\in\mathrm{Br}_\infty(\mathcal{T})\}$ coincides $\mathbb{N}^\psi$-a.e. with the set of discontinuity times of the mapping $a\mapsto\ell^a$. Moreover, $\mathbb{N}^\psi$-a.e., for every such discontinuity time b, there is a unique $x_b\in\mathrm{Br}_\infty(\mathcal{T})\cap\mathcal{T}(b)$, and*

$$\ell^b=\ell^{b-}+\Delta_b\delta_{x_b},$$

where $\Delta_b>0$ is called mass *of the node x_b and can be obtained by the approximation*

$$\Delta_b=\lim_{\varepsilon\to 0}\frac{1}{b(\varepsilon)}n(x_b,\varepsilon), \tag{3.19}$$

where $n(x_b,\varepsilon)=\int \mathbf{1}_{\{x=x_b\}}(x)\mathbf{1}_{\{H_{\max}(\mathcal{T}')>\varepsilon\}}(\mathcal{T}')\mathcal{N}_b^{\mathcal{T}}(dx,d\mathcal{T}')$ is the number of sub-trees originating from x_b with height larger than ε.

Decomposition of the Lévy tree

We will frequently use the following notation for the following measure on $\mathbb{T}$:

$$\mathbf{N}^\psi[\mathcal{T}\in\bullet]=2\beta\mathbb{N}^\psi[\mathcal{T}\in\bullet]+\int_{(0,+\infty)} r\Pi(dr)\,\mathbb{P}_r^\psi[\mathcal{T}\in\bullet]. \tag{3.20}$$

where ψ is given by (3.8).

The decomposition of a (sub)critical Lévy tree $\mathcal{T}$ according to a spine $[\![\emptyset,x]\!]$, where $x\in\mathcal{T}$ is a leaf picked at random at level $a>0$, that is according to the local time $\ell^a(dx)$, is given in Theorem 4.5 in [DL05]. Then by integrating with respect to a, we get the decomposition of $\mathcal{T}$ according to a spine $[\![\emptyset,x]\!]$, where $x\in\mathcal{T}$ is a leaf picked at random on $\mathcal{T}$, that is according to the mass measure $\mathbf{m}^{\mathcal{T}}$. Therefore, we will state this decomposition without proof.

Let $x\in\mathcal{T}$ and $\{x_i, i\in I_x\}$ the set $\mathrm{Br}(\mathcal{T})\cap[\![\emptyset,x]\!]$ of branching points on the spine $[\![\emptyset,x]\!]$. For $i\in I_x$, we set:

$$\mathcal{T}^i=\mathcal{T}\setminus\left(\mathcal{T}^{(x,x_i)}\cup\mathcal{T}^{(\emptyset,x_i)}\right),$$

where $\mathcal{T}^{(y,x_i)}$ is the connected component of $\mathcal{T}\setminus\{x_i\}$ containing y. We let x_i be the root of $\mathcal{T}^i$. The metric and measure on $\mathcal{T}^i$ are respectively the restriction of $d^{\mathcal{T}}$ to $\mathcal{T}^i$ and the restriction of $\mathbf{m}^{\mathcal{T}}$ to $\mathcal{T}^i\setminus\{x_i\}$. By construction, if x is a leaf, we have:

$$\mathcal{T} = [\![\emptyset, x]\!] \circledast_{i\in I_x} (\mathcal{T}^i, x_i),$$

where $[\![\emptyset, x]\!]$ is a w-tree with root $\emptyset$, metric and mass measure the restrictions of $d^{\mathcal{T}}$ and $\mathbf{m}^{\mathcal{T}}$ to $[\![\emptyset, x]\!]$.

We consider the point measure on $[0, H_x]\times\mathbb{T}$ defined by:

$$\mathcal{M}_x = \sum_{i\in i_x} \delta_{(H_{x_i}, \mathcal{T}^i)}.$$

Theorem 3.16 ([DL05]). *Let ψ be a (sub)critical branching mechanism satisfying Assumptions 1 and 2. We have for any non-negative measurable function F defined on $[0,+\infty)\times\mathbb{T}$:*

$$\mathbb{N}^{\psi}\left[\int \mathbf{m}^{\mathcal{T}}(dx)F(H_x, \mathcal{M}_x)\right] = \int_0^{\infty} da\, \mathrm{e}^{-\psi'(0)a}\, \mathbb{E}\left[F\left(a, \sum_{i\in I} \mathbf{1}_{\{z_i\le a\}}\delta_{(z_i, \overline{\mathcal{T}^i})}\right)\right],$$

where under $\mathbb{E}$, $\sum_{i\in I}\delta_{(z_i,\overline{\mathcal{T}^i})}(dz, dT)$ is a Poisson point measure on $[0,+\infty)\times\mathbb{T}$ with intensity $dz\,\mathbf{N}^{\psi}[dT]$.

CSBP process in the Lévy trees

Lévy trees give a genealogical structure for CSBPs, which is precised in the next Theorem. We consider the process $\mathcal{Z} = (\mathcal{Z}_a, a\ge 0)$ defined by:

$$\mathcal{Z}_a = \langle \ell^a, 1\rangle.$$

If needed we will write $\mathcal{Z}_a(\mathcal{T})$ to emphasize that $\mathcal{Z}_a$ corresponds to the tree $\mathcal{T}$.

Theorem 3.17 (CSBP in Lévy trees, [DL02] and [DL05]). *Let ψ be a (sub)critical branching mechanism satisfying Assumptions 1 and 2, and let $x>0$. The process $\mathcal{Z}$ under $\mathbb{P}_x^{\psi}$ is distributed as the CSBP Z under $\mathbf{P}_x^{\psi}$.*

Remark 3. *This theorem can be stated in terms of the height process without Assumption 2.*

Super-critical Lévy trees

Let us now briefly recall the construction from [AD12a] for super-critical Lévy trees using a Girsanov transformation similar to the one used for CSBPs, see Theorem 3.10.

Let ψ be a super-critical branching mechanism satisfying Assumptions 1 and 2. Recall θ^* is the unique positive root of ψ' and that the branching mechanism ψ_θ is sub-critical if $\theta>\theta^*$, critical if $\theta=\theta^*$ and super-critical otherwise. We consider the filtration $\mathcal{H} = (\mathcal{H}_a,\ a\ge 0)$, where $\mathcal{H}_a$ is the σ-field generated by the random variable $\pi_a(\mathcal{T})$ and the $\mathbb{P}_x^{\psi_{\theta^*}}$-negligible sets. For $\theta\ge\theta^*$, we define the process $M^{\psi,\theta} = (M_a^{\psi,\theta}, a\ge 0)$ with:

$$M_a^{\psi,\theta} = \exp\left(\theta x - \theta\mathcal{Z}_a - \psi(\theta)\int_0^a \mathcal{Z}_s ds\right)$$

By absolute continuity of the measures $\mathbb{P}_x^{\psi_\theta}$ (resp. $\mathbb{N}^{\psi_\theta}$) with respect to $\mathbb{P}_x^{\psi_{\theta^*}}$ (resp. $\mathbb{N}^{\psi_{\theta^*}}$), all the processes $M^{\psi_\theta,-\theta}$ for $\theta > \theta^*$ are $\mathscr{H}$-adapted. Moreover, all these processes are $\mathscr{H}$-martingales (see [AD12a] for the proof). Theorem 3.14 shows that $M^{\psi_{\theta^*},-\theta^*}$ is $\mathscr{H}$-adapted. Let us now define the ψ-Lévy tree, cut at level a by the following Girsanov transformation.

Definition 3.18. *Let ψ be a super-critical branching mechanism satisfying Assumptions 1 and 2. Let $\theta \geq \theta^*$. For $a \geq 0$, we define the distribution $\mathbb{P}_x^{\psi,a}$ (resp. $\mathbb{N}^{\psi,a}$) by: if F is a non-negative, measurable functional defined on $\mathbb{T}$,*

$$\mathbb{E}_x^{\psi,a}[F(\mathscr{T})] = \mathbb{E}_x^{\psi_\theta}\Big[M_a^{\psi_\theta,-\theta}F(\pi_a(\mathscr{T}))\Big], \tag{3.21}$$

$$\mathbb{N}^{\psi,a}[F(\mathscr{T})] = \mathbb{N}^{\psi_\theta}\Big[\exp\Big(\theta\mathcal{Z}_a + \psi(\theta)\int_0^a \mathcal{Z}_s(ds)F(\pi_a(\mathscr{T})\Big)\Big]. \tag{3.22}$$

It can be checked that the definition of $\mathbb{P}_x^{\psi,a}$ (and of $\mathbb{N}^{\psi,a}$) does not depend on $\theta \geq \theta^*$.

The probability measures $\mathbb{P}_x^{\psi,a}$ satisfy a consistence property, allowing us to define the super-critical Lévy tree in the following way.

Theorem 3.19. *Let ψ be a super-critical branching mechanism satisfying assumptions 1 and 2. There exists a probability measure $\mathbb{P}_x^\psi$ (resp. a σ-finite measure $\mathbb{N}^\psi$) on $\mathbb{T}$ such that for $a > 0$, we have, if F is a measurable non-negative functional on $\mathbb{T}$,*

$$\mathbb{E}_x^\psi[F(\pi_a(\mathscr{T}))] = \mathbb{E}_x^{\psi,a}[F(\mathscr{T})],$$

the same being true under $\mathbb{N}^\psi$.

The w-tree $\mathscr{T}$ under $\mathbb{P}_x^\psi$ or $\mathbb{N}^\psi$ is called a ψ-Lévy w-tree or simply a Lévy tree.

Proof. For $n \geq 1$, $0 < a_1 < \cdots < a_n$, we define a probability measure on $\mathbb{T}^n$ by:

$$\mathbb{P}_x^{\psi,a_1,\ldots,a_n}(\mathscr{T}_1 \in A_1,\ldots,\mathscr{T}_n \in A_n) = \mathbb{P}_x^{\psi,a_n}(\mathscr{T} \in A_n, \pi_{a_{n-1}}(\mathscr{T}) \in A_{n-1},\ldots,\pi_{a_1}(\mathscr{T}) \in A_1)$$

if $A_1,\ldots,A_n$ are Borel subsets of $\mathbb{T}$. The probability measures

$$(\mathbb{P}_x^{\psi,a_1,\ldots,a_n},\ n \geq 1,\ 0 < a_1 < \cdots < a_n)$$

then form a projective family. This is a consequence of the martingale property of $M^{\psi_\theta,-\theta}$ and the fact that the projectors π_a satisfy the obvious compatibility relation $\pi_b \circ \pi_a = \pi_b$ if $0 < b < a$.

By the Daniell-Kolmogorov theorem, there exists a probability measure $\tilde{\mathbb{P}}_x^\psi$ on the product space $\mathbb{T}^{\mathbf{R}_+}$ such that the finite-dimensional distributions of a $\tilde{\mathbb{P}}_x^\psi$-distributed family are described by the measures defined above. It is easy to construct a version of a $\tilde{\mathbb{P}}_x^\psi$-distributed process that is a.s. increasing. Indeed, almost all sample paths of a $\tilde{\mathbb{P}}_x^\psi$-distributed process are increasing when restricted to rational numbers. We can then define a w-tree $\mathscr{T}^a$ for any $a > 0$ by considering a decreasing sequence of rational numbers $a_n \downarrow a$ and defining $\mathscr{T}^a = \cap_{n\geq 1}\mathscr{T}^{a_n}$. Notice that $\mathscr{T}^a$ is closed for all $a \in \mathbf{R}_+$. It is easy to check that the finite-dimensional distributions of this new process are unchanged by this procedure. Let us then consider $\mathscr{T} = \cup_{a>0}\mathscr{T}^a$, endowed with the obvious metric $d^{\mathscr{T}}$ and mass measure $\mathbf{m}$. It is clear that $\mathscr{T}$ is a real tree, rooted at the common root of the $\mathscr{T}^a$. All the $\mathscr{T}^a$ are compact, so that $\mathscr{T}$ is locally compact and complete. The measure $\mathbf{m}$ is locally finite since all the $\mathbf{m}^{\mathscr{T}^a}$ are finite measures. Therefore, $\mathscr{T}$ is a.s. a w-tree. Then, if we define $\mathbb{P}^\psi_x$ to be the distribution of $\mathscr{T}$, the conclusion follows. Similar arguments hold under $\mathbb{N}^\psi$. □

Remark 4. *Another definition of super-critical Lévy trees was given by Duquesne and Winkel [DW07, DW12]: they consider increasing families of Galton-Watson trees with exponential edge lengths which satisfy a certain hereditary property (such as uniform Bernoulli coloring of the leaves). Lévy trees are then defined to be the Gromov-Hausdorff limits of these processes. Another approach via backbone decompositions is given in [BKMS11].*

All the definitions we made for sub-critical Lévy trees then carry over to the super-critical case. In particular, the level set measure ℓ^a, which is $\pi_a(\mathscr{T})$-measurable, can be defined using the Girsanov formula. Thanks to Theorem 3.10, it is easy to show that the mass process $(\mathscr{Z}_a = \langle \ell^a, 1\rangle,\ a \geq 0)$ is under $\mathbb{P}^\psi_x$ a CSBP with branching mechanism ψ. In particular, with u defined in (3.11) and b by (3.12), we have:

$$\mathbb{N}^\psi\left[1-\mathrm{e}^{-\lambda \mathscr{Z}_a}\right] = u(a,\lambda) \quad \text{and} \quad \mathbb{N}^\psi\left[H_{\max}(\mathscr{T})>a\right] = \mathbb{N}^\psi\left[\mathscr{Z}_a>0\right] = b(a). \tag{3.23}$$

Notice that b is finite only under Assumption 2. We set:

$$\sigma = \int_0^{+\infty} \mathscr{Z}_a\, da = \mathbf{m}^{\mathscr{T}}(\mathscr{T}) \tag{3.24}$$

for the total mass of the Lévy tree $\mathscr{T}$. Notice this is consistent with (3.16) and (3.14) which are defined for (sub)critical Lévy trees. Thanks to (3.24), notice that σ is distributed as the total population size of a CSBP with branching mechanism ψ. In particular, its Laplace transform is given for $\lambda > 0$ by:

$$\mathbb{N}^\psi\left[1-\mathrm{e}^{-\lambda\sigma}\right] = \psi^{-1}(\lambda). \tag{3.25}$$

Notice that $\mathbb{N}^\psi[\sigma = +\infty] = \psi^{-1}(0) > 0$.

We recall the following Theorem, from [AD12a], which sums up the situation for any branching mechanisme ψ.

Theorem 3.20 ([AD12a]). *Let ψ be any branching mechanism satisfying Assumptions 1 and 2, and let $q>0$ such that $\psi(q) \geq 0$. Then, the probability measure $\mathbb{P}^{\psi_q}_x$ on $\mathbb{T}$ is absolutely continuous w.r.t. $\mathbb{P}^\psi_x$, with*

$$\frac{d\mathbb{P}^{\psi_q}_x}{d\mathbb{P}^\psi_x} = M^{\psi,q}_\infty = \mathrm{e}^{qx-\psi(q)\sigma}\mathbf{1}_{\{\sigma<+\infty\}}. \tag{3.26}$$

Similarly, the excursion measure $\mathbb{N}^{\psi_q}$ on $\mathbb{T}$ is absolutely continuous w.r.t. $\mathbb{N}^\psi$ and we have

$$\frac{d\mathbb{N}^{\psi_q}}{d\mathbb{N}^\psi} = \mathrm{e}^{-\psi(q)\sigma}\mathbf{1}_{\{\sigma<+\infty\}}. \tag{3.27}$$

When applying Girsanov formula (3.27) to $q=\bar\theta$ defined by (3.10), we get the following remarkable Corollary, due to the fact that $\psi_\theta(\bar\theta-\theta)=0$.

Corollary 3.21. *Let ψ be a critical branching mechanism satisfying Assumptions 1 and 2, and $\theta \in \Theta^\psi$ with $\theta < 0$. Let F be a non-negative measurable functional defined on $\mathbb{T}$. We have:*

$$\begin{aligned} \mathrm{e}^{(\bar\theta-\theta)x}\, \mathbb{E}_x^{\psi_\theta}[F(\mathcal{T})\mathbf{1}_{\{\sigma<+\infty\}}] &= \mathbb{E}_x^{\psi_{\bar\theta}}[F(\mathcal{T})], \\ \mathbb{N}^{\psi_\theta}[F(\mathcal{T})\mathbf{1}_{\{\sigma<+\infty\}}] &= \mathbb{N}^{\psi_{\bar\theta}}[F(\mathcal{T})]. \end{aligned} \tag{3.28}$$

We deduce from Proposition 3.12 and Theorem 3.19 that the point process $\mathcal{N}_0^{\mathcal{T}}(dx, d\mathcal{T}')$ defined by (3.15) with $a=0$ is under $\mathbb{P}_x^\psi(d\mathcal{T})$ a Poisson point measure on $\{\emptyset\}\times\mathbb{T}$ with intensity $\sigma\delta_\emptyset(dx)\mathbb{N}^\psi[d\mathcal{T}']$. Then we deduce from (3.21), with $F=1$, that for $\theta \geq \theta^*$:

$$\mathbb{N}^{\psi_\theta}\left[1-\exp\left(\theta\mathcal{Z}_a + \psi(\theta)\int_0^a \mathcal{Z}_s ds\right)\right] = -\theta. \tag{3.29}$$

Pruning Lévy trees

We recall the construction from [ADV10] on the pruning of Lévy trees. Let $\mathcal{T}$ be a random Lévy w-tree under $\mathbb{P}_x^\psi$ (or under $\mathbb{N}^\psi$), with ψ conservative. Let

$$m^{(\mathrm{ske})}(dx, d\theta) = \sum_{i\in I^{\mathrm{ske}}} \delta_{(x_i,\theta_i)}(dx, d\theta)$$

be, conditionally on $\mathcal{T}$, a Poisson point measure on $\mathcal{T}\times\mathbf{R}_+$ with intensity $2\beta l^{\mathcal{T}}(dx)d\theta$. Since there is a.s. a countable number of branching points (which have $l^{\mathcal{T}}$-measure 0), the atoms of this measure are distributed on $\mathcal{T}\setminus(\mathrm{Br}(\mathcal{T})\cup\mathrm{Lf}(\mathcal{T}))$.

If $\Pi = 0$, we have $\mathrm{Br}_\infty(\mathcal{T}) = \emptyset$ a.s. whereas if $\Pi(\mathbf{R}_+) = \infty$, $\mathrm{Br}_\infty(\mathcal{T})$ is a.s. a countable dense subset of $\mathcal{T}$. If the latter condition holds, we consider, conditionally on $\mathcal{T}$, a Poisson point measure

$$m^{(\mathrm{nod})}(dx, d\theta) = \sum_{i\in I^{\mathrm{nod}}} \delta_{(x_i,\theta_i)}(dx, d\theta)$$

on $\mathcal{T}\times\mathbf{R}_+$ with intensity

$$\sum_{y\in\mathrm{Br}_\infty(\mathcal{T})} \Delta_y\delta_y(dx)\,d\theta$$

where Δ_x is the mass of the node x, defined by (3.19). Hence, if $\theta > 0$, a node $x\in\mathrm{Br}_\infty(\mathcal{T})$ is an atom of $m^{(\mathrm{nod})}(dx,[0,\theta])$ with probability $1-\exp(-\theta\Delta_x)$. The set

$$\left\{x_i,\ i\in I^{\mathrm{nod}}\right\} = \left\{x\in\mathcal{T},\ m^{(\mathrm{nod})}\left(\{x\}\times\mathbf{R}_+\right) > 0\right\}$$

of marked branching points corresponds $\mathbb{P}_x^\psi$-a.s or $\mathbb{N}^\psi$-a.e. to $\mathrm{Br}_\infty(\mathcal{T})$. For $i \in I^{\mathrm{nod}}$, we set

$$\theta_i = \inf\left\{\theta > 0,\ m^{(\mathrm{nod})}\left(\{x_i\}\times[0,\theta]\right) > 0\right\}$$

the first mark on x_i (which is conditionally on $\mathcal{T}$ exponentially distributed with parameter θ_{x_i}), and we set

$$\left\{\theta_j,\ j\in J_i^{\mathrm{nod}}\right\} = \left\{\theta > \theta_i,\ m^{(\mathrm{nod})}\left(\{x_i\}\times\{\theta\}\right) > 0\right\}$$

so that we can write

$$m^{(\mathrm{nod})}(dx,d\theta)=\sum_{i\in I^{\mathrm{nod}}}\delta_{x_i}(dx)\left(\delta_{\theta_i}(d\theta)+\sum_{j\in J_i^{\mathrm{nod}}}\delta_{\theta_j}(d\theta)\right).$$

We set the measure of marks:

$$\mathcal{M}(dx,d\theta)=m^{(\mathrm{ske})}(dx,d\theta)+m^{(\mathrm{nod})}(dx,d\theta), \tag{3.30}$$

and consider the family of w-trees $\Lambda(\mathcal{T},\mathcal{M})=(\Lambda_\theta(\mathcal{T},\mathcal{M}),\theta\geq 0)$, where the θ-pruned w-tree Λ_θ is defined by:

$$\Lambda_\theta(\mathcal{T},\mathcal{M})=\{x\in\mathcal{T},\ \mathcal{M}(\llbracket\emptyset,x\llbracket\times[0,\theta])=0\},$$

rooted at $\emptyset^{\Lambda_\theta(\mathcal{T},\mathcal{M})}=\emptyset^{\mathcal{T}}$, and the metric $d^{\Lambda_\theta(\mathcal{T},\mathcal{M})}$ and the mass measure $\mathbf{m}^{\Lambda_\theta(\mathcal{T},\mathcal{M})}$ are the restrictions of $d^{\mathcal{T}}$ and $\mathbf{m}^{\mathcal{T}}$ to $\Lambda_\theta(\mathcal{T},\mathcal{M})$. In particular, we have $\Lambda_0(\mathcal{T},\mathcal{M})=\mathcal{T}$. The family of w-trees $\Lambda(\mathcal{T},\mathcal{M})$ is a non-increasing family of real trees, in a sense that $\Lambda_\theta(\mathcal{T},\mathcal{M})$ is a subtree of $\Lambda_{\theta'}(\mathcal{T},\mathcal{M})$ for $0\leq\theta'\leq\theta$, see Figure 3.1. In particular, we have that the pruning operators satisfy a cocycle property, for $\theta_1\geq 0$ and $\theta_2\geq 0$:

$$\Lambda_{\theta_2}\big(\Lambda_{\theta_1}(\mathcal{T},\mathcal{M}),\mathcal{M}_{\theta_1}\big)=\Lambda_{\theta_2+\theta_1}(\mathcal{T},\mathcal{M}),$$

where $\mathcal{M}_\theta(A\times[0,q])=\mathcal{M}(A\times[\theta,\theta+q])$. Abusing notation, we write $\mathbb{N}^\psi(d\mathcal{T},d\mathcal{M})$ for the distribution of the pair $(\mathcal{T},\mathcal{M})$ when $\mathcal{T}$ is distributed according to $\mathbb{N}^\psi(d\mathcal{T})$ and conditionally on $\mathcal{T}$, $\mathcal{M}$ is distributed as described above.

The following result can be deduced from [AD12a].

Theorem 3.22. *Let ψ be a branching mechanism satisfying Assumptions 1 and 2. There exists a non-increasing $\mathbb{T}$-valued Markov process $(\mathcal{T}_\theta,\theta\in\Theta^\psi)$ such that for all $q\in\Theta^\psi$, the process $(\mathcal{T}_{\theta+q},\theta\geq 0)$ is distributed as $\Lambda(\mathcal{T},\mathcal{M})$ under $\mathbb{N}^{\psi_q}[d\mathcal{T},d\mathcal{M}]$.*

In particular, this Theorem implies that $\mathcal{T}_\theta$ is distributed as $\mathbb{N}^{\psi_\theta}$ for $\theta\in\Theta^\psi$ and that for $\theta_0\geq 0$, under $\mathbb{N}^\psi$, the process of pruned trees $(\Lambda_{\theta_0+\theta}(\mathcal{T}),\theta\geq 0)$ has the same distribution as $(\Lambda_\theta(\mathcal{T}),\theta\geq 0)$ under $\mathbb{N}^{\psi_{\theta_0}}[d\mathcal{T}]$.

We want to study the time-reversed process $(\mathcal{T}_{-\theta},\theta\in-\Theta^\psi)$, which can be seen as a growth process. This process grows by attaching sub-trees at a random point, rather than slowly growing uniformly along the branches. We recall some results from [AD12a] on the growth process. From now on, we will assume in this section that the **branching mechanism ψ is critical**, so that ψ_θ is sub-critical iff $\theta>0$ and super-critical iff $\theta<0$.

We will use the following notation for the total mass of the tree $\mathcal{T}_\theta$ at time $\theta\in\Theta^\psi$:

$$\sigma_\theta=\mathbf{m}^{\mathcal{T}_\theta}(\mathcal{T}_\theta). \tag{3.31}$$

The total mass process $(\sigma_\theta,\theta\in\Theta^\psi)$ is a pure-jump process taking values in $(0,+\infty]$.

Lemma 3.23 ([AD12a]). *Let ψ be a critical branching mechanism satisfying Assumptions 1 and 2. If $0\leq\theta_2<\theta_1$, then we have:*

$$\mathbb{N}^\psi[\sigma_{\theta_2}|\mathcal{T}_{\theta_1}]=\sigma_{\theta_1}\frac{\psi'(\theta_1)}{\psi'(\theta_2)}.$$

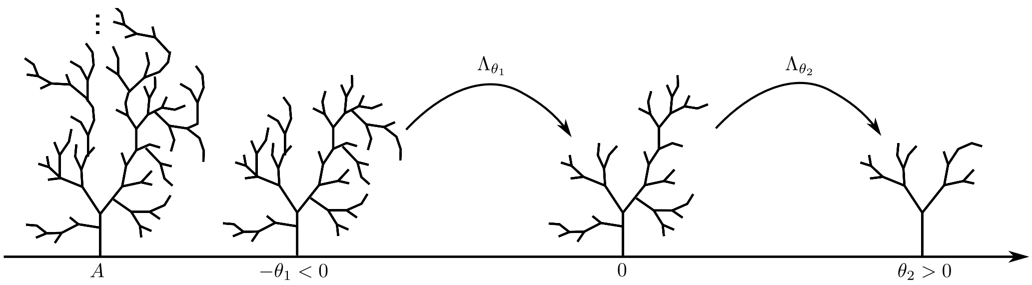

Figure 3.1: The pruning process, starting from explosion time A defined in (3.32).

Consider the ascension time (or explosion time):

$$A = \inf\{\theta \in \Theta^{\psi},\ \sigma_\theta < \infty\}, \tag{3.32}$$

where we use the convention $\inf \emptyset = \theta_\infty$. The following Theorem gives the distribution of the ascension time A and the distribution of the tree at this random time. Recall that $\bar{\theta} = \psi^{-1}(\psi(\theta))$ is defined in (3.10).

Theorem 3.24 ([AD12a]). *Let ψ be a critical branching mechanism satisfying Assumptions 1 and 2.*

1. *For all $\theta \in \Theta^{\psi}$, we have $\mathbb{N}^{\psi}[A > \theta] = \bar{\theta} - \theta$.*

2. *If $\theta_\infty < \theta < 0$, under $\mathbb{N}^{\psi}$, we have, for any non-negative measurable functional F,*

$$\mathbb{N}^{\psi}[F(\mathscr{T}_{A+\theta'}, \theta' \geq 0) | A = \theta] = \psi'(\bar{\theta})\mathbb{N}^{\psi}\left[F(\mathscr{T}_{\theta'}, \theta' \geq 0)\sigma_0 \mathrm{e}^{-\psi(\theta)\sigma_0}\right].$$

3. *For all $\theta \in \Theta^{\psi}$, we have $\mathbb{N}^{\psi}[\sigma_A < +\infty | A = \theta] = 1$.*

In other words, at the ascension time, the tree can be seen as a size-biased critical Lévy tree. A precise description of $\mathscr{T}_A$ is given in [AD12a]. Notice that in the setting of [AD12a], there is no need of Assumption 2.

3.3 The growing tree-valued process

Special Markov Property of pruning

In [ADV10], the authors prove a formula describing the structure of a Lévy tree, conditionally on the θ-pruned tree obtained from it in the (sub)critical case. We will give a general version of this result. From the measure of marks, $\mathscr{M}$ in (3.30), we define a measure of increasing marks by:

$$\mathscr{M}^{\uparrow}(dx, d\theta') = \sum_{i \in I^{\uparrow}} \delta_{(x_i, \theta_i)}(dx, d\theta'), \tag{3.33}$$

with

$$I^{\uparrow} = \Big\{ i \in I^{\text{ske}} \cup I^{\text{nod}}; \mathscr{M}([\![\emptyset, x_i]\!] \times [0, \theta_i]) = 1 \Big\}.$$

The atoms (x_i, θ_i) for $i \in I^{\uparrow}$ correspond to marks such that there are no marks of $\mathscr{M}$ on $[\![\emptyset, x_i]\!]$ with a θ-component smaller than θ_i. In the case of multiple θ_j for a given node $x_i \in \mathrm{Br}_{\infty}(\mathscr{T})$, we only keep the smallest one. In the case $\Pi = 0$, the measure $\mathscr{M}^{\uparrow}$ describes the jumps of a record process on the tree, see [AD11] for further work in this direction. The θ-pruned tree can alternatively be defined using $\mathscr{M}^{\uparrow}$ instead of $\mathscr{M}$ as for $\theta \geq 0$:

$$\Lambda_\theta(\mathscr{T}, \mathscr{M}) = \Big\{ x \in \mathscr{T},\ \mathscr{M}^{\uparrow}([\![\emptyset, x[\![\times [0, \theta]) = 0 \Big\}.$$

We set:

$$I_\theta^{\uparrow} = \Big\{ i \in I^{\uparrow}, x_i \in \mathrm{Lf}(\Lambda_\theta(\mathscr{T}, \mathscr{M})) \Big\} = \Big\{ i \in I^{\uparrow}, \theta_i < \theta \quad \text{and} \quad \mathscr{M}^{\uparrow}([\![\emptyset, x_i[\![\times [0, \theta]) = 0 \Big\}$$

and for $i \in I_\theta^{\uparrow}$:

$$\mathscr{T}^i = \mathscr{T} \setminus \mathscr{T}^{\emptyset, x_i} = \{ x \in \mathscr{T},\ x_i \in [\![\emptyset, x]\!]\},$$

where $\mathscr{T}^{y,x}$ is the connected component of $\mathscr{T} \setminus \{x\}$ containing y. For $i \in I_\theta^{\uparrow}$, $\mathscr{T}^i$ is a real tree, and we will consider x_i as its root. The metric and mass measure on $\mathscr{T}^i$ are the restriction of the metric and mass measure of $\mathscr{T}$ on $\mathscr{T}^i$. By construction, we have:

$$\mathscr{T} = \Lambda_\theta(\mathscr{T}, \mathscr{M}) \circledast_{i \in I_\theta^{\uparrow}} (\mathscr{T}^i, x_i). \tag{3.34}$$

Now we can state the general special Markov property.

Theorem 3.25 (Special Markov Property). *Let ψ be a branching mechanism satisfying Assumptions 1 and 2. Let $\theta > 0$. Conditionally on $\Lambda_\theta(\mathscr{T}, \mathscr{M})$, the point measure:*

$$\mathscr{M}_\theta^{\uparrow}(dx, d\mathscr{T}', d\theta') = \sum_{i \in I_\theta^{\uparrow}} \delta_{(x_i, \mathscr{T}^i, \theta_i)}(dx, d\mathscr{T}', d\theta')$$

under $\mathbb{P}_{r_0}^{\psi}$ (or under $\mathbb{N}^{\psi}$) is a Poisson point measure on $\Lambda_\theta(\mathscr{T}, \mathscr{M}) \times \mathbb{T} \times (0, \theta]$ with intensity:

$$\mathbf{m}^{\Lambda_\theta(\mathscr{T}, \mathscr{M})}(dx) \Big(2\beta \mathbb{N}^{\psi}[d\mathscr{T}'] + \int_{(0,+\infty)} \Pi(dr)\, r e^{-\theta' r} \mathbb{P}_r^{\psi}(d\mathscr{T}') \Big) \mathbf{1}_{(0,\theta]}(\theta')\, d\theta'. \tag{3.35}$$

Proof. It is not difficult to adapt the proof of the special Markov property in [ADV10] to get Theorem 3.25 in the (sub)critical case by taking into account the pruning times θ_i and the w-tree setting; and we omit this proof which can be found in [ADH12b]. We prove how to extend the result to the super-critical Lévy trees using the Girsanov transform of Definition 3.18.

Assume that ψ is super-critical. For $a > 0$, we shall write $\Lambda_{\theta,a}(\mathscr{T}, \mathscr{M}) = \pi_a(\Lambda_\theta(\mathscr{T}, \mathscr{M}))$ for short. According to (3.34) and the definition of super-critical Lévy trees, we have that for any $a > 0$, the truncated tree $\pi_a(\mathscr{T})$ can be written as:

$$\pi_a(\mathscr{T}) = \Lambda_{\theta,a}(\mathscr{T}, \mathscr{M}) \circledast_{\substack{i \in I_\theta^{\uparrow}, \\ H_{x_i} \leq a}} \Big(\pi_{a - H_{x_i}}(\mathscr{T}^i), x_i \Big)$$

and we have to prove that $\sum_{i\in I_\theta^\uparrow}\delta_{(x_i,\mathcal{T}^i,\theta_i)}(dx,d\mathcal{T}',d\theta')$ is conditionally on $\Lambda_\theta(\mathcal{T},\mathcal{M})$ a Poisson point measure with intensity (3.35). Since a is arbitrary, it is enough to prove that the point measure $\mathcal{M}_a$, defined by

$$\mathcal{M}_a(dx,d\mathcal{T}',d\theta') = \sum_{i\in I_\theta^\uparrow} \mathbf{1}_{\{H_{x_i}\le a\}}\,\delta_{(x_i,\pi_{a-H_{x_i}}(\mathcal{T}^i),\theta_i)}(dx,d\mathcal{T}',d\theta'),$$

is conditionally on $\Lambda_{\theta,a}(\mathcal{T},\mathcal{M})$ a Poisson point measure with intensity :

$$\mathbf{1}_{[0,a]}(H_x)\,\mathbf{m}^{\Lambda_\theta(\mathcal{T},\mathcal{M})}(dx)\,\mathbf{1}_{(0,\theta]}(\theta')\,d\theta' \\ \left(2\beta(\pi_{a-H_x})_*\mathbb{N}^\psi(d\mathcal{T}') + \int_{(0,+\infty)} \Pi(dr)\,re^{-\theta' r}(\pi_{a-H_x})_*\mathbb{P}_r^\psi(d\mathcal{T}')\right). \quad (3.36)$$

Recall θ^* is the unique real number such that $\psi'_{\theta^*}(0)=0$, that is, such that ψ_{θ^*} is critical. Let Φ be a non-negative, measurable functional on $\Lambda_{\theta,a}(\mathcal{T},\mathcal{M})\times\mathbb{T}\times(0,\theta]$ and let F be a non-negative measurable functional on $\mathbb{T}$. Let

$$B = \mathbb{N}^\psi\left[F(\Lambda_{\theta,a}(\mathcal{T},\mathcal{M}))\exp(-\langle\mathcal{M}_a,\Phi\rangle)\right].$$

Thanks to Girsanov formula (3.22) and the special Markov property for critical branching mechanisms, we get:

$$\begin{aligned} B &= \mathbb{N}^{\psi_{\theta^*}}\left[F(\Lambda_{\theta,a}(\mathcal{T},\mathcal{M}))\exp(-\langle\mathcal{M}_a,\Phi\rangle)\exp\left(\theta^*\mathcal{Z}_a(\mathcal{T})+\psi(\theta^*)\int_0^a \mathcal{Z}_h(\mathcal{T})\,dh\right)\right] \\ &= \mathbb{N}^{\psi_{\theta^*}}\left[F(\Lambda_{\theta,a}(\mathcal{T},\mathcal{M}))\exp\left(\theta^*\mathcal{Z}_a(\Lambda_\theta(\mathcal{T},\mathcal{M}))+\psi(\theta^*)\int_0^a \mathcal{Z}_h(\Lambda_\theta(\mathcal{T},\mathcal{M}))\,dh\right)\right. \\ &\qquad\qquad \left.\exp\left(-\int \mathbf{m}^{\Lambda_{\theta,a}(\mathcal{T},\mathcal{M})}(dx)G(H_x,x,\theta)\right)\right], \end{aligned}$$

with $\mathbf{m}^{\Lambda_{\theta,a}(\mathcal{T},\mathcal{M})}(dx)=\mathbf{1}_{[0,a]}(H_x)\,\mathbf{m}^{\Lambda_\theta(\mathcal{T},\mathcal{M})}(dx)$ and $G(h,x,\theta)$ equal to:

$$\begin{aligned} \int_0^\theta d\theta' \Bigg\{ & 2\beta\mathbb{N}^{\psi_{\theta^*}}\left[1-\exp\left(-\Phi(x,\pi_{a-h}(\mathcal{T}),\theta')+\theta^*\mathcal{Z}_{a-h}(\mathcal{T})+\psi(\theta^*)\int_0^{a-h}\mathcal{Z}_t(\mathcal{T})\,dt\right)\right] \\ & + \int_{(0,+\infty)} \Pi_{\theta^*}(dr)re^{-\theta' r}\mathbb{E}_r^{\psi_{\theta^*}}\Bigg[1-\exp\Bigg(-\Phi(x,\pi_{a-h}(\mathcal{T}),\theta') \\ & \qquad +\theta^*\mathcal{Z}_{a-h}(\mathcal{T})+\psi(\theta^*)\int_0^{a-h}\mathcal{Z}_t(\mathcal{T})\,dt\Bigg)\Bigg]\Bigg\}. \end{aligned}$$

By using the Poisson decomposition of $\mathbb{P}_r^{\psi_{\theta^*}}$ (Proposition 3.12), we see that $G(h,x,\theta)$ can be written as:

$$G(h,x,\theta) = \int_0^\theta d\theta'\left\{2\beta g(h,x,\theta') + \int_{(0,\infty)} \Pi_{\theta^*}(dr)\;re^{-\theta' r}\left(1-\exp(-rg(h,x,\theta'))\right)\right\},$$

with

$$g(h,x,\theta') = \mathbb{N}^{\psi_{\theta^*}}\left[1-\exp\left(-\Phi(x,\pi_{a-h}(\mathcal{T}),\theta')+\theta^*\mathcal{Z}_{a-h}(\mathcal{T})+\psi(\theta^*)\int_0^{a-h}\mathcal{Z}_t(\mathcal{T})\,dt\right)\right].$$

Thanks to the Girsanov formula and (3.29), we get:

$$\begin{aligned}g(h,x,\theta') &= \mathbb{N}^{\psi_{\theta^*}}\left[(1-\exp(-\Phi(x,\pi_{a-h}(\mathcal{T}),\theta')))\exp\left(\theta^*\mathcal{Z}_{a-h}(\mathcal{T})+\psi(\theta^*)\int_0^{a-h}\mathcal{Z}_t(\mathcal{T})dt\right)\right]\\ &\qquad\qquad + \mathbb{N}^{\psi_{\theta^*}}\left[1-\exp\left(\theta^*\mathcal{Z}_{a-h}(\mathcal{T})+\psi(\theta^*)\int_0^{a-h}\mathcal{Z}_t(\mathcal{T}))dt\right)\right]\\ &= \mathbb{N}^{\psi}\left[1-\exp(-\Phi(x,\pi_{a-h}(\mathcal{T}),\theta'))\right]-\theta^*.\end{aligned}$$

With $\tilde{g}(h,x,\theta') = \mathbb{N}^{\psi}\left[1-\exp(-\Phi(x,\pi_{a-h}(\mathcal{T}),\theta'))\right]$ and thanks to (3.7), we get:

$$\begin{aligned}G(h,x,\theta) = \int_0^\theta d\theta'\Big\{2\beta\tilde{g}(h,x,\theta') + \int_{(0,\infty)}\Pi(dr)\; re^{-\theta' r}\left(1-\exp(-r\tilde{g}(h,x,\theta'))\right)\Big\}\\ +\psi(\theta^*)-\psi_\theta(\theta^*).\end{aligned}$$

Notice that from the definition of G we have g replaced by $\tilde{g}$, Π_{θ^*} replaced by Π and the additional term $\psi(\theta^*)-\psi_\theta(\theta^*)$. As $\int \mathbf{m}^{\Lambda_{\theta,a}(\mathcal{T},\mathcal{M})}(dx) = \int_0^a \mathcal{Z}_h(\Lambda_\theta(\mathcal{T}))dh$, we get:

$$\begin{aligned}B = \mathbb{N}^{\psi_{\theta^*}}\big[F(\Lambda_{\theta,a}(\mathcal{T},\mathcal{M}))R(\Lambda_{\theta,a}(\mathcal{T},\mathcal{M}))\\ \exp\left(\theta^*\mathcal{Z}_a(\Lambda_\theta(\mathcal{T},\mathcal{M}))+\psi_\theta(\theta^*)\int_0^a\mathcal{Z}_h(\Lambda_\theta(\mathcal{T},\mathcal{M}))dh\right)\Big],\end{aligned} \tag{3.37}$$

with

$$\begin{aligned}R(\mathcal{T}) = \exp\Big(-\int \mathbf{m}^{\mathcal{T}}(dx)\int_0^\theta d\theta'\Big[2\beta\tilde{g}(H_x,x,\theta')+\\ \int_{(0,\infty)}\Pi(dr)\; re^{-\theta' r}\left(1-\exp(-r\tilde{g}(H_x,x,\theta'))\right)\Big]\Big).\end{aligned} \tag{3.38}$$

Taking $\Phi = 0$ (and thus $R=1$) in (3.37) yields:

$$\begin{aligned}&\mathbb{N}^{\psi}[F(\Lambda_{\theta,a}(\mathcal{T},\mathcal{M}))]\\ &\quad = \mathbb{N}^{\psi_{\theta^*}}\left[F(\Lambda_{\theta,a}(\mathcal{T},\mathcal{M}))\exp\left(\theta^*\mathcal{Z}_a(\Lambda_\theta(\mathcal{T},\mathcal{M}))+\psi_\theta(\theta^*)\int_0^a\mathcal{Z}_h(\Lambda_\theta(\mathcal{T},\mathcal{M}))dh\right)\right].\end{aligned} \tag{3.39}$$

Using (3.39) with F replaced by FR gives:

$$\mathbb{N}^{\psi}\left[\exp(-\langle\mathcal{M}_a,\Phi\rangle)F(\Lambda_{\theta,a}(\mathcal{T},\mathcal{M}))\right] = B = \mathbb{N}^{\psi}\left[F(\Lambda_{\theta,a}(\mathcal{T},\mathcal{M}))R(\Lambda_{\theta,a}(\mathcal{T},\mathcal{M}))\right].$$

This implies that $\mathcal{M}_a$ is, conditionally on $\Lambda_{\theta,a}(\mathcal{T},\mathcal{M})$, a Poisson point measure with intensity (3.36). This ends the proof. □

An explicit construction of the growing process

In this section, we will construct the growth process using a family of Poisson point measures. Let ψ be a branching mechanism satisfying Assumptions 1 and 2. Let $\theta \in \Theta^{\psi}$. According to (3.20) and (3.7), we have:

$$\mathbf{N}^{\psi_\theta}[\mathscr{T} \in \bullet] = 2\beta \mathbb{N}^{\psi_\theta}[\mathscr{T} \in \bullet] + \int_{(0,+\infty)} \Pi(dr) r e^{-\theta r} \mathbb{P}_r^{\psi_\theta}(\mathscr{T} \in \bullet). \tag{3.40}$$

Let $\mathscr{T}^{(0)} \in \mathbb{T}$ with root $\emptyset$. For $q \in \Theta^{\psi}$ and $q \leq \theta$, we set:

$$\mathfrak{T}_q^{(0)} = \mathscr{T}^{(0)} \quad \text{and} \quad \mathbf{m}_q^{(0)} = \mathbf{m}^{\mathscr{T}^{(0)}}.$$

We define the w-trees grafted on $\mathscr{T}^{(0)}$ by recursion on their generation. We suppose that all the random point measures used for the next construction are defined on $\mathbb{T}$ under a probability measure $Q^{\mathscr{T}^{(0)}}(d\omega)$.

Suppose that we have constructed the family $((\mathfrak{T}_q^{(k)}, \mathbf{m}_q^{(n)}), 0 \leq k \leq n, q \in \Theta^{\psi} \cap (-\infty, \theta))$. We write

$$\mathfrak{T}^{(n)} = \bigsqcup_{q \in \Theta^{\psi},\ q \leq \theta} \mathfrak{T}_q^{(n)}.$$

We define the $(n+1)$-th generation as follows. Conditionally on all trees from generations smaller than n, $(\mathfrak{T}_q^{(k)},\ 0 \leq k \leq n,\ q \in \Theta^{\psi} \cap (-\infty, \theta))$, let

$$\mathcal{N}_\theta^{n+1}(dx, d\mathscr{T}, dq) = \sum_{j \in J^{(n+1)}} \delta_{(x_j, \mathscr{T}^j, \theta_j)}(dx, d\mathscr{T}, dq)$$

be a Poisson point measure on $\mathfrak{T}^{(n)} \times \mathbb{T} \times \Theta^{\psi}$ with intensity:

$$\mu_\theta^{n+1}(dx, d\mathscr{T}, dq) = \mathbf{m}_q^{(n)}(dx) \mathbf{N}^{\psi_q}[d\mathscr{T}]\, \mathbf{1}_{\{q \leq \theta\}}\, dq.$$

For $q \in \Theta^{\psi}$ and $q \leq \theta$, we set

$$J_q^{(n+1)} = \{j \in J^{(n+1)},\ q < \theta_j\}$$

and we define the tree $\mathfrak{T}_q^{(n+1)}$ and the mass measure $\mathbf{m}_q^{(n+1)}$ by:

$$\mathfrak{T}_q^{(n+1)} = \mathfrak{T}_q^{(n)} \circledast_{j \in J_q^{(n+1)}} (\mathscr{T}^j, x_j) \quad \text{and} \quad \mathbf{m}_q^{(n+1)} = \sum_{j \in J_q^{(n+1)}} \mathbf{m}^{\mathscr{T}^j}(dx).$$

Notice that by construction, $(\mathfrak{T}_q^{(n)}, n \in \mathbb{N})$ is a non-decreasing sequence of trees. We set $\mathfrak{T}_q$ the completion of $\cup_{n \in \mathbb{N}} \mathfrak{T}_q^{(n)}$, which is a real tree with root $\emptyset$ and obvious metric $d^{\mathfrak{T}_q}$, and we define a mass measure on $\mathfrak{T}_q$ by $\mathbf{m}^{\mathfrak{T}_q} = \sum_{n \in \mathbb{N}} \mathbf{m}_q^{(n)}$.

For $q \in \Theta^{\psi}$ and $q < \theta$, we consider $\mathscr{F}_q$ the σ-field generated by $\mathfrak{T}^{(0)}$ and the sequence of random point measures $(\mathbf{1}_{\{q' \in [q,\theta]\}} \mathcal{N}_\theta^{(n)}(dx, d\mathscr{T}, dq'), n \in \mathbb{N})$. We set $\mathcal{N}_\theta = \sum_{n \in \mathbb{N}} \mathcal{N}_\theta^n$. The backward random point process $q \mapsto \mathbf{1}_{\{q \leq q'\}} \mathcal{N}_\theta(dx, d\mathscr{T}, dq')$ is by construction adapted to the backward filtration $(\mathscr{F}_q, q \in \Theta^{\psi} \cap (-\infty, \theta])$.

The proof of the following result is postponed to Section 3.3.

Theorem 3.26. *Let ψ be a branching mechanism satisfying Assumptions 1 and 2. Under $Q^{\psi_\theta} := \mathbb{N}^{\psi_\theta}[d\mathscr{T}^{(0)}]Q^{\mathscr{T}^{(0)}}(d\omega)$, the process*

$$\left(\left(\mathfrak{T}_q, d^{\mathfrak{T}_q}, \emptyset, \mathbf{m}^{\tilde{\mathfrak{T}}_q}\right), q \in \Theta^\psi \cap (-\infty, \theta]\right)$$

is a $\mathbb{T}$-valued backward Markov process with respect to the backward filtration $\mathscr{F}^\theta = (\mathscr{F}_q, q \in \Theta^\psi \cap (-\infty, \theta])$. It is distributed as $((\mathscr{T}_q, \mathbf{m}^{\mathscr{T}_q}), q \in \Theta^\psi \cap (-\infty, \theta])$ under $\mathbb{N}^\psi$.

Notice the Theorem in particular entails that $(\mathfrak{T}_q, d^{\mathfrak{T}_q}, \emptyset, \mathbf{m}^{\tilde{\mathfrak{T}}_q})$ is a w-tree for all q.

We shall use the following lemma.

Lemma 3.27. *Let ψ be a branching mechanism satisfying Assumptions 1 and 2. Let K be a measurable non-negative process (as a function of q) defined on $\mathbf{R}_+ \times \mathbb{T} \times \mathbb{T}$ which is predictable with respect to the backward filtration $\mathscr{F}^\theta$. We have:*

$$Q^{\psi_\theta}\left[\int \mathscr{N}_\theta(dx, d\mathscr{T}, dq)\, K(q, \mathfrak{T}_q, \mathfrak{T}_{q-})\right] = Q^{\psi_\theta}\left[\int K\Big(q, \mathfrak{T}_q, \mathfrak{T}_q \circledast (\mathscr{T}, x)\Big)\, \mu_\theta(dx, d\mathscr{T}, dq)\right],$$

where $\mu_\theta(dx, d\mathscr{T}, dq) = \sum_{n \in \mathbb{N}^*} \mu^n(dx, dT, dq) = \mathbf{m}^{\mathfrak{T}_q}(dx)\mathbf{N}^{\psi_q}[d\mathscr{T}]\, \mathbf{1}_{\{q \in \Theta^\psi, q \le \theta\}}\, dq$.

This means that the predictable compensator of $\mathscr{N}_\theta$ is given by:

$$\mu_\theta(dx, d\mathscr{T}, dq) = \mathbf{m}^{\mathfrak{T}_q}(dx)\mathbf{N}^{\psi_q}[d\mathscr{T}]\, \mathbf{1}_{\{q \in \Theta^\psi, q \le \theta\}}\, dq.$$

Notice this construction does not fit in the usual framework of random point measures as the support at time q of the predictable compensator is the (predictable backward in time) random set $\mathfrak{T}_q \times \mathbb{T} \times \Theta^\psi$.

Proof. Based on the recursive construction, we have:

$$\begin{aligned} &Q^{\psi_\theta}\left[\int \mathscr{N}_\theta(dx, d\mathscr{T}, dq)\, K(q, \mathfrak{T}_q, \mathfrak{T}_{q-})\right] \\ &\quad = \sum_{n=0}^{+\infty} Q^{\psi_\theta}\left[Q^{\psi_\theta}\left[\int \mathscr{N}_\theta^n(dx, d\mathscr{T}, dq)\, K(q, \mathfrak{T}_q, \mathfrak{T}_q \circledast (\mathscr{T}, x))\,\Big|\, (\mathfrak{T}_s^{(k)},\ k \le n,\ s \le \theta)\right]\right]. \end{aligned}$$

Now, by construction, we have that:

$$\mathfrak{T}_q = \mathfrak{T}_q^{(n)} \circledast_{j \in J_q^{(n)}} (\tilde{\mathscr{T}}_j, x_j)$$

for $\tilde{\mathscr{T}}_j = \mathfrak{T}_q \backslash \mathfrak{T}_q^{(x_j, \emptyset)}$ which is a measurable function of $\mathbf{1}_{\{q' > q\}} \mathscr{N}_\theta^n(dx, d\mathscr{T}, dq')$ and of the point measures $\mathbf{1}_{\{q' > q\}} \mathscr{N}_\theta^\ell(dx, d\mathscr{T}, dq')$ for $\ell \ge n+1$. Therefore, applying a Palm formula with the function

$$\begin{aligned} F_n\Big(q, \mathscr{T}, x, \sum_{j \in J^{(n)}, q_j > q} \delta_{(x_j, \mathscr{T}^j, \theta_j)}\Big) = Q^{\psi_\theta}\Big[K\Big(q, \mathfrak{T}_q^{(n)} \circledast_{j \in J_q^{(n)}} (\tilde{\mathscr{T}}_j, x_j),& \\ \mathfrak{T}_q^{(n)} \circledast_{j \in J_q^{(n)}} (\tilde{\mathscr{T}}_j, x_j) \circledast (\mathscr{T}, x)\Big)\,\Big|\, (\mathfrak{T}_s^{(k)},\ k \le n,\ s \le \theta), \mathscr{N}_\theta^n\Big],& \end{aligned}$$

we get:

$$
\begin{aligned}
&Q^{\psi_\theta}\left[\int \mathcal{N}_\theta(dx,d\mathcal{T},dq)\, K(q,\mathfrak{T}_q,\mathfrak{T}_{q-})\right]\\
&= \sum_{n=0}^{+\infty} Q^{\psi_\theta}\Big[Q^{\psi_\theta}\Big[\int \mathcal{N}_\theta^n(dx,d\mathcal{T},dq)\\
&\qquad\qquad F_n\Big(q,\mathcal{T},x,\sum_{j\in J^{(n)},q_j>q}\delta_{(x_j,\mathcal{T}^j,\theta_j)}\Big)\Big|(\mathfrak{T}_s^{(k)},\ k\le n,\ s\le\theta)\Big]\Big]\\
&= \sum_{n=0}^{+\infty} Q^{\psi_\theta}\Big[Q^{\psi_\theta}\Big[\int \mu_\theta^n(dx,d\mathcal{T},dq)\\
&\qquad\qquad F_n\Big(q,\mathcal{T},x,\sum_{j\in J^{(n)},q_j>q}\delta_{(x_j,\mathcal{T}^j,\theta_j)}\Big)\Big|(\mathfrak{T}_s^{(k)},\ k\le n,\ s\le\theta)\Big]\Big]\\
&= \sum_{n=0}^{+\infty} Q^{\psi_\theta}\left[Q^{\psi_\theta}\left[\int \mu_\theta^n(dx,d\mathcal{T},dq)\, K\Big(q,\mathfrak{T}_q^{(n)}\circledast_{j\in J_q^{(n)}}(\tilde{\mathcal{T}}_j,x_j),\right.\right.\\
&\qquad\qquad \left.\left.\mathfrak{T}_q^{(n)}\circledast_{j\in J_q^{(n)}}(\tilde{\mathcal{T}}_j,x_j)\circledast(\mathcal{T},x)\Big)\,\Big|\,(\mathfrak{T}_s^{(k)},\ k\le n,\ s\le\theta)\right]\right]\\
&= \sum_{n=0}^{+\infty} Q^{\psi_\theta}\left[\int \mu_q^n(dx,d\mathcal{T},dq)\, K\Big(q,\mathfrak{T}_q,\mathfrak{T}_q\circledast(\mathcal{T},x)\Big)\right]\\
&= Q^{\psi_\theta}\left[\int K\Big(q,\mathfrak{T}_q,\mathfrak{T}_q\circledast(T,x)\Big)\,\mu_\theta(dx,d\mathcal{T},dq)\right].
\end{aligned}
$$

□

It can be noticed that the map $q\mapsto\mathfrak{T}_q$ is non-decreasing càdlàg (backwards in time) and that we have, for $j\in\cup_{n\in\mathbb{N}}J^{(n)}$, $x_j\in\mathfrak{T}_{\theta_j}$: $\mathfrak{T}_{\theta_j-}=\mathfrak{T}_{\theta_j}\circledast(\mathcal{T}^j,x_j)$. In particular, we can recover the random measure $\mathcal{N}_\theta$ from the jumps of the process $(\mathfrak{T}_q,q\in\Theta^\psi\cap(-\infty,\theta])$. This and the natural compatibility relation of $\mathcal{N}_\theta$ with respect to θ gives the next Corollary.

Corollary 3.28. *Let ψ be a branching mechanism satisfying Assumptions 1 and 2. Let $(\mathcal{T}_\theta,\theta\in\Theta^\psi)$ be defined under $\mathbb{N}^\psi$. Let*

$$\mathcal{N}=\sum_{j\in J}\delta_{(x_j,\mathcal{T}^j,\theta_j)}$$

be the random point measure defined as follows:

- *The set $\{\theta_j;j\in J\}$ is the set of jumping times of the process $(\mathcal{T}_\theta,\theta\in\Theta^\psi)$: for $j\in J$, $\mathcal{T}_{\theta_j-}\neq\mathcal{T}_{\theta_j}$.*
- *The real tree $\mathcal{T}^j$ is the closure of $\mathcal{T}_{\theta_j-}\setminus\mathcal{T}_{\theta_j}$.*
- *The point x_j is the root of $\mathcal{T}^j$ (that is x_j is the only element $y\in\mathcal{T}_{\theta_j-}$ such that $x\in\mathcal{T}^j$ implies $[\![y,x]\!]\subset\mathcal{T}^j$).*

Then the backward point process $\theta\mapsto\mathbf{1}_{\{\theta\le q'\}}\mathcal{N}(dx,d\mathcal{T},dq')$ defined on Θ^ψ has predictable compensator:

$$\mu(dx,d\mathcal{T},dq)=\mathbf{m}^{\mathcal{T}_q}(dx)\mathbf{N}^{\psi_q}[d\mathcal{T}]\,\mathbf{1}_{\{q\in\Theta^\psi\}}\,dq,$$

with respect to the backward left-continuous filtration $\mathscr{F} = (\mathscr{F}_\theta, \theta \in \Theta^\psi)$ *defined by:*

$$\mathscr{F}_\theta = \sigma((x_j, \mathscr{T}^j, \theta_j); \theta \le \theta_j) = \sigma(\mathscr{T}_{q-}; \theta \le q).$$

More precisely, for any non-negative predictable process K *with respect to the backward filtration* $\mathscr{F}$*, we have:*

$$\mathbb{N}^\psi\left[\int \mathcal{N}(dx, d\mathscr{T}, dq)\, K\Big(q, \mathscr{T}_q, \mathscr{T}_{q-}\Big)\right] = \mathbb{N}^\psi\left[\int \mu(dx, dT, dq)\, K\Big(q, \mathscr{T}_q, \mathscr{T}_q \circledast (T, x)\Big)\right]. \quad (3.41)$$

Remark 5. *Notice that Assumption 2 is assumed only for technical measurability condition, see Remark 2. We conjecture that this results holds also if Assumption 2 is not in force.*

As a consequence, thanks to property 3 of Theorem 3.24, we get, with the convention $\sup \emptyset = \theta_\infty$, that:

$$A = \sup\{\theta_j, j \in J \text{ and } \sigma^j = +\infty\} \quad \text{with} \quad \sigma_j = \mathbf{m}^{\mathscr{T}^j}(\mathscr{T}^j).$$

Proof of Theorem 3.26

By construction, it is clear that the process $(\mathfrak{T}_q, q \in \Theta^\psi \cap (-\infty, \theta])$ is a backward Markov process with respect to the backward filtration $(\mathscr{F}_q, q \in \Theta^\psi \cap (-\infty, \theta])$. By construction this process is càglàd in backward time. Since the process $(\mathscr{T}_q, q \in \Theta^\psi)$ is a forward càdlàg Markov process, it is enough to check that for $\theta_0 \in \Theta^\psi$, such that $\theta_0 < \theta$, the two dimensional marginals $(\mathfrak{T}_{\theta_0}, \mathfrak{T}_\theta)$ and $(\mathscr{T}_{\theta_0}, \mathscr{T}_\theta)$ have the same distribution.

Replacing ψ by ψ_{θ_0}, we can assume that $\theta_0 = 0$ and $0 < \theta$. We shall decompose the big tree $\mathscr{T}_0$ conditionally on the small tree $\mathscr{T}_\theta$ by iteration. This decomposition is similar to the one which appears in [AD07] or [Voi10] for the fragmentation of the (sub)critical Lévy tree, but roughly speaking the fragmentation is here frozen but for the fragment containing the root.

We set $\mathscr{T}^{(0)} = \mathscr{T}_\theta$ and $\tilde{\mathbf{m}}^{(0)} = \mathbf{m}^{\mathscr{T}_\theta}$, so that $(\mathfrak{T}^{(0)}, \mathbf{m}^{(0)})$ and $(\mathscr{T}^{(0)}, \tilde{\mathbf{m}}^{(0)})$ have the same distribution. Recall notation $\mathcal{M}^\uparrow$ from (3.33) as well as (3.34): $\mathscr{T}_0 = \mathscr{T}^{(0)} \circledast_{i \in I_\theta^{\uparrow,1}} (\mathscr{T}^i, x_i)$, where we write $I_\theta^{\uparrow,1} = I_\theta^\uparrow$ and where $\mathscr{P}^1 = \sum_{i \in I_\theta^{\uparrow,1}} \delta_{(x_i, \mathscr{T}^i, \theta_i)}$ is, conditionally on $\mathscr{T}^{(0)}$, a Poisson point measure with intensity:

$$\nu^1(dx, d\mathscr{T}', dq) = \tilde{\mathbf{m}}^{(0)}(dx)\Big(2\beta\mathbb{N}^\psi[d\mathscr{T}'] + \int_{(0,+\infty)} \Pi(dr)\, r e^{-qr}\mathbb{P}_r^\psi(d\mathscr{T}')\Big)\, \mathbf{1}_{(0,\theta]}(q)\, dq.$$

For $i \in I_\theta^{\uparrow,1}$, we define the sub-tree of $\mathscr{T}^i$:

$$\tilde{\mathscr{T}}^i = \Big\{x \in \mathscr{T}^i; \mathcal{M}^\uparrow(]\!]x_i, x[\![\times [0, \theta_i]) = 0\Big\}.$$

Since $\mathscr{T}^i$ is distributed according to $\mathbb{N}^\psi$ (or to $\mathbb{P}_{r_i}^\psi$ for some $r_i > 0$), using the property of Poisson point measures, we have that conditionally on $\mathscr{T}^0$ and θ_i, the tree $\tilde{\mathscr{T}}^i$ is distributed as

$\Lambda_{\theta_i}(\mathcal{T},\mathcal{M})$ under $\mathbb{N}^\psi$ (or under $\mathbb{P}^\psi_{r_i}$) that is the distribution of $\tilde{\mathcal{T}}^i$ is $\mathbb{N}^{\psi_{\theta_i}}[d\mathcal{T}]$ (or $\mathbb{P}^{\psi_{\theta_i}}_{r_i}(d\mathcal{T})$), thanks to the special Markov property. Furthermore we have $\mathcal{T}^i = \tilde{\mathcal{T}}^i \circledast_{i'\in I^{\uparrow,2}_{\theta,i}} (\mathcal{T}^{i'}, x_{i'})$ where

$$\sum_{i'\in I^{\uparrow,2}_{\theta,i}} \delta_{(x_{i'},\mathcal{T}^{i'},\theta_{i'})}$$

is, conditionally on $\mathcal{T}^{(0)}$ and $\tilde{\mathcal{T}}^i$ a Poisson point measure on $\tilde{\mathcal{T}}^i \times \mathbb{T} \times (0,\theta]$ with intensity:

$$\mathbf{m}^{\tilde{\mathcal{T}}^i}(dx)\Big(2\beta\mathbb{N}^\psi(d\mathcal{T}') + \int_{(0,+\infty)} \Pi(dr)\, r\mathrm{e}^{-qr}\mathbb{P}^\psi_r(d\mathcal{T}')\Big)\, \mathbf{1}_{[0,\theta_i)}(q)\, dq.$$

Thus we deduce, using again the special Markov property, that:

$$\tilde{\mathcal{N}}^1_\theta(dx,d\mathcal{T},dq) = \sum_{i\in I^{\uparrow,1}} \delta_{(x_i,\tilde{\mathcal{T}}^i,\theta_i)}(dx,d\mathcal{T},dq)$$

is conditionally on $\mathcal{T}^0$ a Poisson point measure on $\mathcal{T}^{(0)} \times \mathbb{T} \times \Theta^\psi$ with intensity:

$$\tilde{\mu}^1(dx,d\mathcal{T},dq) = \tilde{\mathbf{m}}^{(0)}_q(dx)\mathbf{N}^{\psi_q}[d\mathcal{T}]\, \mathbf{1}_{[0,\theta)}(q)\, dq,$$

with $\tilde{\mathbf{m}}^{(0)}_q(dx) = \tilde{\mathbf{m}}^{(0)}(dx)$. We set $\mathcal{T}^{(1)} = \mathcal{T}^{(0)} \circledast_{i\in I^{\uparrow,1}_\theta} (\tilde{\mathcal{T}}^i, x_i)$ for the first generation tree and for $q\in[0,\theta]$:

$$\tilde{\mathbf{m}}^{(1)}_q(dx) = \sum_{i\in I^{\uparrow,1}_\theta} \mathbf{m}^{\tilde{\mathcal{T}}^i}(dx)\mathbf{1}_{[0,\theta_i)}(q).$$

See Figure 3.2 for a simplified representation. We get that $(\mathfrak{T}^{(1)}_\theta, (\mathbf{m}^{(1)}_q, q\in[0,\theta]), \mathfrak{T}^{(0)}, \mathbf{m}^{\mathfrak{T}^{(0)}})$ and $(\mathcal{T}^{(1)}, (\tilde{\mathbf{m}}^{(1)}_q, q\in[0,\theta]), \mathcal{T}^{(0)}, \tilde{\mathbf{m}}^{(0)})$ have the same distribution.

Furthermore, by collecting all the trees grafted on $\mathcal{T}^{(1)}$, we get that

$$\mathcal{T} = \mathcal{T}^{(1)} \circledast_{i'\in I^{\uparrow,2}_\theta} (\mathcal{T}^{i'}, x_{i'}),$$

where $I^{\uparrow,2}_\theta = \cup_{i\in I^{\uparrow,1}_\theta} I^{\uparrow,2}_{\theta,i}$ and where

$$\mathcal{P}^2 = \sum_{i'\in I^{\uparrow,2}_\theta} \delta_{(x_{i'},\mathcal{T}^{i'},\theta_{i'})}$$

is, conditionally on $(\mathcal{T}^{(1)}, (\tilde{\mathbf{m}}^{(1)}_q, q\in[0,\theta]), \mathcal{T}^{(0)}, \tilde{\mathbf{m}}^{(0)})$ a Poisson point measure on $\mathcal{T}^{(1)} \times \mathbb{T} \times (0,\theta]$ with intensity:

$$\nu^2(dx,d\mathcal{T},dq) = \tilde{\mathbf{m}}^{(1)}_q(dx)\, \Big(2\beta\mathbb{N}^\psi(d\mathcal{T}') + \int_{(0,+\infty)} \Pi(dr)\, r\mathrm{e}^{-qr}\mathbb{P}^\psi_r(d\mathcal{T}')\Big)\, \mathbf{1}_{[0,\theta]}(q)\, dq.$$

Notice that:

$$\mathcal{T}^{(1)} = \{x\in\mathcal{T}_0; \mathcal{M}^\uparrow([\![\emptyset,x[\![\times[0,\theta]) \le 1\} \quad \text{and} \quad \tilde{\mathbf{m}}^{(1)}_\theta(dx) + \tilde{\mathbf{m}}^{(0)}(dx) = \mathbf{1}_{\mathcal{T}^{(1)}}(x)\, \mathbf{m}^{\mathcal{T}_0}(dx). \tag{3.42}$$

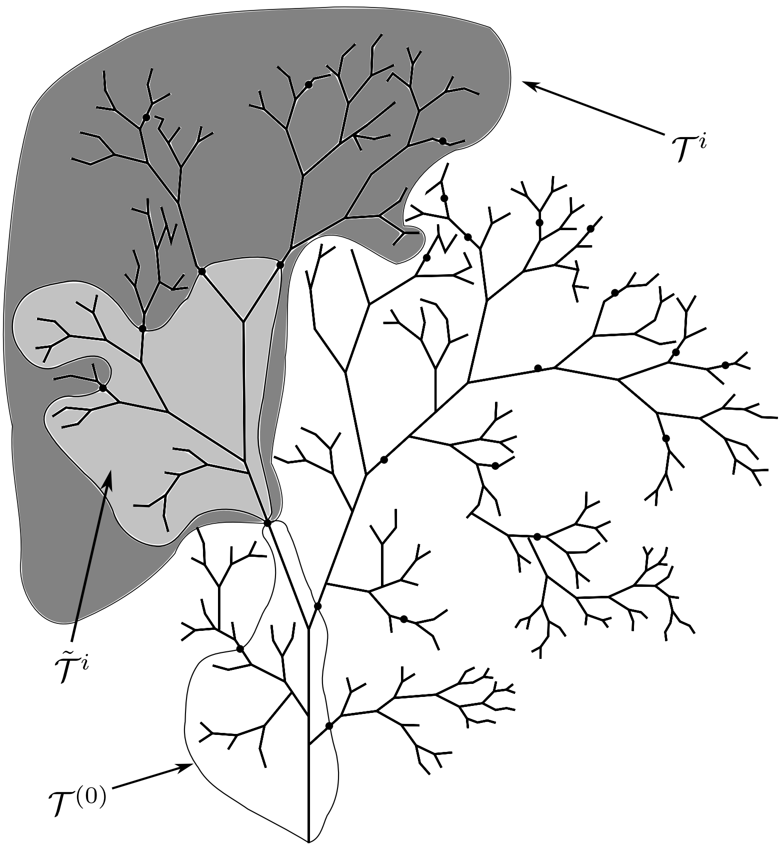

Figure 3.2: The tree $\mathscr{T}_0$, $\mathscr{T}^{(0)}$, and a tree $\mathscr{T}^i$ and its sub-tree $\tilde{\mathscr{T}}^i$ belonging to the first generation tree $\mathscr{T}^{(1)} \setminus \mathscr{T}^{(0)}$.

Then we can iterate this construction, and by taking increasing limits we obtain that the pair $((\cup_{n\in\mathbb{N}} \mathfrak{T}_\theta^{(n)}, \sum_{n\in\mathbb{N}} \mathbf{m}_\theta^{(n)}), \mathfrak{T}_0)$ has the same distribution as $(\mathscr{T}', \mathscr{T}^{(0)})$, where:

$$\mathscr{T}' = \left\{x \in \mathscr{T}_0; \mathscr{M}^{\uparrow}([\![\emptyset, x[\![\times [0,\theta]) < +\infty\right\} \quad \text{and} \quad \tilde{\mathbf{m}}'(dx) = \mathbf{1}_{\mathscr{T}'}(x)\, \mathbf{m}^{\mathscr{T}_0}(dx).$$

To conclude, we need to check first that the completion of $\mathscr{T}'$ is $\mathscr{T}_0$, or as $\mathscr{T}_0$ is complete that the closure of $\mathscr{T}'$ as a subset of $\mathscr{T}_0$ is exactly $\mathscr{T}_0$ and then that $\mathbf{m}^{\mathscr{T}_0}(\mathscr{T}'^c) = 0$.

Notice that $\mathscr{M}^{\uparrow}$ has less marks than $\mathscr{M}$. Then Proposition 1.2 in [AD07] in the case when $\beta = 0$ or an elementary adaptation of it in the general framework of [Voi10], gives there is no loss of mass in the fragmentation process. This implies that, if ψ is (sub)critical, then:

$$\mathbf{m}^{\mathscr{T}_0}(\{x \in \mathscr{T}_0; \mathscr{M}([\![\emptyset, x[\![\times [0,\theta]) = \infty\} = 0. \tag{3.43}$$

Then, if ψ is super-critical, by considering the restriction of $\mathscr{T}_0$ up to level a, $\pi_a(\mathscr{T}_0)$, and using a Girsanov transformation from Definition 3.18 with $\theta = \theta^*$ and (3.43), we deduce that (3.43) holds for $\pi_a(\mathscr{T}_0)$. Since a is arbitrary, we deduce by monotone convergence that (3.43) holds also in the super-critical case. Thus we have $\mathbf{m}^{\mathcal{T}_0}(\mathcal{T}'^c) = 0$. Since the closed support of $\mathbf{m}^{\mathcal{T}_0}$ is the set of leaves $\mathrm{Lf}(\mathcal{T}_0)$, we deduce that $\mathrm{Lf}(\mathcal{T}')$ is dense in $\mathrm{Lf}(\mathcal{T}_0)$ and, as $\mathcal{T}'$ and $\mathcal{T}_0$ have the same root, that $\mathrm{Sk}(\mathcal{T}') = \mathrm{Sk}(\mathcal{T}_0)$. This implies that the closure of $\mathcal{T}'$ is $\mathcal{T}_0$. This ends the proof.

3.4 Application to overshooting

We assume that ψ is critical, $\theta_\infty < 0$ and Assumptions 1 and 2 hold. We shall write u^θ (resp. b^θ) for the solution of (3.11) (resp. (3.12)) when ψ is replaced by ψ_θ, for $a \geq 0$, $h > 0$ and $t \in [0, h)$:

$$\int_{u^\theta(a,\lambda)}^{\lambda} \frac{dr}{\psi_\theta(r)} = a, \quad \text{and} \quad b_h^\theta(t) = b^\theta(h-t) \quad \text{with} \quad \int_{b^\theta(h)}^{\infty} \frac{dr}{\psi_\theta(r)} = h. \tag{3.44}$$

We have $u^\theta(a, b^\theta(h-a)) = b^\theta(h)$. Notice that $\partial_h b^\theta(h)/\psi_\theta(b^\theta(h)) = -1$ and also that we have $\partial_\lambda u^\theta(a,\lambda) = \psi_\theta(u^\theta(a,\lambda))/\psi_\theta(\lambda)$ which implies that:

$$\partial_\lambda u^\theta\left(a, b^\theta(h-a)\right) = \frac{\psi_\theta(b^\theta(h))}{\psi_\theta(b^\theta(h-a))} = -\frac{\psi_\theta(b^\theta(h))}{\psi_\theta(b^\theta(h-a))^2}\partial_h b^\theta(h-a). \tag{3.45}$$

We set for $\theta \in \Theta^\psi$ and $\lambda \geq 0$:

$$\gamma_\theta(\lambda) = \psi_\theta'(\lambda) - \psi_\theta'(0) = \psi'(\lambda+\theta) - \psi'(\theta) = \partial_\theta \psi_\theta(\lambda). \tag{3.46}$$

Notice the function γ_θ is non-negative and non-decreasing.

Recall that $\bar\theta = \psi^{-1} \circ \psi(\theta)$. We deduce from (3.44) that for $\theta \in \Theta^\psi$, $\theta < 0$ and $h > 0$:

$$\bar\theta + b^{\bar\theta}(h) = \theta + b^\theta(h) \quad \text{and} \quad \psi_{\bar\theta}(b^{\bar\theta}(h)) = \psi_\theta(b^\theta(h)). \tag{3.47}$$

Exit times

Let $h > 0$. We are interested in the first time when the process of growing trees exceeds height h, in the following sense.

Definition 3.29. *The first exit time out of h is the (possibly infinite) number A_h defined by*

$$A_h = \sup\left\{\theta \in \Theta^\psi,\ H_{\max}(\mathcal{T}_\theta) > h\right\},$$

with the convention that $\sup \emptyset = \theta_\infty$.

The constraint not to be higher than h will be coded by the function $b^\theta(h)$ which is the probability (under $\mathbb{N}^\psi$) for the tree $\mathcal{T}^\theta$ of having maximal height larger than h. By definition of the function b, we have for $\theta \in \Theta^\psi$:

$$\mathbb{N}^\psi[\theta \leq A_h] = \mathbb{N}^\psi\left[H_{\max}(\mathcal{T}_\theta) \geq h\right] = b^\theta(h). \tag{3.48}$$

Proposition 3.30. *Let ψ be a critical branching mechanism with $\theta_\infty < 0$ and satisfying Assumptions 1 and 2. The function $\theta \mapsto b_h^\theta$ is of class $\mathscr{C}^1$ on $(\theta_\infty, +\infty)$. And, under $\mathbb{N}^\psi$, the distribution of A_h on $(\theta_\infty, +\infty)$ has density $\theta \mapsto -\partial_\theta b^\theta(h)$ with respect to the Lebesgue measure. We also have the following expression for the density of A_h on $(\theta_\infty, +\infty)$. Let $\theta_\infty < \theta$ and $h > 0$. Then:*

$$-\partial_\theta b^\theta(h) = \psi_\theta\left(b^\theta(h)\right)\int_0^h da\,\frac{\gamma_\theta(b^\theta(a))}{\psi_\theta(b^\theta(a))} = \int_0^h da\,\gamma_\theta\left(b^\theta(h-a)\right)\,\mathrm{e}^{-\psi'(\theta)a-\int_0^a dx\,\gamma_\theta(b^\theta(h-x))}.$$

Notice that the distribution of A_h might have an atom at θ_∞.

Proof. Notice that for $\theta_\infty < \theta$, we have $\lim_{\lambda\to+\infty}\psi''(\lambda) = \beta$ and $\lim_{\lambda\to+\infty}\psi'(\lambda) = +\infty$. In particular $\psi'_\theta(\lambda)/\psi_\theta(\lambda)$ is bounded for λ large enough. This implies that $\int^{+\infty} dr\,\psi'_\theta(r)/\psi_\theta(r)^2$ is finite thanks to Assumption 2. We deduce that the function $\theta \mapsto b_h^\theta$ is of class $\mathscr{C}^1$ on $(\theta_\infty, +\infty)$ and, thanks to (3.48), that under $\mathbb{N}^\psi$, the distribution of A_h on $(\theta_\infty, +\infty)$ has density $\theta \mapsto -\partial_\theta b^\theta(h)$ with respect to the Lebesgue measure.

Taking the derivative with respect to θ in the last term of (3.44), using (3.46) and the change of variable $r = b^\theta(a)$ gives the first equality of the Proposition:

$$-\partial_\theta b^\theta(h) = \psi_\theta\left(b^\theta(h)\right)\int_{b^\theta(h)}^{+\infty} dr\,\frac{\gamma_\theta(r)}{\psi_\theta(r)^2} = \psi_\theta\left(b^\theta(h)\right)\int_0^h da\,\frac{\gamma_\theta(b^\theta(a))}{\psi_\theta(b^\theta(a))}. \tag{3.49}$$

From (3.44) we get that $\partial_t b_h^\theta(t) = \psi_\theta(b_h^\theta(t))$. Hence, we have:

$$\int_0^t \psi'_\theta\left(b_h^\theta(r)\right) dr = \int_0^t \frac{\psi'_\theta(b_h^\theta(r))}{\psi_\theta(b_h^\theta(r))}\partial_t b_h^\theta(r)\,dr = \log\left(\frac{\psi_\theta(b_h^\theta(t))}{\psi_\theta(b_h^\theta(0))}\right).$$

This gives:

$$\int_0^t \gamma_\theta\left(b_h^\theta(r)\right) dr = \int_0^t \psi'_\theta\left(b_h^\theta(r)\right) dr - t\psi'(\theta) = \log\left(\frac{\psi_\theta(b_h^\theta(t))}{\psi_\theta(b_h^\theta(0))}\right) - t\psi'(\theta). \tag{3.50}$$

We deduce that:

$$\int_0^h da\,\gamma_\theta\left(b^\theta(h-a)\right)\mathrm{e}^{-\psi'(\theta)a-\int_0^a dx\,\gamma_\theta(b^\theta(h-x))} = \psi_\theta\left(b^\theta(h)\right)\int_0^h da\,\frac{\gamma_\theta(b^\theta(a))}{\psi_\theta(b^\theta(a))}.$$

This proves the second equality of the Proposition. □

Since we will also be dealing with super-critical trees, there is always the positive probability that in the Poisson process of trees an infinite tree arises, which will be grafted onto the process, effectively making it infinite and thus outgrowing height h. In the next proposition, we will compute the conditional distribution of overshooting time A_h, given A. Note that we always have $A \le A_h$.

Proposition 3.31. *Let ψ be a critical branching mechanism with $\theta_\infty < 0$ and satisfying Assumptions 1 and 2. For $\theta_\infty < \theta_0 < \theta$ and $\theta_0 < 0$ (that is ψ_{θ_0} super-critical), we have, with $\hat{\theta} = \bar{\theta}_0 - \theta_0 + \theta$:*

$$\mathbb{N}^\psi[A_h \geq \theta | A = \theta_0] = 1 - \psi'(\hat{\theta})\psi_{\hat{\theta}}\left(b^{\hat{\theta}}(h)\right)\int_{b^{\hat{\theta}}(h)}^{+\infty} \frac{dr}{\psi_{\hat{\theta}}(r)^2},$$
$$\mathbb{N}^\psi[A_h = A | A = \theta_0] = \psi'(\bar{\theta}_0)\psi_{\bar{\theta}_0}\left(b^{\bar{\theta}_0}(h)\right)\int_{b^{\bar{\theta}_0}(h)}^{+\infty} \frac{dr}{\psi_{\bar{\theta}_0}(r)^2}.$$

Since $\psi_{\bar{\theta}_0}$ is sub-critical, we have $\psi'(\bar{\theta}_0) > 0$ and $\psi_{\bar{\theta}_0}(r) \sim r\psi'(\bar{\theta}_0)$ when r goes down to 0. Since $\lim_{h\to+\infty} b^{\bar{\theta}_0}(h) = 0$, we deduce that:

$$\lim_{h\to+\infty} \mathbb{N}^\psi[A_h = A | A = \theta_0] = 1.$$

This has a straightforward explanation. If h is very large, with high probability the process up to A will not have crossed height h, so that the first jump to cross height h will correspond to the grafting time of the first infinite tree which happens at the ascension time A.

We also deduce from (3.47) that:

$$\mathbb{N}^\psi[A_h = A | A = \theta_0] = \psi'(\bar{\theta}_0)\psi_{\theta_0}\left(b^{\theta_0}(h)\right)\int_{b^{\theta_0}(h)}^{+\infty} \frac{dr}{\psi_{\theta_0}(r)^2}. \tag{3.51}$$

Proof. We use the notation $\mathcal{Z}_h^\theta = \mathcal{Z}_h(\mathcal{T}^\theta)$ and $\mathcal{Z}_h = \mathcal{Z}_h(\mathcal{T}^0)$. We have:

$$\begin{aligned}
\mathbb{N}^\psi[A_h \geq \theta | A = \theta_0] = \mathbb{N}^\psi[\mathcal{Z}_h^\theta > 0 | A = \theta_0] &= \mathbb{N}^\psi[\mathcal{Z}_h^{A+(\theta-\theta_0)} > 0 | A = \theta_0] \\
&= \psi'(\bar{\theta}_0)\mathbb{N}^\psi\left[\sigma_0 \mathbf{1}_{\{\mathcal{Z}_h^{(\theta-\theta_0)}>0\}} \mathrm{e}^{-\psi(\theta_0)\sigma_0}\right] \\
&= \psi'(\bar{\theta}_0)\mathbb{N}^{\psi_{\theta_0}}\left[\sigma_0 \mathbf{1}_{\{\mathcal{Z}_h^{(\theta-\theta_0)}>0\}}\right] \\
&= \psi'(\bar{\theta}_0)\mathbb{N}^\psi\left[\sigma_{\bar{\theta}_0} \mathbf{1}_{\{\mathcal{Z}_h^{\bar{\theta}_0+(\theta-\theta_0)}>0\}}\right] \\
&= \psi'(\bar{\theta}_0)\mathbb{N}^\psi\left[\sigma_{\bar{\theta}_0} \mathbf{1}_{\{\mathcal{Z}_h^{\hat{\theta}}>0\}}\right],
\end{aligned}$$

where we used (2) of Theorem 3.24 for the third equality, Girsanov formula (3.27) for the fourth and the homogeneity property of Theorem 3.22 in the fifth. We now condition with respect to $\mathcal{T}^{\hat{\theta}}$. The indicator function being measurable, the only quantity left to compute is the conditional expectation of $\sigma_{\bar{\theta}_0}$ given $\mathcal{T}^{\hat{\theta}}$. Thanks to Lemma 3.23, the fact that $\hat{\theta} > 0$ and the homogeneity property, we get:

$$\mathbb{N}^\psi[A_h \geq \theta | A = \theta_0] = \psi'(\hat{\theta})\mathbb{N}^\psi\left[\sigma_{\hat{\theta}} \mathbf{1}_{\{\mathcal{Z}_h^{\hat{\theta}}>0\}}\right] = \psi'(\hat{\theta})\mathbb{N}^{\psi_{\hat{\theta}}}\left[\sigma \mathbf{1}_{\{\mathcal{Z}_h>0\}}\right].$$

Using that $\mathbb{N}^{\psi_{\hat{\theta}}}[\sigma] = 1/\psi'(\hat{\theta})$, which can be deduced from (3.25), we get:

$$\begin{aligned}
\mathbb{N}^\psi[A_h \geq \theta | A = \theta_0] &= \psi'(\hat{\theta})\mathbb{N}^{\psi_{\hat{\theta}}}[\sigma] - \psi'(\hat{\theta})\mathbb{N}^{\psi_{\hat{\theta}}}\left[\int_0^h \mathcal{Z}_a da \mathbf{1}_{\{\mathcal{Z}_h=0\}}\right] \\
&= 1 - \psi'(\hat{\theta})\int_0^h da \lim_{\lambda\to\infty} \mathbb{N}^{\psi_{\hat{\theta}}}\left[\mathcal{Z}_a \mathrm{e}^{-\lambda \mathcal{Z}_h}\right].
\end{aligned}$$

Now, conditioning by $\mathcal{Z}_a$ and using $\lim_{\lambda\to\infty} u^{\hat\theta}(h-t,\lambda) = b_h^{\hat\theta}(t)$ as well as (3.23), we get:

$$\lim_{\lambda\to\infty} \mathbb{N}^{\psi_{\hat\theta}}\left[\mathcal{Z}_a \mathrm{e}^{-\lambda \mathcal{Z}_h}\right] = \lim_{\lambda\to\infty} \mathbb{N}^{\psi_{\hat\theta}}\left[\mathcal{Z}_a \mathrm{e}^{-\mathcal{Z}_a u^{\hat\theta}(h-a,\lambda)}\right] = \mathbb{N}^{\psi_{\hat\theta}}\left[\mathcal{Z}_a \mathrm{e}^{-\mathcal{Z}_a b_h^{\hat\theta}(a)}\right] = \partial_\lambda u^{\hat\theta}(s, b_h^{\hat\theta}(a)).$$

Then use (3.45) to get:

$$\begin{aligned}\int_0^h da\, \lim_{\lambda\to\infty} \mathbb{N}^{\psi_{\hat\theta}}\left[\mathcal{Z}_a \mathrm{e}^{-\lambda \mathcal{Z}_h}\right] = \int_0^h da\, \partial_\lambda u^{\hat\theta}(s, b_h^{\hat\theta}(a)) &= \psi_{\hat\theta}(b^{\hat\theta}(h)) \int_0^h da\, \frac{|\partial_h b^{\hat\theta}(h-a)|}{\psi_{\hat\theta}(b^{\hat\theta}(h-a))^2} \\ &= \psi_{\hat\theta}(b^{\hat\theta}(h)) \int_{b^{\hat\theta}(h)}^{+\infty} \frac{dr}{\psi_{\hat\theta}(r)^2},\end{aligned}$$

and thus deduce the first equality of the Proposition. Notice $\int^{+\infty} dr/\psi_\theta(r)^2 < +\infty$ thanks to Assumption 2 (in fact this is true in general). Let θ go down to θ_0 and use the fact that $\mathbb{N}^\psi$-a.e. $A \le A_h$ to get the second equality. $\square$

Remark 6. *In the quadratic case $\psi(u) = \beta u^2$, we can obtain closed formulæ. For all $\theta > 0$, we have:*

$$u^\theta(t,\lambda) = \frac{2\theta\lambda}{(2\theta+\lambda)\exp(2\beta\theta t) - \lambda} \quad \text{and} \quad b^\theta(t) = \frac{2\theta}{\mathrm{e}^{2\beta\theta t} - 1}.$$

We have the following exact expression of the conditional distribution for $\theta_0 < \theta$, $\theta_0 < 0$ and with $\bar\theta_0 = |\theta_0| = -\theta_0$ and $\hat\theta = \theta + 2|\theta_0|$:

$$\begin{aligned}\mathbb{N}^\psi[A_h \ge \theta | A = \theta_0] &= 1 + (\beta\hat\theta h)/\sinh^2(\beta\hat\theta h) - \mathrm{cotanh}(\beta\hat\theta h), \\ \mathbb{N}^\psi[A_h = A | A = \theta_0] &= \beta\theta_0 h/\sinh^2(\beta\theta_0 h) - \mathrm{cotanh}(\beta\theta_0 h).\end{aligned}$$

Notice that $\lim_{\theta_0\to-\infty} \mathbb{N}^\psi[A_h = A|A = \theta_0] = 1$. This corresponds to the fact that if A is large, then the tree $\mathcal{T}_A$ is small and has little chance to cross level h. (Notice that $\mathcal{T}_A$ has finite height but $\mathcal{T}_{A-}$ has infinite height.) Thus the time A_h is equal to the time when an infinite tree is grafted, that is to the ascension time A.

Distribution of the tree at the exit time

Before stating the theorem describing the tree before it overshoots a given height $h > 0$ under the form of a spinal decomposition, we shall explain how this spine is distributed. Recall (3.46) for the definition of γ_θ.

Lemma 3.32. *Let ψ be a critical branching mechanism satisfying Assumptions 1 and 2. Let $\theta \in \Theta^\psi$. The non-negative function*

$$f : t \mapsto \gamma_\theta(b_h^\theta(t)) \exp\left(-\int_0^t \gamma_\theta(b_h^\theta(r))\, dr\right) \tag{3.52}$$

is a probability density on $[0,h)$ with respect to Lebesgue measure. If ξ is a random variable whose distribution is f, then we have $\mathbb{E}[\exp(-\psi'(\theta)\xi)] < +\infty$.

Notice the integrability property on ξ is trivial if $\theta \ge 0$.

Proof. Notice that $f = g' \mathrm{e}^{-g}$ with $g(t) = \int_0^t \gamma_\theta(b_h^\theta(r))\, dr$. Thus we have

$$\int_0^h f = \int_0^h g' \mathrm{e}^{-g} = \mathrm{e}^{-g(0)} - \mathrm{e}^{-g(h)}$$

and f is a density if and only if $g(h) = \infty$. We deduce from (3.50) that $\int_0^t \gamma_\theta(b_h^\theta(r))dr$ diverges as t goes to h. The last part of Proposition 3.30 implies that $\mathrm{e}^{-\psi'(\theta)\xi}$ is integrable. □

Recall Equation (3.5) defining the grafting procedure.

Theorem 3.33. *Let ψ be a critical branching mechanism satisfying Assumptions 1 and 2. Let $\theta_\infty < \theta$ and let F be a non-negative measurable functional on $\mathbb{T}^2$. Then, we have:*

$$\begin{aligned}\mathbb{N}^\psi &\left[F(\mathscr{T}_{A_h}\,;\,\mathscr{T}_{A_h-})|A_h = \theta\right]\\ &= \frac{1}{\mathbf{E}\left[\mathrm{e}^{-\psi'(\theta)H_\mathbf{x}}\right]}\mathbf{E}\left[F\Big([\![\emptyset,\mathbf{x}]\!] \circledast_{i\in I}(\mathscr{T}^i, x_i)\,;\,([\![\emptyset,\mathbf{x}]\!] \circledast_{i\in I}(\mathscr{T}^i, x_i)) \circledast (T,\mathbf{x})\Big)\,\mathrm{e}^{-\psi'(\theta)H_\mathbf{x}}\right],\end{aligned}$$

where the spine $[\![\emptyset,\mathbf{x}]\!]$ is identified with the interval $[0, H_\mathbf{x}]$ (and thus $y \in [\![\emptyset,\mathbf{x}]\!]$ is identified with H_y) and:

- *The random variable $H_\mathbf{x}$ is distributed with density given by (3.52).*
- *Conditionally on $H_\mathbf{x}$, sub-trees are grafted on the spine $[0, H_\mathbf{x}]$ according to a Poisson point measure $\mathscr{N} = \sum_{i\in I} \delta_{(x_i,\mathscr{T}^i)}$ on $[0, H_\mathbf{x}] \times \mathbb{T}$ with intensity:*

$$\begin{aligned}\nu_\theta(da, d\mathscr{T}) = da\,\Big(&2\beta(\theta + b_h^\theta(a))\mathbb{N}^{\psi_\theta}[d\mathscr{T}, H_{\max}(\mathscr{T}) < h-a]\\ &+ \int_{(0,+\infty)} r\Pi_{\theta+b_h^\theta(x)}(dr)\mathbb{P}_r^{\psi_\theta}(d\mathscr{T}, H_{\max}(\mathscr{T}) < h-a)\Big).\end{aligned} \quad (3.53)$$

- *Conditionally on $H_\mathbf{x}$ and on $\mathscr{N}$, T is a random variable on $\mathbb{T}$ with distribution*

$$\mathbf{N}^{\psi_\theta}[dT|H_{\max}(T) > h - H_\mathbf{x}].$$

In other words, conditionally on $\{A_h = \theta\}$, we can describe the tree before overshooting height h by a spinal decomposition along the ancestral branch of the point at which the overshooting sub-tree is grafted. Conditionally on the height of this point, the overshooting tree has distribution $\mathbf{N}^{\psi_\theta}[dT]$, conditioned on overshooting.

If $\theta > 0$ then $\psi'(\theta) > 0$, and we can understand the weight $\mathrm{e}^{-\psi'(\theta)H_\mathbf{x}}/\mathbf{E}\left[\mathrm{e}^{-\psi'(\theta)H_\mathbf{x}}\right]$ as a conditioning of the random variable $H_\mathbf{x}$ to be larger than an independent exponential random variable with parameter $\psi'(\theta)$.

Remark 7. *When h goes to infinity, we have, for $\theta \geq 0$, $\lim_{h\to+\infty} b^\theta(h) = 0$ and thus the distribution of A_h concentrates on $\Theta^\psi \cap (-\infty, 0)$. For $\theta < 0$ and $\theta \in \Theta^\psi$, we deduce from (3.47) that $\lim_{h\to+\infty} b^\theta(h) = \bar\theta - \theta > 0$. And the distribution of ξ in Lemma 3.32 clearly converges to the exponential distribution with parameter $\gamma_\theta(b^\theta(+\infty)) = \psi'(\bar\theta) - \psi'(\theta)$. Then the weight $\mathrm{e}^{-\psi'(\theta)H_\mathbf{x}}/\mathbf{E}\left[\mathrm{e}^{-\psi'(\theta)H_\mathbf{x}}\right]$ changes this distribution. In the end, $H_\mathbf{x}$ is asymptotically distributed as an exponential random variable with parameter $\psi'(\bar\theta)$. Notice this is exactly the distribution of the height of a random leaf taken in $\mathscr{T}_A$, conditionally on $\{A = \theta\}$, see Lemma 7.6 in [ADH12a].*

Remark 8. *A direct application of Theorem 3.33 with $F(\mathcal{T};\mathcal{T}')$ chosen equal to*

$$G(\mathcal{T};\mathcal{T}') = \mathbf{1}_{\{\mathbf{m}^{\mathcal{T}}(\mathcal{T})<+\infty, \mathbf{m}^{\mathcal{T}'}(\mathcal{T}')=+\infty\}}, \tag{3.54}$$

allows to compute for $\theta < 0$:

$$\mathbb{N}^{\psi}[A = A_h | A_h = \theta] = \left(\psi'(\bar{\theta}) - \psi'(\theta)\right) \frac{C(\theta,h)}{\psi'(\bar{\theta}) - \psi'(\theta) C(\theta,h)},$$

where $C(\theta,h) = \psi'(\bar{\theta})\psi_\theta(b^\theta(h)) \int_{b^\theta(h)}^{+\infty} dr\ \psi_\theta(r)^{-2} = \mathbb{N}^{\psi}[A = A_h | A = \theta]$. The last equality is a consequence of (3.51). *As $\lim_{h\to+\infty} \mathbb{N}^{\psi}[A = A_h | A = \theta] = 1$, we get that*

$$\lim_{h\to+\infty} \mathbb{N}^{\psi}[A = A_h | A_h = \theta] = 1.$$

Remark 9. *By considering the function G in* (3.54) *instead of F in the proof of Theorem 3.33, we can recover the distribution of $\mathcal{T}_A$ given in [ADH12a], but we also can get the joint distribution of $(\mathcal{T}_{A-}, \mathcal{T}_A)$. Roughly speaking (and unsurprisingly), conditionally on $\{A = \theta\}$, $\mathcal{T}_{A-}$ is obtained from $\mathcal{T}_A$ by grafting an independent random tree T on a independent leaf x chosen according to $\mathbf{m}^{\mathcal{T}_A}(dx)$ and the distribution of T is $\mathbf{N}^{\psi_\theta}[dT, H_{\max}(T) = +\infty]$. Notice that choosing a leaf at random on $\mathcal{T}_A$ gives that the distribution of $\mathcal{T}_A$ is a size-biased distribution of $\mathbb{N}^{\psi_\theta}[d\mathcal{T}]$.*

Proof of Theorem 3.33. Thanks to the compensation formula (3.41), we can write, if g is any measurable functional $\mathbf{R} \mapsto \mathbf{R}_+$ with support in $(\theta_\infty, +\infty)$:

$$\begin{aligned}
&\mathbb{N}^{\psi}[F(\mathcal{T}_{A_h}; \mathcal{T}_{A_h-}) g(A_h)] \\
&\qquad = \mathbb{N}^{\psi}\left[\sum_{j\in J} \mathbf{1}_{\{H_{\max}(\mathcal{T}_{\theta_j})<h\}} F(\mathcal{T}_{\theta_j}; \mathcal{T}_{\theta_j} \circledast (\mathcal{T}^j, x_j)) g(\theta_j) \mathbf{1}_{\{H_{x_j} + H_{\max}(\mathcal{T}^j)>h\}}\right] \\
&\qquad = \int_{\Theta^\psi} d\theta\ g(\theta) B(\theta,h),
\end{aligned}$$

where, using the homogeneity property and the Girsanov transformation (3.28):

$$\begin{aligned}
B(\theta,h) &= \mathbb{N}^{\psi}\left[\mathbf{1}_{\{H_{\max}(\mathcal{T}_\theta)<h\}} \int \mathbf{m}^{\mathcal{T}_\theta}(dx) \int \mathbf{N}^{\psi_\theta}[dT] F(\mathcal{T}_\theta; \mathcal{T}_\theta \circledast (T,x)) \mathbf{1}_{\{H_x + H_{\max}(T)>h\}}\right] \\
&= \mathbb{N}^{\psi_\theta}\left[\mathbf{1}_{\{H_{\max}(\mathcal{T})<h\}} \int \mathbf{m}^{\mathcal{T}}(dx) \int \mathbf{N}^{\psi_\theta}[dT] F(\mathcal{T}; \mathcal{T} \circledast (T,x)) \mathbf{1}_{\{H_x + H_{\max}(T)>h\}}\right] \\
&= \mathbb{N}^{\psi_{\bar{\theta}}}\left[\mathbf{1}_{\{H_{\max}(\mathcal{T})<h\}} \int \mathbf{m}^{\mathcal{T}}(dx) \int \mathbf{N}^{\psi_\theta}[dT] F(\mathcal{T}; \mathcal{T} \circledast (T,x)) \mathbf{1}_{\{H_x + H_{\max}(T)>h\}}\right].
\end{aligned}$$

Notice we only replaced $\mathbb{N}^{\psi_\theta}$ by $\mathbb{N}^{\psi_{\bar{\theta}}}$ in the last equality.

We explain how the term $\mathbf{1}_{\{H_{\max}(\mathcal{T})<h\}}$ changes the decomposition of $\mathcal{T}$ according to the spine given in Theorem 3.16. Let Φ a non-negative measurable function defined on $[0,+\infty) \times \mathbb{T}$ and φ a non-negative measurable function defined on $[0,+\infty)$. Using Theorem

3.16 and notations therein, we get:

$$\begin{aligned}
&\mathbb{N}^{\psi_{\bar\theta}}\left[\int \mathbf{m}^{\mathcal{T}}(dx)\ \varphi(H_x)\mathrm{e}^{-\langle \mathcal{M}_x,\Phi\rangle}\mathbf{1}_{\{H_{\max}(\mathcal{T})<h\}}\right]\\
&\qquad\qquad = \int_0^\infty da\, \varphi(a)\mathrm{e}^{-\psi'_{\bar\theta}(0)a}\,\mathbb{E}\left[\mathrm{e}^{-\sum_{i\in I}\mathbf{1}_{\{z_i\le a\}}\Phi(z_i,\tilde{\mathcal{T}}^i)}\prod_{i\in I, z_i\le a}\mathbf{1}_{\{H_{\max}(\tilde{\mathcal{T}}^i)<h-z_i\}}\right]\\
&\qquad\qquad = \int_0^h da\, \varphi(a)\ \exp\left(-\psi'(\bar\theta)a-\int_0^a dx\,\mathbf{N}^{\psi_{\bar\theta}}\left[1-\mathrm{e}^{-\Phi(x,\mathcal{T})}\mathbf{1}_{\{H_{\max}(\mathcal{T})<h-x\}}\right]\right).
\end{aligned}$$

Using the definition of $\mathbf{N}^{\psi_{\bar\theta}}$, see (3.40), (3.46) and the Girsanov transformation (3.28), we get:

$$\begin{aligned}
\mathbf{N}^{\psi_{\bar\theta}}\left[1-\mathrm{e}^{-\Phi(x,\mathcal{T})}\mathbf{1}_{\{H_{\max}(\mathcal{T})<h-x\}}\right] &= \gamma_{\bar\theta}\left(\mathbb{N}^{\psi_{\bar\theta}}\left[1-\mathrm{e}^{-\Phi(x,\mathcal{T})}\mathbf{1}_{\{H_{\max}(\mathcal{T})<h-x\}}\right]\right)\\
&= \gamma_{\bar\theta}\left(b^{\bar\theta}(h-x)+\mathbb{N}^{\psi_\theta}\left[\left(1-\mathrm{e}^{-\Phi(x,\mathcal{T})}\right)\mathbf{1}_{\{H_{\max}(\mathcal{T})<h-x\}}\right]\right).
\end{aligned}$$

Thanks to (3.46) and (3.47), we have for $\lambda\ge 0$:

$$\gamma_{\bar\theta}(b^{\bar\theta}(h-x)+\lambda)=\gamma_{\theta+b^\theta(h-x)}(\lambda)+\gamma_\theta(b^\theta(h-x))+\psi'(\theta)-\psi'(\bar\theta).$$

Take $\lambda=\mathbb{N}^{\psi_\theta}\left[\left(1-\mathrm{e}^{-\Phi(x,\mathcal{T})}\right)\mathbf{1}_{\{H_{\max}(\mathcal{T})<h-x\}}\right]$, to deduce that:

$$\begin{aligned}
&\mathbb{N}^{\psi_{\bar\theta}}\left[\int \mathbf{m}^{\mathcal{T}}(dx)\ \varphi(H_x)\mathrm{e}^{-\langle \mathcal{M}_x,\Phi\rangle}\mathbf{1}_{\{H_{\max}(\mathcal{T})<h\}}\right]\\
&\qquad = \int_0^h da\, \varphi(a)\ \exp\left(-\psi'(\theta)a-\int_0^a dx\,\gamma_\theta(b^\theta(h-x))\right)\\
&\qquad\qquad\qquad \exp\left(-\int_0^a dx\,\gamma_{\theta+b^\theta(h-x)}\left(\mathbb{N}^{\psi_\theta}\left[\left(1-\mathrm{e}^{-\Phi(x,\mathcal{T})}\right)\mathbf{1}_{\{H_{\max}(\mathcal{T})<h-x\}}\right]\right)\right)\\
&\qquad = \int_0^h da\, \varphi(a)\ \exp\left(-\psi'(\theta)a-\int_0^a dx\,\gamma_\theta(b^\theta(h-x))\right)\mathbb{E}\left[\mathrm{e}^{-\sum_{i\in I}\mathbf{1}_{\{z_i\le a\}}\Phi(z_i,\tilde{\mathcal{T}}^i)}\right],
\end{aligned}$$

where under $\mathbb{E}$, $\sum_{i\in I}\delta_{(z_i,\tilde{\mathcal{T}}^i)}(dz,d\mathcal{T})$ is a Poisson point measure on $[0,h]\times\mathbb{T}$ with intensity ν_θ in (3.53). Since Laplace transforms characterize random measure distributions, we get that for any non-negative measurable function $\tilde F$, we have:

$$\begin{aligned}
&\mathbb{N}^{\psi_{\bar\theta}}\left[\int \mathbf{m}^{\mathcal{T}}(dx)\tilde F(H_x,\mathcal{M}_x)\mathbf{1}_{\{H_{\max}(\mathcal{T})<h\}}\right]\\
&\qquad\qquad = \int_0^h da\,\mathrm{e}^{-\psi'(\theta)a-\int_0^a dx\,\gamma_\theta(b^\theta(h-x))}\,\mathbb{E}\left[\tilde F\left(a,\sum_{i\in I}\mathbf{1}_{\{z_i\le a\}}\delta_{(z_i,\tilde{\mathcal{T}}^i)}\right)\right].
\end{aligned}$$

If we identify the spine $[\![\emptyset,x]\!]$ (with its metric) to the interval $[0,H_x]$ (with the Euclidean metric), we can use this result to compute $B(\theta,h)$ with:

$$\tilde F(H_x,\mathcal{M}_x)=\int \mathbf{N}^{\psi_\theta}[dT\mid H_x+H_{\max}(T)>h]F(\mathcal{T}\,;\mathcal{T}\circledast(T,x)),$$

$\mathcal{M}_x = \sum_{i\in I_x} \delta_{(H_{x_i},\mathcal{T}^i)}$ and $\mathcal{T} = [0, H_x] \circledast_{i\in I_x} (\mathcal{T}^i, H_{x_i})$. Since $\mathbf{N}^{\psi_\theta}[H_{\max}(\mathcal{T}) > h] = \gamma_\theta(b^\theta(h))$, we have:

$$\gamma_\theta(b^\theta(h - H_x))\tilde{F}(H_x, \mathcal{M}_x) = \int \mathbf{N}^{\psi_\theta}[dT] F(\mathcal{T} \,;\, \mathcal{T} \circledast (T, x))\mathbf{1}_{\{H_x + H_{\max}(T) > h\}}.$$

Therefore, we have:

$$\begin{aligned} B(\theta, h) &= \mathbb{N}^{\psi_{\bar\theta}}\left[\mathbf{1}_{\{H_{\max}(\mathcal{T})<h\}} \int \mathbf{m}^{\mathcal{T}}(dx) \int \mathbf{N}^{\psi_\theta}[dT] F(\mathcal{T} \,;\, \mathcal{T} \circledast (T, x))\mathbf{1}_{\{H_x + H_{\max}(T) > h\}}\right] \\ &= \int_0^h da\, \gamma_\theta(b^\theta(h-a)) \mathrm{e}^{-\psi'(\theta)a - \int_0^a dx\, \gamma_\theta(b^\theta(h-x))} \mathbb{E}\left[\tilde{F}\left(a, \sum_{i\in I} \mathbf{1}_{\{z_i \le a\}} \delta_{(z_i, \tilde{\mathcal{T}}^i)}\right)\right]. \end{aligned}$$

Thus, we get:

$$\begin{aligned} &\mathbb{N}^{\psi}[F(\mathcal{T}_{A_h} \,;\, \mathcal{T}_{A_h -}) g(A_h)] \\ &\qquad = \int_{\Theta^\psi} d\theta\; g(\theta) \int_0^h da\, \gamma_\theta(b^\theta(h-a)) \mathrm{e}^{-\psi'(\theta)a - \int_0^a dx\, \gamma_\theta(b^\theta(h-x))} \\ &\qquad\qquad \mathbb{E}\left[\tilde{F}\left(a, \sum_{i\in I} \mathbf{1}_{\{z_i \le a\}} \delta_{(z_i, \tilde{\mathcal{T}}^i)}\right)\right]. \end{aligned}$$

Then use the distribution of A_h under $\mathbb{N}^\psi$ given in Proposition 3.30 to conclude. □

CHAPTER 4

Fluctuations for the number of records on CRT subtrees

Introduction

The Continuum Random Tree (CRT) is a random metric measure space, introduced by Aldous ([Ald91a, Ald93]) as a scaling limit of various discrete random tree models. In particular, if we consider μ, a critical probability measure on $\mathbb{N}$, with variance $0 < \sigma^2 < \infty$ and if we consider a random Galton-Watson tree $\mathscr{T}_n$ with offspring distribution μ, conditioned on having n vertices, then we have the following convergence in distribution:

$$\lim_{n\to\infty} \frac{\sigma}{\sqrt{n}} \mathscr{T}_n = \mathscr{T}, \tag{4.1}$$

in the sense of Gromov-Hausdorff convergence of compact metric spaces (see for instance [DL05] for more information about the Gromov-Hausdorff topology), where $\mathscr{T}$ is a CRT. The family of conditioned Galton-Watson trees turns out to be quite large, since it contains for instance uniform rooted planar binary trees (take $\mu(0) = \mu(2) = 1/2$) or uniform rooted labelled trees (Cayley trees, take $\mu(k) = e^1/k!,\ k \geq 0$). There is a combinatorial characterization of conditioned Galton-Watson trees: they correspond to the class of so-called *simply generated trees* (see [Jan12] for a detailed survey).

In their 1970 paper ([MM70]), Meir and Moon considered the problem of isolating the root through uniform cuts in random Cayley trees. The problem is as follows: start with a rooted discrete tree $\mathscr{T}_n$, having exactly n edges (in our context, *rooted* means that, among the $n+1$ vertices of $\mathscr{T}_n$, one has been distinguished). At each step, remove an edge, selected uniformly among all edges, then discard the connected component not containing the root. This procedure is iterated on the remaining tree until the root is the only remaining vertex. The number $X(\mathscr{T}_n)$ of cuts that is needed to isolate the root is random, with values in $[\![1, n]\!]$. Meir and Moon showed that when $\mathscr{T}_n$ is a uniform Cayley tree with n edges,

$$\mathbb{E}[X(\mathscr{T}_n)] \sim \sqrt{\pi n/2} \quad \text{and} \quad \mathrm{Var}(X(\mathscr{T}_n)) \sim (2 - 1/\pi)n.$$

Later, the limiting distribution was found to be the Rayleigh distribution (the distribution on $[0,\infty)$ with density $x\exp(-x^2/2)dx$) by Panholzer for (a subset of) the class of simply generated trees ([Pan06]) and, using a different proof, by Janson for the class of critical, finite-variance, conditioned Galton-Watson trees ([Jan06]).

In [Jan06], the distribution of the limiting Rayleigh variable was obtained using a moment problem, but the question arose whether it had a connection with the convergence (4.1) above. Indeed, it is well-known that the distance from the root to a uniform leaf of the CRT is Rayleigh-distributed. As a consequence, several approaches were used to describe a cutting procedure on the CRT that could account for the convergence of $X(\mathcal{T}_n)/\sqrt{n}$. All these works are relying on the Aldous-Pitman fragmentation of the CRT, first described in ([AP98a]). We will give below a brief descriptions of this procedure, as it will be central in this work. Using an extension of the Aldous-Broder algorithm, Addario-Berry, Broutin and Holmgren described a fragmentation-reconstruction procedure for Cayley trees and its analog for the CRT. The invariance they prove shows that the limiting random variable in Janson's result can indeed be realized as the height of a uniform leaf in a CRT. However, it is not the same CRT as the one arising from the scaling limit of $\mathcal{T}_n/\sqrt{n}$. Indeed, the random variables $n^{-1/2}\mathcal{T}_n$ and $n^{-1/2}X(\mathcal{T}_n)$ do *not* converge jointly to a CRT $\mathcal{T}$ and the height of a random leaf $H(\mathcal{T})$. Bertoin and Miermont ([BM12]) describe the so-called *cut-tree* $\mathrm{cut}(\mathcal{T})$ of a given CRT $\mathcal{T}$ following the genealogy of fragments in the Aldous-Pitman fragmentation. The limiting variable can then be described as the height of a uniform leaf in $\mathrm{cut}(\mathcal{T})$, which is again a CRT, thus recovering Rayleigh distribution.

Following Abraham and Delmas ([AD11]), we shall use a different point of view, based on the theory of records of Poisson point processes. We shall now review some of their results, in order to set the notations and to describe the framework.

The Brownian CRT

In this section, we shall recall some basic facts about the Brownian CRT. For details, see [Ald91a, DL05, Eva08]. We will write $\mathbb{T}$ for the set of (pointed isometry classes of) compact, rooted real trees endowed with a finite Borel measure. Recall that real trees are metric spaces (X,d) such that

(i) For every $s,t\in X$, there is a unique isometric map $f_{s,t}$ from $[0,d(s,t)]$ to X such that $f_{s,t}(0)=s$ and $f_{s,t}(d(s,t))=t$. The image of $f_{s,t}$ is noted $[\![s,t]\!]$.

(ii) For every $s,t\in X$, if q is a continuous injective map from $[0,1]$ to X such that $q(0)=s$ and $q(1)=t$, then $q([0,1])=f_{s,t}([0,d(s,t)])$.

There exists a metric on $\mathbb{T}$ that makes it a Polish metric space, but we will not attempt to describe it here. For more details, see [ADH12b, Eva08].

The Brownian CRT (or Aldous's CRT) is a random element of $\mathbb{T}$, defined using the so-called *contour process* description: if f is a continuous nonnegative map $f:[0,\sigma]\to\mathbf{R}_+$, such that $f(0)=f(\sigma)=0$, then the real tree encoded by f is defined by $\mathcal{T}_f=[0,\sigma]_{/\sim_f}$, where $\sim_f$ is the equivalence relation

$$x\sim_f y \;\Leftrightarrow\; f(x)=f(y)=\min_{u\in[x\wedge y,x\vee y]} f(u),\qquad x,y\in[0,\sigma].$$

The metric on $\mathcal{T}_f$ is defined by

$$d_f(x,y) = f(x) + f(y) - 2 \min_{u \in [x\wedge y, x\vee y]} f(u), \qquad x,y \in [0,\sigma],$$

so that $d_f(x,y) = 0$ if and only if $x \sim_f y$. Hence, d_f is definite-positive on $\mathcal{T}_f$ and defines a true metric. It can be checked (see [DL05]) that $(\mathcal{T}_f, d_f)$ is indeed a real tree. We define the *mass-measure* $\mathbf{m}^{\mathcal{T}_f}(ds)$ on $\mathcal{T}_f$ as the image of Lebesgue measure on $[0,\sigma]$ by the canonical projection $[0,\sigma] \to \mathcal{T}_f$. Thus, $\mathbf{m}^{\mathcal{T}_f}$ is a finite measure on $\mathcal{T}_f$, with total mass $\mathbf{m}^{\mathcal{T}_f}(\mathcal{T}_f) = \sigma$. When the context is clear, we will usually drop the reference to the tree and write $\mathbf{m}$ for the mass-measure $\mathbf{m}^{\mathcal{T}}$.

Now, the Brownian Continuum Random Tree (CRT) corresponds to the real tree encoded by $f = 2B^{\mathrm{ex}}$, twice the normalized Brownian excursion. Since the length of the normalized Brownian excursion is 1 a.s., the CRT has total mass 1, *i.e.* the mass measure $\mathbf{m}$ is a probability measure. The distribution of the CRT will be noted $\mathbb{P}$, or sometimes $\mathbb{P}^{(1)}$ if we want to emphasize the fact that $\mathbf{m}$ has mass 1. Sometimes, we will consider *scaled* versions of the CRT. If $r > 0$, we consider the scaled Brownian excursion

$$B_t^{\mathrm{ex},r} = \sqrt{r} B_{t/r}^{\mathrm{ex}}, \ \ t \in [0,r]$$

and the associated real tree $\mathcal{T}_{2B^{\mathrm{ex},r}}$, whose distribution will be noted $\mathbb{P}^{(r)}$. Note that the transformation above corresponds to rescaling all the distances in a $\mathbb{P}^{(1)}$-distributed tree by a factor $\sqrt{r}$.

The measure $\mathbf{m}$ is supported by the set of *leaves* of $\mathcal{T}$, which are the points $\mathrm{x} \in \mathcal{T}$ such that $\mathcal{T} \setminus \{\mathrm{x}\}$ is connected. There is another natural measure ℓ defined on the CRT, called *length measure*, which is σ-finite and such that $\ell([\![x,y]\!]) = d(x,y)$. Also, the CRT is rooted at one particular vertex $\emptyset$, which is the equivalence class of 0, but it can be shown (see Proposition 4.8 in [DL05]) that if x is chosen according to $\mathbf{m}$, then, if $\mathcal{T}^{\mathrm{x}}$ is the tree $\mathcal{T}$ re-rooted at x, $(\mathcal{T}, \mathrm{x})$ has same distribution as $(\mathcal{T}^{\mathrm{x}}, \emptyset)$.

When we consider the real tree $\mathcal{T}$ encoded by $2B$, where B is an *excursion* of Brownian motion, distributed under the (σ-finite) excursion measure $\mathbb{N}$, we get that $\mathcal{T}$ is a compact metric space, with a length measure ℓ and with a finite measure $\mathbf{m}$. We will write σ for the (random) total mass of $\mathbf{m}$. Under $\mathbb{N}$, σ is distributed as the length of a random excursion of Brownian Motion, that is

$$\mathbb{N}[\sigma \geq t] = \sqrt{\frac{2}{\pi t}}.$$

The Brownian CRT can be seen as a conditioned version of the tree distributed as $\mathbb{N}[d\mathcal{T}]$, in the sense that, if F is some nonnegative measurable functional defined on the tree space $\mathbb{T}$, then

$$\mathbb{N}[F(\mathcal{T})] = \int_0^\infty \frac{d\sigma}{\sqrt{2\pi}\ \sigma^{3/2}} \mathbb{E}^{(\sigma)}[F(\mathcal{T})].$$

In the sequel, we shall make use of this disintegration of $\mathbb{N}$, since some computations are easier to do under $\mathbb{N}$ (see Proposition 4.5).

The Aldous-Pitman fragmentation

Given a CRT $\mathcal{T}$, we consider a Poisson point process

$$\mathcal{N}(ds, dt) = \sum_{i \in I} \delta_{(s_i, t_i)}(ds, dt)$$

on $\mathcal{T} \times \mathbf{R}_+$, with intensity $\ell(ds) \otimes dt$. We will sometimes refer to $\mathcal{N}$ as the *fragmentation measure*. If (s_i, t_i) is an atom of $\mathcal{N}$, we will say that the point s_i was *marked* at time t_i. For $t \geq 0$, we can consider the connected components of $\mathcal{T}$ separated by the atoms of $\mathcal{N}(\cdot \times [0, t])$. They define a random forest $\mathcal{F}_t$ of subtrees of $\mathcal{T}$. Aldous and Pitman proved that if we consider the trees $(\mathcal{T}_k(t), k \geq 1)$ composing $\mathcal{F}_t$, ranked by decreasing order of their mass, then the process

$$((\mathbf{m}(\mathcal{T}_1(t)), \mathbf{m}(\mathcal{T}_2(t)), \ldots), \ t \geq 0)$$

is a binary, self-similar fragmentation process, with index 1/2 and erosion coefficient 0, according to the terminology later framed by Bertoin ([Ber02]).

Separation times

In order to give a continuous analogue to the cutting procedure on discrete trees described above, we will use the Aldous-Pitman fragmentation on the CRT. Given a CRT $\mathcal{T}$ and a fragmentation measure $\mathcal{N}$, we will define, for any $s \in \mathcal{T}$, the ***separation time*** from the root $\emptyset$ by

$$\theta(s) = \inf \{t \geq 0, \ \mathcal{N}([\![\emptyset, s]\!] \times [0, t]) \geq 1\},$$

with the convention $\inf \emptyset = +\infty$. This separation process will be our main object of study. Note that, under the definition above, conditionally on $\mathcal{T}$, $\theta(\emptyset) = \infty$ a.s., and $\theta(s) < \infty$ a.s. for all $s \neq \emptyset$, since $\theta(s)$ is then exponentially distributed with parameter $\ell([\![\emptyset, s]\!]) = d(\emptyset, s)$. Note also that $\theta(s) \to \infty$ when $s \to \emptyset$, which justifies our convention for $\theta(\emptyset)$.

It is also possible to define the separation process started from any $q \geq 0$, rather than from infinity. In order to do this, we consider only the marks whose t-component is smaller than q:

$$\theta(s) = \inf \{0 \leq t \leq q, \ \mathcal{N}([\![\emptyset, s]\!] \times [0, t]) \geq 1\}, \tag{4.2}$$

with the convention $\inf \emptyset = q$. Note that, under this definition, we always have $\theta(\emptyset) = q$, as well as $\lim \theta(s)_{s \to \emptyset} = q$ a.s. In the case where $q = \infty$, we recover the same distribution as the separation process defined earlier. The (quenched) distribution of the separation process started at $q \in [0, \infty]$ on a given CRT $\mathcal{T}$ will be noted $\mathbb{P}_q^{\mathcal{T}}$.

We will also note $\mathbb{P}_q^{(r)}$ the (annealed) distribution of the process $(\theta(s), s \in \mathcal{T})$ started at $q \in [0, \infty]$, when $\mathcal{T}$ is distributed as a Brownian CRT with mass $r > 0$:

$$\mathbb{P}_q^{(r)} = \int_{\mathbb{T}} \mathbb{P}^{(r)}(d\mathcal{T}) \ \mathbb{P}_q^{\mathcal{T}}.$$

Again, to keep things simple, we will usually work under $\mathbb{P}_\infty = \mathbb{P}_\infty^{(1)}$. The jump points of the separation process correspond to points s marked by the fragmentation measure at a time t

where they belong to the connected component of the root. This implies that they accumulate in the neighbourhood of the root if $q=\infty$. If T is a subtree of $\mathscr{T}$, we note $X(T)$ the number of jumps of the separation process on T. This number can be finite or infinite, according to whether T contains the root or not, in the case $q=\infty$.

Linear record process

One can consider the record process on the real line (*i.e.* when $\mathscr{T}=\mathbf{R}_+$), defined using a Poisson point measure with intensity $ds\otimes dt$. We get, for any $q\in(0,\infty]$, a random process $(\theta(s),\ s\geq 1)$ such that $\theta(0)=q$, $\mathbb{P}_q^{\mathbf{R}_+}$-a.s. The distribution of this process will be noted $\mathbf{P}_q=\mathbb{P}_q^{\mathbf{R}_+}$. We can consider the jump process

$$X_t=\sum_{s\in[0,t]}\mathbf{1}_{\{\theta(s-)>\theta(s)\}},$$

counting the number of jumps of θ on $[0,t]$. It should be noted that if $q=\infty$, then θ jumps infinitely often in the neighbourhood of the root, so that a.s. $X_t=\infty$ for any $t>0$. It is easy to check that, for any bounded, measurable functional g defined on $[0,q]$, we have

$$\mathbf{E}_q\left[g(\theta(s))\right]=\mathrm{e}^{-qs}g(q)+\int_0^q g(x)s\mathrm{e}^{-sx}\ dx.$$

In particular,

$$\mathbf{E}_q\left[\theta(s)\right]=\frac{1-\mathrm{e}^{-qs}}{s}. \tag{4.3}$$

When $q<\infty$, if $t\geq 0$, and conditionally on $\theta(t)=q'$, the next jump of θ can be seen to be equal to $\inf\{s\geq t,\ \mathcal{N}([0,q'],[t,s])\geq 1\}$, which is exponentially distributed, with parameter q'. Thus, X is the counting process of a point measure on $\mathbf{R}_+$ with intensity $\theta(s)ds$. Elementary properties of counting processes of point measures (see [AD11] for more details) then show that, for any $q\in(0,\infty)$, the processes

$$\left(N_t=X_t-\int_0^t\theta(s)\ ds,\ t\geq 0\right) \tag{4.4}$$

$$\left(N_t^2-\int_0^t\theta(s)\ ds,\ t\geq 0\right) \tag{4.5}$$

$$\left(N_t^4-3\left(\int_0^t\theta(s)\ ds\right)^2-\int_0^t\theta(s)\ ds,\ t\geq 0\right) \tag{4.6}$$

are $\mathbf{P}_q$-martingales in the natural filtration of θ.

Number of records on subtrees

Given a CRT $\mathscr{T}$, let $(\mathrm{x}_n,\ n\geq 1)$ be an iid sequence of leaves of $\mathscr{T}$, sampled according to $\mathbf{m}$. If $n\geq 1$, we consider T_n, the subtree spanned by the leaves $(\varnothing,\mathrm{x}_1,\dots,\mathrm{x}_n)$. The tree T_n is a random rooted binary tree with edge-lengths, whose distribution is explicitly known (see [Ald93]). Its length $L_n = \ell(\mathrm{T}_n)$ is known to be distributed according to the Chi$(2n)$-distribution, that is

$$\mathbb{P}(L_n \in dx) = \frac{2^{1-n}}{(n-1)!} x^{2n-1} \exp(-x^2/2) \mathbf{1}_{\{x>0\}}. \tag{4.7}$$

Note that the case $n = 1$ gives a Rayleigh distribution, as was mentioned earlier. It is proven in [AD11] that, a.s.:

$$\lim_{n\to\infty} \frac{L_n}{\sqrt{2n}} = 1. \tag{4.8}$$

The tree T_n has exactly $2n-1$ edges. The edge adjacent to the root will be noted $[\![\emptyset, s_{\emptyset,n}]\!]$, where $s_{\emptyset,n}$ is the first branching point in T_n; the height of $s_{\emptyset,n}$ is noted $h_{\emptyset,n} = \ell([\![\emptyset, s_{\emptyset,n}]\!])$. Recall from Proposition 5.3 in [AD11] that $\sqrt{n} h_{\emptyset,n}$ converges in distribution to a nondegenerate random variable, and that we have the following moment computation, for $\alpha > -1$:

$$\mathbb{E}\left[h_{\emptyset,n}^{\alpha}\right] = \frac{\Gamma(\alpha+1)}{2^{\alpha/2}} \frac{\Gamma(n-1/2)}{\Gamma(n+\alpha/2-1/2)} \sim_{n\to\infty} \Gamma(\alpha+1) 2^{-\alpha/2} n^{-\alpha/2}. \tag{4.9}$$

We will also use the notation $\mathrm{T}_n^* = (\mathrm{T}_n \setminus [\![\emptyset, s_{\emptyset,n}]\!]) \cup \{s_{\emptyset,n}\}$ for the subtree above the lowest branching point in T_n. When a new leaf x_n is sampled, it gets attached to the tree T_{n-1} through a new edge, that connects to T_{n-1} at the vertex $s_n \in \mathrm{T}_{n-1}$. We write

$$\mathrm{B}_n = (\mathrm{T}_n \setminus \mathrm{T}_{n-1}) \cup \{s_n\} = [\![s_n, \mathrm{x}_n]\!].$$

The quantity X_n^* is the continuum counterpart of the edge-cutting number $X(T_n)$ that can be found in the literature. Indeed, as soon as a jump appears on the first edge $[\![\emptyset, s_{\emptyset,n}]\!]$, all subsequent jumps will be on this edge, even closer to the root. Thus, X_n^* can be seen as the number of cuts before the first cut on $[\![\emptyset, s_{\emptyset,n}]\!]$ was made. In some sense, the first mark appearing on $[\![\emptyset, s_{\emptyset,n}]\!]$ is analog to the last cut needed to isolate the root in the discrete case.

The following theorem is the analog of the convergence (in distribution) that can be found in [Jan06] $X(\mathcal{T}_n)/\sqrt{n} \to \mathcal{R}$, where $\mathcal{R}$ is Rayleigh-distributed. We will write Θ for the mean separation time $\int_{\mathcal{T}} \theta(ds)\mathbf{m}(ds)$.

Theorem ([AD11]). *We have $\mathbb{P}_\infty$-a.s:*

$$\lim_{n\to\infty} \frac{X_n^*}{\sqrt{2n}} = \Theta. \tag{4.10}$$

Furthermore, under $\mathbb{P}_\infty$, Θ has Rayleigh distribution.

Note that T_n^* has $2n-2$ edges, so that the rescaling is $\sqrt{2n}$. In comparison, Janson considers random trees with n edges, which explains the difference between the two results. It should be noted that Abraham and Delmas show a slightly more general result, since they consider scaled versions of the CRT, proving the result under all the measures $\mathbb{P}_\infty^{(r)}$, $r > 0$. While our main result, Theorem 4.1 below is still true in these cases, we restrict ourselves to the case of Aldous's tree ($r = 1$) for convenience.

The purpose of this work is to investigate the fluctuations of $X_n^*/\sqrt{2n}$ around its limit Θ. It is shown in Theorem 4.1, which is the main result of this work, that these fluctuations are typically of the order $n^{1/4}$.

Theorem 4.1. *Under $\mathbb{P}_\infty$, we have the following convergence in distribution:*

$$\lim_{n\to\infty} n^{1/4}\left(\frac{X_n^*}{\sqrt{2n}} - \Theta\right) = Z, \tag{4.11}$$

where Z is a random variable which is, conditionally on Θ, distributed according to

$$\mathbb{E}_\infty^{(1)}\left[e^{itZ}\middle|\Theta\right] = e^{-t^2\Theta/\sqrt{2}}. \tag{4.12}$$

In other words, Z is distributed as $2^{1/4}\sqrt{\Theta}G$, where G is an independent standard normal random variable. As Θ is Rayleigh-distributed under $\mathbb{E}_\infty^{(1)}$, the Laplace transform (4.12) can be explicitly computed, but does not correspond to any known distribution.

The proof of Theorem 4.1 will be carried out in two steps: we write

$$\left(\frac{X_n^*}{\sqrt{2n}} - \Theta\right) = \frac{1}{\sqrt{2n}}\left(X_n^* - \int_{\mathrm{T}_n^*}\theta(s)\ell(ds)\right) + \left(\frac{1}{\sqrt{2n}}\int_{\mathrm{T}_n^*}\theta(s)\ell(ds) - \Theta\right). \tag{4.13}$$

In Section 1, we will show that, when averaging over $\mathscr{T}$, the variance arising from the random choice of the leaves $(\mathrm{x}_n,\ n\geq 1)$ does not bring any significant contribution to (4.11). We prove this by decomposing $\mathscr{T}$ conditionally on its subtree T_n and by proving a general disintegration formula (Lemma 4.5). Therefore, the second term in (4.13) converges to 0 when suitably renormalized.

In Section 2, we prove Theorem 4.1 by showing that, when properly rescaled, the difference $(X_n^* - \int_{\mathrm{T}_n^*}\theta(s)\mathbf{m}(ds))$ is asymptotically normally distributed (Proposition 4.4). This is a consequence of the classical martingale convergence theorems of [HH80].

In the Appendix, we collect several technical lemmas.

4.1 Variance in the weak convergence of length measure to mass measure

The main result of this section is Proposition 4.2.

Proposition 4.2. *As $n\to\infty$, we have the following convergence in probability:*

$$\lim_{n\to\infty} n^{1/4}\left(\int_{\mathrm{T}_n^*}\theta(s)\frac{\ell(ds)}{\sqrt{2n}} - \Theta\right) = 0. \tag{4.14}$$

Recall that, conditionally on $\mathscr{T}$, we sample independent leaves $(\mathrm{x}_n,\ n\geq 1)$ with common distribution $\mathbf{m}(d\mathrm{x})$. We will consider the filtration $(\mathscr{F}_n,\ n\geq 1)$ defined by

$$\mathscr{F}_n = \sigma\left(\{(\mathrm{T}_1,\ldots,\mathrm{T}_n),\ (\theta(s),\ s\in\mathrm{T}_n)\}\right),\quad n\geq 1.$$

A key step in the proof of the a.s. convergence of $X_n^*/\sqrt{2n}$ to Θ in [AD11] is the convergence of $M_n = \mathbb{E}_\infty^{(1)}[\Theta|\mathscr{F}_n]$. Since $(M_n,\ n\geq 1)$ is a closed L^2 martingale, it converges $\mathbb{P}_\infty^{(1)}$-a.s. (and in L^2) towards $M_\infty = \Theta$ (notice that Θ is indeed $\mathscr{F}_\infty$-measurable, since $\cup_{n\geq 1}\mathrm{T}_n$ is dense in $\mathscr{T}$, and since θ is continuous $\mathbf{m}$-almost everywhere). The proof of Proposition 4.2 will be divided in two. First, we prove the next proposition:

Proposition 4.3. *We have the following convergence in probability:*

$$\lim_{n\to\infty} n^{1/4}\left(\frac{1}{\sqrt{2n}}\int_{\mathrm{T}_n^*}\theta(s)\ell(ds) - \mathbb{E}_\infty^{(1)}[\Theta|\mathscr{F}_n]\right) = 0. \tag{4.15}$$

Then, we prove a more precise statement than the convergence of $\mathbb{E}_\infty^{(1)}[\Theta|\mathscr{F}_n]$ towards Θ.

Proposition 4.4. *We have*

$$\lim_{n\to\infty} n^{1/4}\left(\mathbb{E}_\infty^{(1)}[\Theta|\mathscr{F}_n] - \Theta\right) = 0, \tag{4.16}$$

in probability, as $n\to\infty$.

Of course, Propositions 4.3 and 4.4 imply Proposition 4.2. Before we can prove Proposition 4.3, we need to describe more precisely how the marked tree $(\mathscr{T},\theta)$ is distributed conditionally on $\mathscr{F}_n$.

Subtree decomposition

Given the subtree T_n, the set $\mathscr{T}\setminus\mathrm{T}_n$ is a random forest; let $(\mathscr{X}_i,\ i\in I_n)$ be the collection of its connected components. For any connected component $\mathscr{X}_i$ of $\mathscr{T}\setminus\mathrm{T}_n$, there is a unique point $s_i\in\mathrm{T}_n$ such that

$$\bigcap_{x\in\mathscr{X}_i}[\![\emptyset,x]\!] = [\![\emptyset,s_i]\!].$$

For any $i\in I_n$, we will write $\mathscr{T}_i$ for the tree $\mathscr{X}_i\cup\{s_i\}$, rooted at $s_i\in\mathrm{T}_n$. We will sometimes use the notation

$$\begin{aligned}\mathscr{E}_n &= \{s\in\mathscr{T},\ [\![\emptyset,s]\!]\cap\mathrm{T}_n^* = \emptyset\} \\ &= \bigcup_{i\in I_n,\ s_i\in[\![\emptyset,s_{\emptyset,n}]\!]}\mathscr{X}_i\end{aligned} \tag{4.17}$$

for the set of all vertices in the tree such that the unique path linking them to the root intersects T_n on $[\![\emptyset,s_{\emptyset,n}]\!]$.

Many things are known about the distribution of the forest $(\mathscr{T}_i,\ i\in I_n)$. For instance, Pitman pointed out (see [DGM06]) that the stickbreaking construction of the CRT in [Ald91a] implied that the sequence $(\mathbf{m}(\mathscr{T}_i),\ i\in I_n)$, ranked in decreasing order, is distributed according to the Poisson-Dirichlet distribution with parameters $\alpha = 1/2$ and $\theta = n-1/2$ (for more background on Poisson-Dirichlet distributions, see [Pit06]). We will give another description, focusing on the tree structure of $\mathscr{T}$ conditionally on $\mathscr{F}_n$. This description can be seen as a conditioned version of Theorem 3 in [Le 93b].

Lemma 4.5. *Let F be a nonnegative functional on $\mathbb{T}\times\mathrm{T}_n$. Then*

$$\mathbb{E}_\infty^{(1)}\left[\sum_{i\in I_n}F(\mathscr{T}_i,s_i)\middle|\mathscr{F}_n\right] = \int_0^1\frac{\mathrm{e}^{-L_n^2 v/(2-2v)}}{\sqrt{2\pi}v^{3/2}(1-v)^{3/2}}\,dv\int_{\mathrm{T}_n}\ell(ds)\mathbb{E}_{\theta(s)}^{(v)}[F(\mathscr{T},s)]. \tag{4.18}$$

Proof. Let Y be a $\mathscr{F}_n$-measurable random variable; let us compute $\mathbb{E}_\infty^{(1)}\left[Y\sum_{i\in I_n}F(\mathscr{T}_i,s_i)\right]$. In order to do this computation, we will perform a disintegration with respect to σ in the following expression: for $\mu\geq 0$,

$$\begin{aligned} I(\mu) &= \mathbb{N}_\infty\left[Y\sum_{i\in I_n}F(\mathscr{T}_i,s_i)\mathrm{e}^{-\mu\sigma}\right] \\ &= \mathbb{N}_\infty\left[Y\sum_{i\in I_n}F(\mathscr{T}_i,s_i)\mathrm{e}^{-\mu\sigma_i}\mathrm{e}^{-\mu\sum_{j\neq i}\sigma_j}\right]. \end{aligned}$$

Using a Palm formula, we get:

$$\begin{aligned} &= \mathbb{N}_\infty\left[\left[\int_{\mathrm{T}_n}\ell(ds)\mathbb{N}_{\theta(s)}[F(\mathscr{T},s)\mathrm{e}^{-\mu\sigma}]\exp\left(-\int_{\mathrm{T}_n}\ell(ds)\int_0^\infty\frac{du}{\sqrt{2\pi}u^{3/2}}(1-\mathrm{e}^{-\mu\sigma})\right)\right]\right. \\ &= \mathbb{N}_\infty\left[Y\int_{\mathrm{T}_n}\ell(ds)\mathbb{N}_{\theta(s)}[F(\mathscr{T},s)\mathrm{e}^{-\mu\sigma}]\mathrm{e}^{-L_n\sqrt{2\mu}}\right], \end{aligned}$$

since $\mathbb{N}[1-\exp(-\mu\sigma)]=\sqrt{2\mu}$. We can disintegrate the σ-finite measure $\mathbb{N}_{\theta(s)}$ according to the total mass σ:

$$\begin{aligned} I(\mu) &= \mathbb{N}_\infty\left[Y\int_{\mathrm{T}_n}\ell(ds)\int_0^\infty\frac{dv}{\sqrt{2\pi}v^{3/2}}\mathbb{E}_{\theta(s)}^{(v)}[F(\mathscr{T},s)\mathrm{e}^{-\mu\sigma}]\mathrm{e}^{-L_n\sqrt{2\mu}}\right] \\ &= \mathbb{N}_\infty\left[Y\int_{\mathrm{T}_n}\ell(ds)\int_0^\infty\frac{dv}{\sqrt{2\pi}v^{3/2}}\mathbb{E}_{\theta(s)}^{(v)}[F(\mathscr{T},s)]\mathrm{e}^{-\mu v}\int_0^\infty L_n\frac{dr}{\sqrt{2\pi r^3}}\mathrm{e}^{-\mu r-L_n^2/(2r)}\right], \end{aligned}$$

using the well-known formula

$$\mathrm{e}^{a\sqrt{2s}}=\int_0^\infty\mathrm{e}^{-sr}\frac{a}{\sqrt{2\pi r^3}}\mathrm{e}^{-a^2/2r}\,dr,$$

for the Laplace transform of the density of the 1/2-stable subordinator (see for instance Chapter III, Proposition (3.7) in [RY05]). By the Fubini-Tonelli theorem, we then get:

$$\begin{aligned} I(\mu) &= \mathbb{N}_\infty\left[Y\int_{\mathrm{T}_n}\ell(ds)\int_0^\infty\frac{dv}{\sqrt{2\pi}v^{3/2}}\mathbb{E}_{\theta(s)}^{(v)}[F(\mathscr{T},s)]\mathrm{e}^{-\mu v}\int_v^\infty\frac{L_n\mathrm{e}^{-\mu(t-v)}\,dt}{\sqrt{2\pi}(t-v)^{3/2}}\mathrm{e}^{-L_n^2/(2t-2v)}\right] \\ &= \int_0^\infty\frac{\mathrm{e}^{-\mu t}dt}{\sqrt{2\pi}t^{3/2}}\mathbb{N}_\infty\left[Y\int_{\mathrm{T}_n}\ell(ds)\int_0^t\frac{L_nt^{3/2}\,dv}{\sqrt{2\pi}v^{3/2}(t-v)^{3/2}}\mathrm{e}^{-L_n^2/(2t-2v)}\mathbb{E}_{\theta(s)}^{(v)}[F(\mathscr{T},s)]\right]. \end{aligned}$$

Now, we can use the scaling property of the marked tree $(\mathscr{T},\theta)$ under $\mathbb{N}_\infty$ and that the fact the total mass σ has density $dt/(\sqrt{2\pi}t^{3/2})$ under $\mathbb{N}_\infty$, to get that, for any $\mathscr{F}_n$-measurable random variable Y,

$$\mathbb{E}_\infty^{(1)}\left[Y\sum_{i\in I_n}F(\mathscr{T}_i,s_i)\right]=\mathbb{N}_\infty\left[Y\frac{1}{\sqrt{2\pi}}\int_0^1\frac{L_n\,dv}{v^{3/2}(1-v)^{3/2}}\mathrm{e}^{-L_n^2/(2-2v)}\times\int_{\mathrm{T}_n}\ell(ds)\mathbb{E}_{\theta(s)}^{(v)}[F(\mathscr{T},s)]\right].$$

Now, recall the absolute continuity relation the distribution of T_n under $\mathbb{N}_\infty$ and under $\mathbb{E}_\infty^{(1)}$ (Corollary 4 in [Le 93b]): for any measurable bounded functional G,

$$\mathbb{E}_\infty^{(1)}[G(\mathrm{T}_n)] = \mathbb{N}_\infty\left[\ell(\mathrm{T}_n)\mathrm{e}^{-\ell(\mathrm{T}_n)^2/2}G(\mathrm{T}_n)\right].$$

Since $\exp(-L_n^2/(2-2v)) = \exp(-L_n^2/2)\cdot\exp(-L_n^2 v/(2-2v))$, we get:

$$\mathbb{E}_\infty^{(1)}\left[Y\sum_{i\in I_n} F(\mathscr{T}_i, s_i)\right] = \mathbb{E}_\infty^{(1)}\left[Y\frac{1}{\sqrt{2\pi}}\int_0^1 \frac{dv}{v^{3/2}(1-v)^{3/2}}\mathrm{e}^{-L_n^2 v/(2-2v)}\int_{\mathrm{T}_n}\ell(ds)\mathbb{E}_{\theta(s)}^{(v)}[F(\mathscr{T}, s)]\right].$$

Taking conditional expectations with respect to $\mathscr{F}_n$ gives the desired result. □

Remark 10. *Notice that if $F(\mathscr{T}, s) = \mathbf{m}(\mathscr{T})$, we find the striking identity*

$$\frac{1}{\sqrt{2\pi}}\int_0^1 \frac{L_n \mathrm{e}^{-L_n^2 v/(2-2v)}}{v^{1/2}(1-v)^{3/2}}\,dv = 1. \tag{4.19}$$

In other words, the function $f_a(v) = a\mathrm{e}^{-a^2 v/(2-2v)}/(\sqrt{2\pi}v^{1/2}(1-v)^{3/2})$ is a probability density on $(0,1)$ for any $a > 0$. This probability distribution has already been described in the context of the Aldous-Pitman fragmentation: if $a > 0$, Aldous and Pitman show that it is the distribution of the size of the fragment containing the root at time a. We refer to [AP98a, Ber06] for more information on the "tagged fragment" process in self-similar fragmentations.

Proof of Proposition 4.3

We now have everything we need to prove Proposition 4.3.

Proof of Proposition 4.3. We will start from Lemma 7.4 in [AD11]: we have a.s. for $n \geq 1$

$$-R_n \leq \mathbb{E}_\infty^{(1)}[\Theta|\mathscr{F}_n] - \frac{1}{L_n}\int_{\mathrm{T}_n^*}\theta(s)\ \ell(ds) \leq V_n, \tag{4.20}$$

where we noted $V_n = \mathbb{E}_\infty^{(1)}[\int_{\mathscr{E}_n}\theta(s)\mathbf{m}(ds)|\mathscr{F}_n]$ (recall the definition of $\mathscr{E}_n$ in (4.17)) and where $R_n = \exp(-L_n^2/4)\theta(h_{\emptyset,n})^2/4$. Furthermore, there $\mathbb{P}_\infty$-a.s. exists a constant $C > 0$ such that

$$R_n \leq Cn^8\mathrm{e}^{-L_n^2/8}.$$

Thus, considering that $L_n/\sqrt{2n}$ converges a.s. to 1 (4.8), we get that $n^{1/4}R_n$ converges a.s. to 0. Therefore, we needn't worry about the left-hand side of (4.20) and the only thing we need to prove is that $n^{1/4}V_n$ converges in distribution to 0 as $n\to\infty$. The proof in [AD11] uses a dominated convergence argument to show that V_n a.s. converges to 0, but we will need a more precise estimate for V_n. By definition, using the notation

$$\Theta_i^{(n)} = \int_{\mathscr{T}_i}\theta(s)\ \mathbf{m}(ds), \quad i\in I_n,$$

we have

$$V_n = \mathbb{E}_\infty^{(1)}\left[\int_{\mathscr{E}_n} \theta(s)\ \mathbf{m}(ds)\middle|\mathscr{F}_n\right] = \mathbb{E}_\infty^{(1)}\left[\sum_{i\in I_n} \Theta_i \mathbf{1}_{\{s_i\in[\![\emptyset,s_{\emptyset,n}]\!]\}}\middle|\mathscr{F}_n\right].$$

Using the disintegration formula from Lemma 4.5, we get:

$$V_n = \frac{1}{\sqrt{2\pi}}\int_0^1 \frac{dv}{v^{3/2}(1-v)^{3/2}}\mathrm{e}^{-L_n^2 v/(2-2v)}\int_{[\![\emptyset,s_{\emptyset,n}]\!]} \mathbb{E}_{\theta(s)}^{(v)}[\Theta]\ \ell(ds).$$

Using the fact that $\theta(s)$ is, conditionally on T_n, exponentially distributed with parameter s, we get:

$$\begin{aligned}
\mathbb{E}_\infty^{(1)}[V_n|\mathrm{T}_n] &= \frac{1}{\sqrt{2\pi}}\int_0^1 \frac{dv}{v^{3/2}(1-v)^{3/2}}\mathrm{e}^{-L_n^2 v/(2-2v)}\int_0^{h_{\emptyset,n}} ds\int_0^\infty s\mathrm{e}^{-st}\mathbb{E}_t^{(v)}[\Theta]\ dt \\
&\le \frac{1}{2}\int_0^1 \frac{dv}{v^{3/2}(1-v)^{3/2}}\mathrm{e}^{-L_n^2 v/(2-2v)}\int_0^{h_{\emptyset,n}} ds\left(\int_0^{v^{-1/2}} stv\mathrm{e}^{-st}\ dt + \int_{v^{-1/2}}^\infty s\sqrt{v}\mathrm{e}^{-st}\ dt\right),
\end{aligned}$$

using the domination $\mathbb{E}_q^{(v)}[\Theta] \le \sqrt{\pi/2}\min(qv,\sqrt{v})$ (Lemma 4.12). For technical reasons, we will restrict ourselves to the event $\{h_{\emptyset,n} < 1/2\}$, but this will not be too restrictive, since $h_{\emptyset,n} \to 0$ a.s. Computing the integrals, we eventually get that $\mathbb{E}_\infty^{(1)}[V_n|\mathrm{T}_n]\mathbf{1}_{\{h_{\emptyset,n}<1/2\}}$ is dominated by

$$W_n = \left(\frac{1}{2}\int_0^1 \frac{dv}{v^{1/2}(1-v)^{3/2}}\mathrm{e}^{-L_n^2 v/(2-2v)}\int_0^{h_{\emptyset,n}} \frac{1-\mathrm{e}^{-s/\sqrt{v}}}{s}\ ds\right)\mathbf{1}_{\{h_{\emptyset,n}<1/2\}}.$$

We will use the domination $(1-\exp(-s))/s \le \mathbf{1}_{[0,1]}(s) + 2/(s+1)\mathbf{1}_{(1,\infty)}(s)$, which gives:

$$\begin{aligned}
W_n &\le \frac{1}{2}\int_0^1 \frac{\mathrm{e}^{-L_n^2 v/(2-2v)}}{v^{1/2}(1-v)^{3/2}}\ dv\left(\frac{h_{\emptyset,n}}{\sqrt{v}}\mathbf{1}_{\{h_{\emptyset,n}/\sqrt{v}\le 1\}}\right. \\
&\qquad\qquad \left.+\left(1+2\log\left(\frac{h_{\emptyset,n}/\sqrt{v}+1}{2}\right)\right)\mathbf{1}_{\{h_{\emptyset,n}/\sqrt{v}\ge 1\}}\right)\mathbf{1}_{\{h_{\emptyset,n}<1/2\}} \\
&= \left(\frac{1}{2}\int_0^{h_{\emptyset,n}^2} \frac{\mathrm{e}^{-L_n^2 v/(2-2v)}}{v^{1/2}(1-v)^{3/2}}\left(1-2\log 2+2\log\left(\frac{h_{\emptyset,n}}{\sqrt{v}}+1\right)\right)\ dv\right)\mathbf{1}_{\{h_{\emptyset,n}<1/2\}} \qquad (4.21)\\
&\qquad\qquad +\left(\frac{1}{2}\int_{h_{\emptyset,n}^2}^1 \frac{\mathrm{e}^{-L_n^2 v/(2-2v)}}{v^{1/2}(1-v)^{3/2}}\frac{h_{\emptyset,n}}{\sqrt{v}}\ dv\right)\mathbf{1}_{\{h_{\emptyset,n}<1/2\}}. \qquad (4.22)
\end{aligned}$$

As far as (4.21) is concerned, we can dominate $\exp(-\alpha L_n^2 v/(1-v))$ by 1 and $(1-v)^{-3/2}$ by its value at $h_{\emptyset,n}^2$, *i.e.* $(1-h_{\emptyset,n}^2)^{-3/2} < (3/4)^{-3/2}$ to get:

$$(4.21) \le \frac{1}{2(3/4)^{3/2}}\int_0^{h_{\emptyset,n}^2} \frac{dv}{\sqrt{v}}\left(1-2\log 2+2\log\left(\frac{h_{\emptyset,n}}{\sqrt{v}}+1\right)\right)\mathbf{1}_{\{h_{\emptyset,n}<1/2\}} = C\cdot h_{\emptyset,n}\mathbf{1}_{\{h_{\emptyset,n}<1/2\}},$$

where C is some deterministic constant. Concerning (4.22), we can bound $1/\sqrt{v}$ by $1/h_{\varnothing,n}$, to get:

$$
\begin{aligned}
(4.22) &\le \left(\frac{1}{L_n}\int_{h^2_{\varnothing,n}}^1 \frac{1}{2}\frac{L_n e^{-L_n^2 v/(2-2v)}}{v^{1/2}(1-v)^{3/2}}\ dv\right)\mathbf{1}_{\{h_{\varnothing,n}<1/2\}}\\
&\le \left(\frac{1}{L_n}\int_{0}^1 \frac{1}{2}\frac{L_n e^{-L_n^2 v/(2-2v)}}{v^{1/2}(1-v)^{3/2}}\ dv\right)\mathbf{1}_{\{h_{\varnothing,n}<1/2\}} = \frac{\sqrt{\pi}}{\sqrt{2}L_n}\mathbf{1}_{\{h_{\varnothing,n}<1/2\}},
\end{aligned}
$$

by equation (4.19). Putting things together, we get that $\mathbb{P}_\infty$-a.s.

$$
\mathbb{E}^{(1)}_\infty\left[V_n|\mathrm{T}_n\right]\mathbf{1}_{\{h_{\varnothing,n}<1/2\}} \le C\cdot h_{\varnothing,n}\mathbf{1}_{\{h_{\varnothing,n}<1/2\}} + \frac{\sqrt{\pi}}{\sqrt{2}L_n}. \tag{4.23}
$$

Now, $n^{1/4}h_{\varnothing,n}\mathbf{1}_{\{h_{\varnothing,n}<1/2\}}$ converges in L^1 to 0 thanks to (4.9). Similarly, an easy moment computation using (4.7) for the density of L_n shows that $n^{1/4}/L_n$ also converges in L^1 to 0, so that the same is true for $n^{1/4}V_n\mathbf{1}_{\{h_{\varnothing,n}<1/2\}}$. Hence, $n^{1/4}V_n\mathbf{1}_{\{h_{\varnothing,n}<1/2\}}$ converges to 0 in probability. Since a.s. there is a (random) $n_0\ge 1$ such that $h_{\varnothing,n}<1/2$ for any $n\ge n_0$, we also get that $n^{1/4}V_n$ converges to 0 in probability. Combining this with the a.s. convergence to 0 for $n^{1/4}R_n$, we indeed get a convergence in probability:

$$
\lim_{n\to\infty} n^{1/4}\left(\mathbb{E}^{(1)}_\infty[\Theta|\mathscr{F}_n] - \frac{1}{L_n}\int_{\mathrm{T}_n^*}\theta(s)\ \ell(ds)\right) = 0. \tag{4.24}
$$

To get the announced result, we still have to prove that

$$
\lim_{n\to\infty} n^{1/4}\left(\frac{1}{L_n}-\frac{1}{\sqrt{2n}}\right)\int_{\mathrm{T}_n^*}\theta(s)\ \ell(ds) = 0. \tag{4.25}
$$

This is not difficult: simply write

$$
n^{1/4}\left(\frac{1}{L_n}-\frac{1}{\sqrt{2n}}\right)\int_{\mathrm{T}_n^*}\theta(s)\ell(ds) = n^{1/4}\left(1-\frac{L_n}{\sqrt{2n}}\right)\left(\frac{1}{L_n}\int_{\mathrm{T}_n^*}\theta(s)\ell(ds)\right).
$$

Now, recall that $1/L_n\int_{\mathrm{T}_n^*}\theta(s)\ell(ds)$ converges to Θ $\mathbb{P}_\infty$-a.s., hence in probability. Furthermore, we can compute

$$
n^{1/2}\mathbb{E}^{(1)}_\infty\left[\left(1-\frac{L_n}{\sqrt{2n}}\right)^2\right] = n^{1/2}\mathbb{E}^{(1)}_\infty\left(1+\frac{L_n^2}{2n}-2\frac{L_n}{\sqrt{2n}}\right),
$$

Using the density (4.7) of L_n, we easily get that

$$
\mathbb{E}^{(1)}_\infty[L_n] = \sqrt{2}\frac{\Gamma(n+1/2)}{\Gamma(n)} \quad ; \quad \mathbb{E}^{(1)}_\infty\left[L_n^2\right] = 2n.
$$

Therefore, after computations, we get $n^{1/2}\mathbb{E}^{(1)}_\infty[(1-L_n/\sqrt{2n})^2]\sim 1/(8\sqrt{n})$, so that in the end, $n^{1/4}(1-L_n/\sqrt{2n})$ converges to 0 in L^2. This implies convergence in probability, hence the convergence of (4.25). □

Rate of convergence in the Martingale Convergence Theorem

Before we can move on to Proposition 4.4, we are going to state a lemma that will be needed in the proof.

Lemma 4.6. *If $1<\alpha<2$, then, the sequence $\int_{\mathrm{T}_n^*}\theta(s)^\alpha \ell(ds)/L_n$ is bounded in $L^1(\mathbb{P}_\infty)$.*

Proof. The main idea is that the measure $\ell(ds)/L_n$ converges a.s. to the mass measure $\mathbf{m}(ds)$, in the sense of weak convergence of probability measures on $\mathscr{T}$. Since the function θ is neither continuous nor bounded on $\mathscr{T}$, we cannot use this fact directly, but it will be the inspiration for the proof. We will compute the first moment of $Z_n=\int_{\mathrm{T}_n^*}\theta(s)^\alpha \ell(ds)/L_n^*$, using the notation $L_n^*=\ell(\mathrm{T}_n^*)$. Since $\theta(s)$ is, conditionally on $\mathscr{T}$, exponentially distributed with parameter $\ell([\![\emptyset,s]\!])$, we get

$$
\begin{aligned}
\mathbb{E}_\infty^{(1)}[Z_n] &= \mathbb{E}_\infty^{(1)}\left[\int_{\mathrm{T}_n^*}\ell([\![\emptyset,s]\!])^{-\alpha}\,\frac{\ell(ds)}{L_n^*}\right]\\
&= \mathbb{E}_\infty^{(1)}\left[\int_{\mathscr{T}}(d(\emptyset,s)-d(s,\mathrm{T}_n^*))^{-\alpha}\mathbf{1}_{\mathscr{T}\setminus\mathscr{E}_n}(s)\ \mathbf{m}(ds)\right],
\end{aligned}
$$

where $d(s,\mathrm{T}_n^*)$ is the distance from the leaf s to the closed subtree T_n^* of $\mathscr{T}$. The last equality comes from the fact that if s is a leaf of $\mathscr{T}$ selected uniformly (according to $\mathbf{m}(ds)$) among all leaves of $\mathscr{T}\setminus\mathscr{E}_n$, then its projection $\pi(s,\mathrm{T}_n)$ on T_n is uniformly distributed (according to length measure) on T_n^*. We will rewrite the last expression so as to make the leaves $\emptyset,\mathrm{x}_1,\dots,\mathrm{x}_n$ apparent. The set $\mathscr{T}\setminus\mathscr{E}_n$ can be written as

$$
\mathscr{T}\setminus\mathscr{E}_n=\{s\in\mathscr{T},\ [\![\emptyset,\pi(\emptyset,\mathrm{T}_n^*)]\!]\cap[\![s,\pi(s,\mathrm{T}_n^*)]\!]=\emptyset\},\tag{4.26}
$$

since $\pi(\emptyset,\mathrm{T}_n^*)=s_{\emptyset,n}$. Note that T_n^* is actually the subtree spanned by the n leaves $\mathrm{x}_1,\dots,\mathrm{x}_n$ and that its definition does not depend on $\emptyset$ or on s.

We then apply the fundamental *re-rooting invariance* of the Brownian CRT, which implies, in this context, that when re-rooting $\mathscr{T}$ at s, the re-rooted tree $\mathscr{T}^s$ is distributed as a CRT, and the sequence $(\emptyset,\mathrm{x}_1,\dots,\mathrm{x}_n)$ is distributed as a sample of $n+1$ uniform leaves in $\mathscr{T}^s$. Thus,

$$
\begin{aligned}
\mathbb{E}_\infty^{(1)}[Z_n] &= \mathbb{E}_\infty^{(1)}\left[\int_{\mathscr{T}}(d(\emptyset,s)-d(s,\mathrm{T}_n^*))^{-\alpha}\mathbf{1}_{\{[\![\emptyset,\pi(\emptyset,\mathrm{T}_n^*)]\!]\cap[\![s,\pi(s,\mathrm{T}_n^*)]\!]=\emptyset\}}(s)\ \mathbf{m}(ds)\right]\\
&= \mathbb{E}_\infty^{(1)}\left[\int_{\mathscr{T}}(d(\emptyset,s)-h_{\emptyset,n})^{-\alpha}\mathbf{1}_{\{[\![\emptyset,\pi(\emptyset,\mathrm{T}_n^*)]\!]\cap[\![s,\pi(s,\mathrm{T}_n^*)]\!]=\emptyset\}}(s)\ \mathbf{m}(ds)\right],
\end{aligned}
$$

since in the re-rooting, $d(s,\mathrm{T}_n^*)$ becomes $d(\emptyset,\mathrm{T}_n^*)=h_{\emptyset,n}$. Therefore, we get, using (4.26) again,

$$
\begin{aligned}
\mathbb{E}_\infty^{(1)}[Z_n] &= \mathbb{E}_\infty^{(1)}\left[\int_{\mathscr{T}}(d(\emptyset,s)-h_{\emptyset,n})^{-\alpha}\mathbf{1}_{\mathscr{T}\setminus\mathscr{E}_n}(s)\ \mathbf{m}(ds)\right]\\
&= \mathbb{E}_\infty^{(1)}\left[\int_{\mathscr{T}_n^{(1)}\cup\mathscr{T}_n^{(2)}} d(s_{\emptyset,n},s)^{-\alpha}\mathbf{m}(ds)\right],
\end{aligned}
$$

where $\mathcal{T}_n^{(1)}$ and $\mathcal{T}_n^{(2)}$ are the connected components of $\mathcal{T} \setminus (\mathcal{E}_n \cup \{s_{\emptyset,n}\})$, joined together by their common root $s_{\emptyset,n}$. We can now use the self-similarity property of the fragmentation at heights of the Brownian CRT (see [Ber02]) which shows that, conditionally on $\sigma_n^{(1)} = \mathbf{m}(\mathcal{T}_n^{(1)})$ and $\sigma_n^{(2)} = \mathbf{m}(\mathcal{T}_n^{(2)})$, the trees $\mathcal{T}_n^{(1)}$ and $\mathcal{T}_n^{(2)}$ are rescaled copies of the Brownian CRT. Thus,

$$\begin{aligned}
\mathbb{E}_\infty^{(1)}[Z_n] &= \mathbb{E}_\infty^{(1)}\left[\int_{\mathcal{T}_n^{(1)}} d(s_{\emptyset,n},s)^{-\alpha}\mathbf{m}(ds)\right] + \mathbb{E}_\infty^{(1)}\left[\int_{\mathcal{T}_n^{(2)}} d(s_{\emptyset,n},s)^{-\alpha}\mathbf{m}(ds)\right] \\
&= \mathbb{E}_\infty^{(1)}\left[\int_{\mathcal{T}_n^{(1)}} (\sigma_n^{(1)})^{-\alpha/2}\left(\frac{d(s_{\emptyset,n},s)}{(\sigma_n^{(1)})^{1/2}}\right)^{-\alpha} \sigma_n^{(1)}\frac{\mathbf{m}(ds)}{\sigma_n^{(1)}}\right] \\
&\quad + \mathbb{E}_\infty^{(1)}\left[\int_{\mathcal{T}_n^{(2)}} (\sigma_n^{(2)})^{-\alpha/2}\left(\frac{d(s_{\emptyset,n},s)}{(\sigma_n^{(2)})^{1/2}}\right)^{-\alpha} \sigma_n^{(2)}\frac{\mathbf{m}(ds)}{\sigma_n^{(2)}}\right] \\
&= \mathbb{E}_\infty^{(1)}\left[\left((\sigma_n^{(1)})^{1-\alpha/2} + (\sigma_n^{(2)})^{1-\alpha/2}\right)\int_{\mathcal{T}} d(\emptyset,s)^{-\alpha}\mathbf{m}(ds)\right],
\end{aligned}$$

using the scaling invariance of the Brownian CRT. Then, as $0 < 1-\alpha/2$, we can simply dominate $(\sigma_n^{(1)})^{1-\alpha/2}$ and $(\sigma_n^{(1)})^{1-\alpha/2}$ by 1 to get that

$$\mathbb{E}_\infty^{(1)}[Z_n] \le 2\cdot\mathbb{E}_\infty^{(1)}\left[\int_{\mathcal{T}} d(\emptyset,s)^{-\alpha}\mathbf{m}(ds)\right].$$

Now, since $d(\emptyset,s)$ is Rayleigh-distributed under $\mathbb{E}_\infty^{(1)}$, we easily see that it has moments of order $-\alpha$ for any $\alpha < 2$, which shows that $\mathbb{E}_\infty^{(1)}[Z_n]$ is indeed bounded, ending our proof. □

We can now turn to the proof of Proposition 4.4.

Proof of Proposition 4.4. Let $M_n = \mathbb{E}_\infty^{(1)}[\Theta|\mathcal{F}_n]$. We will use the fact that

$$n^{1/4}\left(\Theta - \mathbb{E}_\infty^{(1)}[\Theta|\mathcal{F}_n]\right) = n^{1/4}\sum_{k=n+1}^{\infty}\mathbb{E}_\infty^{(1)}[M_k - M_{k-1}|\mathcal{F}_{k-1}]. \tag{4.27}$$

Let $\varepsilon > 0$ be small enough, and consider the events

$$E_k^1 = \{L_k \ge k^{1/2-\varepsilon}\} \quad ; \quad E_k^2 = \{k^{-2} \le h_{\emptyset,k} \le 1/2\}, \quad k \ge 1.$$

Recalling that L_k^2 is distributed as the sum of k independent exponential random variables with parameter 1, a simple application of Chernoff's inequality shows that

$$\mathbb{P}_\infty\left(E_k^1\right) \ge 1 - k^{-\varepsilon k}. \tag{4.28}$$

For E_k^2, we can use the moment estimation (4.9) for $h_{\emptyset,k}$ to find that, for any $0 \le \alpha \le 1$, and for any $\beta > 0$,

$$\begin{aligned}
1 - \mathbb{P}_\infty\left(E_k^2\right) &= \mathbb{P}_\infty\left(\{h_{\emptyset,k} > 1/2\} \cup \{h_{\emptyset,k} < k^{-2}\}\right) \\
&\le \mathbb{P}_\infty(h_{\emptyset,k} > 1/2) + \mathbb{P}_\infty\left(h_{\emptyset,k}^{-1} \ge k^2\right) \\
&\le 2^\beta\mathbb{E}_\infty^{(1)}\left[h_{\emptyset,k}^\beta\right] + k^{-2\alpha}\mathbb{E}_\infty^{(1)}\left[h_{\emptyset,k}^{-\alpha}\right] \\
&\sim C\cdot k^{-\beta/2} + C'\cdot k^{-2\alpha}k^{\alpha/2}.
\end{aligned}$$

Hence, by taking $\alpha = 1-\eta$ (and $\beta > 3\alpha$), we get, for sufficiently large k,

$$\mathbb{P}_\infty\left(E_k^2\right) \geq 1 - k^{-3/2+3/2\eta}. \tag{4.29}$$

Thus, combining equations (4.28) and (4.29), we get that

$$\sum_{k\geq 1} \mathbb{P}_\infty\left(\left(E_k^1\right)^c \cup \left(E_k^2\right)^c\right) \leq \sum_{k\geq 1} \mathbb{P}_\infty\left((E_k^1)^c\right) + \mathbb{P}_\infty\left((E_k^2)^c\right) < \infty.$$

Thus, by the Borel-Cantelli lemma, there a.s. exists $k_0 \geq 1$ such that for $k \geq k_0$, $L_k \geq k^{1/2-\varepsilon}$ and $k^{-2} \leq h_{\emptyset,k} \leq 1/2$. We will use this truncating events in the following way: since the event $E_k = E_k^1 \cap E_k^2$ is $\mathscr{F}_k$-measurable, the usual martingale computations show that

$$\mathbb{E}_\infty^{(1)}\left[\left(n^{1/4}\sum_{k=n}^{\infty}(M_k - M_{k-1})\mathbf{1}_{E_{k-1}}\right)^2\right] = n^{1/2}\sum_{k=n}^{\infty}\mathbb{E}_\infty^{(1)}[(M_k - M_{k-1})^2\mathbf{1}_{E_{k-1}}].$$

We will now give precise estimations of $\mathbb{E}_\infty^{(1)}[(M_k - M_{k-1})^2|\mathscr{F}_{k-1}]$ using the disintegration formula from Lemma 4.5. By definition, for all $k \geq 1$, we can write

$$\Theta = \int_{\mathscr{T}} \theta(s)\ \mathbf{m}(ds) = \sum_{i\in I_k}\Theta_i^{(k)}.$$

Then,

$$\begin{aligned}
M_k &= \mathbb{E}_\infty^{(1)}[\Theta|\mathscr{F}_k] \\
&= \mathbb{E}_\infty^{(1)}\left[\sum_{i\in I_{k-1}}\Theta_i^{(k-1)}\middle|\mathscr{F}_k\right] \\
&= \mathbb{E}_\infty^{(1)}[\Theta_{i_k}|\mathscr{F}_k] + \mathbb{E}_\infty^{(1)}\left[\sum_{i\in I_{k-1}\setminus\{i_k\}}\Theta_i\middle|\mathscr{F}_k\right],
\end{aligned}$$

where i_k is the unique index in I_{k-1} such that $\mathrm{x}_k \in \mathscr{T}_{i_k}$. We then define:

$$\begin{aligned}
G_k &= \mathbb{E}_\infty^{(1)}\left[\Theta_{i_k}^{(k-1)}|\mathscr{F}_k\right] \\
H_k &= \mathbb{E}_\infty^{(1)}\left[\sum_{i\in I_{k-1}\setminus\{i_k\}}\Theta_i^{(k-1)}\middle|\mathscr{F}_k\right] - \mathbb{E}_\infty^{(1)}\left[\sum_{i\in I_{k-1}}\Theta_i^{(k-1)}\middle|\mathscr{F}_{k-1}\right],
\end{aligned}$$

so that we have $M_k - M_{k-1} = G_k + H_k$ and

$$\mathbb{E}_\infty^{(1)}\left[(M_k - M_{k-1})^2|\mathscr{F}_{k-1}\right] \leq 2\mathbb{E}_\infty^{(1)}\left[G_k^2|\mathscr{F}_{k-1}\right] + 2\mathbb{E}_\infty^{(1)}\left[H_k^2|\mathscr{F}_{k-1}\right].$$

As far as G_k is concerned, we note that, conditionally on $\mathscr{F}_k$, $\Theta_{i_k}^{(k-1)}$ can be written as $\sum_{i\in I_k}\Theta_i^{(k)}\mathbf{1}_{\{s_i\in \mathrm{B}_k\}}$, so that we can use the disintegration formula of Lemma 4.5 to get:

$$G_k = \frac{1}{\sqrt{2\pi}}\int_0^1 \frac{\mathrm{e}^{-L_k^2 v/(2-2v)}}{v^{3/2}(1-v)^{3/2}}\ dv\int_{\mathrm{B}_k}\mathbb{E}_{\theta(s)}^{(v)}[\Theta]\ \ell(ds).$$

Hence, using this expression, we can now compute:

$$\begin{aligned}\mathbb{E}_\infty^{(1)}[G_k^2|\mathscr{F}_{k-1}] &= \mathbb{E}_\infty^{(1)}\left[\left(\frac{1}{\sqrt{2\pi}}\int_0^1 \frac{\mathrm{e}^{-L_k^2 v/(2-2v)}}{v^{3/2}(1-v)^{3/2}}\,dv\int_{\mathrm{B}_k}\mathbb{E}_{\theta(s)}^{(v)}[\Theta]\;\ell(ds)\right)^2\Big|\mathscr{F}_{k-1}\right]\\ &\le \mathbb{E}_\infty^{(1)}\left[\left(\frac{1}{\sqrt{2\pi}}\int_0^1 \frac{\mathrm{e}^{-L_k^2 v/(2-2v)}}{v^{1/2}(1-v)^{3/2}}\,dv\int_{\mathrm{B}_k}\ell(ds)\theta(s)\right)^2\Big|\mathscr{F}_{k-1}\right],\end{aligned}$$

since $\mathbb{E}_{\theta(s)}^{(v)}[\Theta]\le v\theta(s)$ (Lemma 4.12). Now, the measure

$$\frac{L_n \mathrm{e}^{-L_n^2 v/(2-2v)}}{\sqrt{2\pi}v^{1/2}(1-v)^{3/2}}dv$$

is a probability density on [0,1] (cf. (4.19)), so that we get, using the fact that $L_{k-1}<L_k$,

$$\begin{aligned}\mathbb{E}_\infty^{(1)}\left[G_k^2|\mathscr{F}_{k-1}\right] &\le \mathbb{E}_\infty^{(1)}\left[\left(\frac{1}{L_k}\int_{\mathrm{B}_k}\ell(ds)\theta(s)\right)^2\Big|\mathscr{F}_{k-1}\right]\\ &\le \frac{1}{L_{k-1}^2}\mathbb{E}_\infty^{(1)}\left[\left(\int_{\mathrm{B}_k}\ell(ds)\theta(s)\right)^2\Big|\mathscr{F}_{k-1}\right].\end{aligned}$$

Now, conditionally on $\mathscr{F}_{k-1}$, the record process on B_k has the distribution of an independent record process on $\mathbf{R}_+$, started from $\theta(s_k)$, stopped at time $\ell(\mathrm{B}_k)$. Furthermore, it is a consequence from the stickbreaking construction of Aldous (see [Ald91a]) that, conditionally on $\mathscr{F}_{k-1}$, the random variables s_k and $\ell(\mathrm{B}_k)$ are independent. Furthermore, s_k is distributed uniformly on T_{k-1}, and $\ell(\mathrm{B}_k)$ can be expressed as the length of the interval between the $(k-1)$th and the kth jump of a Poisson process with intensity $t\mathbf{1}_{[0,\infty)}(t)dt$. Therefore, conditionally on $\mathscr{F}_{k-1}$, $\ell(\mathrm{B}_k)$ has density

$$r_{L_{k-1}}(dx)=(L_{k-1}+x)\mathrm{e}^{-x^2/2-L_{k-1}x}\,dx. \tag{4.30}$$

Thus, using the notation $F(q,t)=\mathbf{E}_q[(\int_0^t\theta(s)ds)^2]$ for $0<q<\infty$ and $t\ge 0$, we get

$$\mathbb{E}_\infty^{(1)}\left[G_k^2|\mathscr{F}_{k-1}\right]\le\frac{1}{L_{k-1}^2}\int_{\mathrm{T}_{k-1}}\frac{\ell(ds)}{L_{k-1}}\int_0^\infty r_{L_{k-1}}(dx)F(\theta(s),x).$$

We will cut the integral in two parts, according to $\mathrm{T}_{k-1}=\mathrm{T}_{k-1}^*\cup(\mathrm{T}_{k-1}\setminus\mathrm{T}_{k-1}^*)$. We then use Lemma 4.10 to dominate $F(\theta(s),x)$: inequality (4.47) for $s\in\mathrm{T}_{k-1}^*$ and (4.48) for $s\in\mathrm{T}_{k-1}\setminus\mathrm{T}_{k-1}^*$. This leads to:

$$\begin{aligned}\mathbb{E}_\infty^{(1)}\left[G_k^2|\mathscr{F}_{k-1}\right] &\le\frac{1}{L_{k-1}^2}\int_{\mathrm{T}_{k-1}^*}\frac{\ell(ds)}{L_{k-1}}\int_0^\infty r_{L_{k-1}}(dx)\left(C_1\theta(s)^{3/2}x^{3/2}+C_2\theta(s)x^2\right)\\ &+\frac{1}{L_{k-1}^2}\int_{\mathrm{T}_{k-1}\setminus\mathrm{T}_{k-1}^*}\frac{\ell(ds)}{L_{k-1}}\int_0^\infty r_{L_{k-1}}(dx)\left(C_3\theta(s)^{1/2}x^{1/2}+C_4\theta(s)^{-1/2}x^{1/2}\right)\end{aligned}$$

$$
\begin{aligned}
&= \frac{C_1}{L_{k-1}^2}\left(\int_0^\infty r_{L_{k-1}}(dx)x^{3/2}\right)\left(\int_{\mathrm{T}_{k-1}^*}\frac{\ell(ds)}{L_{k-1}}\theta(s)^{3/2}\right)\\
&+\frac{C_2}{L_{k-1}^2}\left(\int_0^\infty r_{L_{k-1}}(dx)x^{2}\right)\left(\int_{\mathrm{T}_{k-1}^*}\frac{\ell(ds)}{L_{k-1}}\theta(s)\right)\\
&+\frac{C_3}{L_{k-1}^3}\left(\int_0^\infty r_{L_{k-1}}(dx)x^{1/2}\right)\left(\int_{\mathrm{T}_{k-1}\setminus\mathrm{T}_{k-1}^*}\ell(ds)\theta(s)^{1/2}\right)\\
&+\frac{C_4}{L_{k-1}^3}\left(\int_0^\infty r_{L_{k-1}}(dx)x^{1/2}\right)\left(\int_{\mathrm{T}_{k-1}\setminus\mathrm{T}_{k-1}^*}\ell(ds)\theta(s)^{-1/2}\right).
\end{aligned}
$$

We can then compute, using Lemma 4.11 for the asymptotic moments of $r_{L_{k-1}}(dx)$:

$$
\begin{aligned}
\mathbb{E}_\infty^{(1)}\left[G_k^2\mathbf{1}_{E_{k-1}}\right] &= \mathbb{E}_\infty^{(1)}\left[\mathbb{E}_\infty^{(1)}\left[G_k^2|\mathscr{F}_{k-1}\right]\mathbf{1}_{E_{k-1}}\right]\\
&\leq \mathbb{E}_\infty^{(1)}\left[\int_{\mathrm{T}_{k-1}^*}\frac{\ell(ds)}{L_{k-1}}\theta(s)^{3/2}\right]\cdot O(k^{-7/4+7/2\varepsilon}) && (4.31)\\
&+\mathbb{E}_\infty^{(1)}\left[\left(\int_{\mathrm{T}_{k-1}^*}\frac{\ell(ds)}{L_{k-1}}\theta(s)\right)\mathbf{1}_{E_{k-1}}\right]\cdot O(k^{-2+4\varepsilon}) && (4.32)\\
&+\mathbb{E}_\infty^{(1)}\left[\left(\int_{\mathrm{T}_{k-1}\setminus\mathrm{T}_{k-1}^*}\ell(ds)\theta(s)^{1/2}\right)\right]\cdot O(k^{-7/4+7/2\varepsilon}) && (4.33)\\
&+\mathbb{E}_\infty^{(1)}\left[\left(\int_{\mathrm{T}_{k-1}\setminus\mathrm{T}_{k-1}^*}\ell(ds)\theta(s)^{-1/2}\right)\right]\cdot O(k^{-7/4+7/2\varepsilon}). && (4.34)
\end{aligned}
$$

Using Lemma 4.6, we see that (4.31) is indeed of the order $k^{-7/4+7/2\varepsilon}$. As far as (4.32) is concerned, we will show the following lemma, which will be useful later on, and which implies in particular that (4.32) is of the order $k^{-2+4\varepsilon}$.

Lemma 4.7. $\mathbb{E}_\infty^{(1)}\left[\left(\int_{\mathrm{T}_{k-1}^*}\theta(s)\ell(ds)/L_{k-1}\right)^2\mathbf{1}_{E_{k-1}}\right]$ *is bounded as* $k\to\infty$.

Proof of Lemma 4.7. Recall (4.20):

$$
-R_{k-1}\leq \mathbb{E}_\infty^{(1)}[\Theta|\mathscr{F}_{k-1}]-\frac{1}{L_{k-1}}\int_{\mathrm{T}_{k-1}^*}\theta(s)\ \ell(ds)\leq V_{k-1}. \tag{4.35}
$$

Therefore, we can write

$$
\begin{aligned}
\mathbb{E}_\infty^{(1)}\left[\left(\int_{\mathrm{T}_{k-1}^*}\frac{\ell(ds)}{L_{k-1}}\theta(s)-\mathbb{E}_\infty^{(1)}[\Theta|\mathscr{F}_{k-1}]\right)^2\mathbf{1}_{E_{k-1}}\right] &\leq \mathbb{E}_\infty^{(1)}\left[(R_{k-1}\vee V_{k-1})^2\mathbf{1}_{E_{k-1}}\right]\\
&\leq \mathbb{E}_\infty^{(1)}\left[R_{k-1}^2\mathbf{1}_{E_{k-1}}\right]+\mathbb{E}_\infty^{(1)}\left[V_{k-1}^2\mathbf{1}_{E_{k-1}}\right].
\end{aligned}
$$

Using (4.23), we can see that, since $E_{k-1}\in\sigma(\{\mathrm{T}_n\})$,

$$
\mathbb{E}_\infty^{(1)}\left[V_{k-1}^2\mathbf{1}_{E_{k-1}}\right]\leq \mathbb{E}_\infty^{(1)}\left[\left(C\cdot h_{\emptyset,k-1}+\frac{\sqrt{\pi}}{\sqrt{2}L_{k-1}}\right)^2\mathbf{1}_{E_{k-1}}\right]\leq \mathbb{E}_\infty^{(1)}\left[(C\cdot h_{\emptyset,k-1}+\sqrt{\pi/2}/L_{k-1})^2\right].
$$

Hence, as $h_{\emptyset,k-1}$ and L_{k-1}^{-1} are integrable and decrease to 0 a.s., $\mathbb{E}_\infty^{(1)}[V_{k-1}^2 \mathbf{1}_{E_{k-1}}]$ converges to 0 by monotone convergence. As for $\mathbb{E}_\infty^{(1)}[R_{k-1}^2 \mathbf{1}_{E_{k-1}}]$, we use the fact that, conditionally on T_{k-1}, $\theta(h_{\emptyset,k-1})$ is exponentially distributed with parameter $h_{\emptyset,k-1}$ to find

$$\begin{aligned}\mathbb{E}_\infty^{(1)}\left[R_{k-1}^2 \mathbf{1}_{E_{k-1}}\right] &= \mathbb{E}_\infty^{(1)}\left[\frac{1}{16}\mathrm{e}^{-L_{k-1}^2/2} h_{\emptyset,k-1}^{-4}\mathbf{1}_{E_{k-1}}\right] \\ &\leq \frac{1}{16}k^8 \mathbb{E}_\infty^{(1)}\left[\mathrm{e}^{-L_{k-1}^2/2}\right],\end{aligned}$$

which easily converges to 0 as $k\to\infty$. Hence, since $\mathbb{E}_\infty^{(1)}[\Theta|\mathscr{F}_{k-1}]$ converges in L^2 to Θ, it is of course L^2-bounded, so that $\mathbb{E}_\infty^{(1)}[(\int_{\mathrm{T}_{k-1}^*}\theta(s)\ell(ds)/L_{k-1})^2\mathbf{1}_{E_{k-1}}]$ is indeed bounded as $k\to\infty$, as announced. □

In the two remaining terms (4.33) and (4.34), the integral is taken on a single branch; therefore, we can use the linear case to get

$$\begin{aligned}\mathbb{E}_\infty^{(1)}\left[\left(\int_{\mathrm{T}_{k-1}\setminus \mathrm{T}_{k-1}^*}\ell(ds)\theta(s)^{1/2}\right)\mathbf{1}_{E_{k-1}}\right] &= \mathbb{E}_\infty^{(1)}\left[\mathbf{E}_\infty\left[\int_0^{h_{\emptyset,k-1}}\theta(s)^{1/2}ds\right]\mathbf{1}_{E_{k-1}}\right] \\ &= C\cdot \mathbb{E}_\infty^{(1)}[h_{\emptyset,k-1}^{1/2}\mathbf{1}_{E_{k-1}}],\end{aligned}$$

which easily converges to 0 as $k\to\infty$. A similar argument shows that (4.34) converges to 0 as $\mathbb{E}_\infty^{(1)}[h_{\emptyset,k-1}^{3/2}\mathbf{1}_{E_{k-1}}]$. Putting everything together, we find that $\mathbb{E}_\infty^{(1)}[G_k^2\mathbf{1}_{E_{k-1}}]$ is of the order $k^{-7/4+7/2\varepsilon}$ as $k\to\infty$, so that the remainder $\sum_{k=n}^\infty \mathbb{E}_\infty^{(1)}[G_k^2\mathbf{1}_{E_{k-1}}]$ is of the order $n^{-3/4+7/2\varepsilon}$.

Turning to H_k, we note that $I_{k-1}\setminus\{i_k\} = \{i\in I_k,\ s_i\notin \mathrm{B}_k\}$, so that, using Lemma 4.5, we get:

$$\begin{aligned}H_k =& \mathbb{E}_\infty^{(1)}\left[\sum_{i\in I_k}\Theta_i^{(k)}\mathbf{1}_{\{s_i\notin \mathrm{B}_k\}}\middle|\mathscr{F}_k\right] - \mathbb{E}_\infty^{(1)}\left[\sum_{i\in I_{k-1}}\Theta_i^{(k-1)}\middle|\mathscr{F}_{k-1}\right] \\ =& \int_0^1 \frac{\mathrm{e}^{-L_k^2 v/(2-2v)}}{\sqrt{2\pi}v^{3/2}(1-v)^{3/2}}dv\int_{\mathrm{T}_k}\ell(ds)\mathbb{E}_{\theta(s)}^{(v)}[\Theta]\mathbf{1}_{\{s\notin \mathrm{B}_k\}} \\ & -\int_0^1 \frac{\mathrm{e}^{-L_{k-1}^2 v/(2-2v)}}{\sqrt{2\pi}v^{3/2}(1-v)^{3/2}}dv\int_{\mathrm{T}_{k-1}}\ell(ds)\mathbb{E}_{\theta(s)}^{(v)}[\Theta],\end{aligned}$$

thus, considering that $\mathrm{T}_k = \mathrm{T}_{k-1}\cup(\mathrm{B}_k\setminus\{s_k\})$, and that of course $\ell(\{s_k\})=0$,

$$H_k = \int_0^1\frac{dv}{\sqrt{2\pi}v^{3/2}(1-v)^{3/2}}\int_{\mathrm{T}_{k-1}}\ell(ds)\mathbb{E}_{\theta(s)}^{(v)}[\Theta]\left(\mathrm{e}^{-L_k^2 v/(2-2v)}-\mathrm{e}^{-L_{k-1}^2 v/(2-2v)}\right).$$

We then use the inequality $|\mathrm{e}^{-at}-\mathrm{e}^{-as}|\leq a\mathrm{e}^{-at}(s-t)$, valid for any $a>0$, and $t\leq s$, to find:

$$\begin{aligned}\mathbb{E}_\infty^{(1)}\left[H_k^2|\mathscr{F}_{k-1}\right] \leq \left(\frac{1}{\sqrt{2\pi}}\int_0^1\frac{\mathrm{e}^{-L_{k-1}^2 v/(2-2v)}}{v^{3/2}(1-v)^{3/2}}\frac{v}{2-2v}dv\int_{\mathrm{T}_{k-1}}\mathbb{E}_{\theta(s)}^{(v)}[\Theta]\ \ell(ds)\right)^2 \\ \times\mathbb{E}_\infty^{(1)}\left[(L_k^2-L_{k-1}^2)^2|\mathscr{F}_{k-1}\right]. \quad (4.36)\end{aligned}$$

On the one hand, we will use the change of variables

$$u = L_{k-1}^2 v/(2-2v) \Leftrightarrow v = u/(L_{k-1}^2/2+u).$$

in the integral, which gives:

$$\left(\frac{1}{\sqrt{2\pi}}\int_0^\infty \frac{\mathrm{e}^{-u}}{(L_{k-1}^2/2)^{1/2}}\frac{L_{k-1}^2/2+u}{\sqrt{u}}\,du\int_{\mathrm{T}_{k-1}}\frac{\ell(ds)}{L_{k-1}^2}\mathbb{E}_{\theta(s)}^{(u/(L_{k-1}^2/2+u))}[\Theta]\right)^2. \tag{4.37}$$

We then cut the integral in two parts, according to $\mathrm{T}_{k-1} = \mathrm{T}_{k-1}^* \cup (\mathrm{T}_{k-1}\setminus \mathrm{T}_{k-1}^*)$, and we use the simple domination $\mathbb{E}_{\theta(s)}^{(v)}[\Theta] \le v\theta(s)$ on T_{k-1}^*, and the domination $\mathbb{E}_{\theta(s)}^{(v)}[\Theta] \le \mathbb{E}_{\infty}^{(v)}[\Theta] = \sqrt{\pi v/2}$ on $\mathrm{T}_{k-1}\setminus \mathrm{T}_{k-1}^*$ to get

$$\begin{aligned}(4.37) \le \Bigg(\frac{1}{\sqrt{\pi}}\int_0^\infty \frac{du}{L_{k-1}^3}\frac{L_{k-1}^2/2+u}{\sqrt{u}}\mathrm{e}^{-u}\int_{\mathrm{T}_{k-1}^*}\ell(ds)\theta(s)\frac{u}{L_{k-1}^2/2+u}\\ +\int_0^\infty \frac{du}{L_{k-1}^3}\frac{L_{k-1}^2/2+u}{\sqrt{2u}}\mathrm{e}^{-u}h_{\emptyset,k-1}\frac{\sqrt{u}}{\sqrt{L_{k-1}^2/2+u}}\Bigg)^2.\end{aligned}$$

The integrals can be computed, giving

$$(4.37) \le \left(\frac{1}{2\sqrt{\pi}L_{k-1}^2}\int_0^\infty \sqrt{u}\mathrm{e}^{-u}\int_{\mathrm{T}_{k-1}^*}\frac{\ell(ds)}{L_{k-1}}\theta(s) + \int_0^\infty \frac{du}{\sqrt{2}L_{k-1}^2}\sqrt{1/2+u/L_{k-1}^2}\,\mathrm{e}^{-u}h_{\emptyset,k-1}\right)^2.$$

On the other hand, the term $\mathbb{E}_\infty^{(1)}[(L_k^2-L_{k-1}^2)^2|\mathscr{F}_{k-1}]$ appearing in the domination (4.36) can be expanded into

$$\mathbb{E}_\infty^{(1)}\left[\ell(\mathrm{B}_k)^4|\mathscr{F}_{k-1}\right] + 4L_{k-1}^2\mathbb{E}_\infty^{(1)}\left[\ell(\mathrm{B}_k)^2|\mathscr{F}_{k-1}\right] + 4L_{k-1}\mathbb{E}_\infty^{(1)}\left[\ell(\mathrm{B}_k)^3|\mathscr{F}_{k-1}\right]$$

Then, recall the density (4.30) of $\ell(\mathrm{B}_k)$ conditionally on $\mathscr{F}_{k-1}$. In the proof of Lemma 4.11, we show that for any $\lambda > 0$, we have a.s.

$$\mathbb{E}_\infty^{(1)}\left[\ell(\mathrm{B}_k)^\lambda|\mathscr{F}_{k-1}\right] = \int r_{L_{k-1}}(dx)x^\lambda \le C_1\cdot L_{k-1}^{-\lambda} + C_2\cdot L_{k-1}^{-\lambda-2}$$

with C_1 and C_2 deterministic constants. Thus, $\mathbb{E}_\infty^{(1)}[(L_k^2-L_{k-1}^2)^2|\mathscr{F}_{k-1}]$ is a.s. bounded by $F(L_{k-1})$, where F is a nonincreasing bounded nonnegative function. In the end, we get

$$\begin{aligned}\mathbb{E}_\infty^{(1)}\left[H_k^2\mathbf{1}_{E_{k-1}}\right] &\le \mathbb{E}_\infty^{(1)}\Bigg[\Bigg(\frac{C}{L_{k-1}^2}\int_{\mathrm{T}_{k-1}^*}\theta(s)\frac{\ell(ds)}{L_{k-1}}\\ &\qquad +\int_0^\infty \frac{\mathrm{e}^{-u}\,du}{2L_{k-1}^2}\sqrt{1/2+u/L_{k-1}^2}\,h_{\emptyset,k}\Bigg)^2 F(L_{k-1})\mathbf{1}_{E_{k-1}}\Bigg]\\ &\le F(k^{2-4\varepsilon})\Bigg(C\cdot k^{-2+4\varepsilon}\mathbb{E}_\infty^{(1)}\left[\left(\int_{\mathrm{T}_{k-1}^*}\theta(s)\frac{\ell(ds)}{L_{k-1}}\right)^2\right]\\ &\qquad +C'\cdot k^{-2+4\varepsilon}\left(\int_0^\infty \mathrm{e}^{-u}\sqrt{1/2+u/k^{1-2\varepsilon}}\right)^2\mathbb{E}_\infty^{(1)}[h_{\emptyset,k}^2]\Bigg).\end{aligned}$$

Hence, using the fact that $\int_{\mathrm{T}^*_{k-1}} \theta(s)\ell(ds)/L_{k-1}$ is bounded in L^2 (Lemma 4.7), we find that $\mathbb{E}^{(1)}_\infty[H_k^2 \mathbf{1}_{E_{k-1}}] = O(k^{-2+4\varepsilon})$. Putting this together with the estimate on $\mathbb{E}^{(1)}_\infty[G_k^2 \mathbf{1}_{E_{k-1}}]$, we get that $\mathbb{E}^{(1)}_\infty[(M_k - M_{k-1})^2 \mathbf{1}_{E_{k-1}}] = O(k^{-7/4+7/2\varepsilon})$. If $\varepsilon < 1/14$,

$$\mathbb{E}^{(1)}_\infty[(M_k - M_{k-1})^2 \mathbf{1}_{E_{k-1}}] = O(k^{-7/4+7/2\varepsilon}) = o(k^{-3/2}).$$

Hence, we get

$$\lim_{n\to\infty} n^{1/2} \sum_{k=n}^{\infty} \mathbb{E}^{(1)}_\infty \left[(M_k - M_{k-1})^2 \mathbf{1}_{E_{k-1}}\right] = 0.$$

This shows that the random sequence $n^{1/4}\sum_{k=n}^\infty (M_k - M_{k-1})\mathbf{1}_{E_{k-1}}$ converges to 0 in L^2, hence in probability. But, since there a.s. exists $k_0 \geq 1$ such that $\mathbf{1}_{E_k} = 1$ for all $k \geq k_0$, the sequence $n^{1/4}\sum_{k=n}^\infty (M_k - M_{k-1})$ also converges to 0 in probability, which is what we wanted to prove. □

4.2 Proof of the main theorem

We can now turn to the proof of the actual convergence towards a nontrivial limit, in the asymptotic $n^{1/4}$. The main idea is to apply the Martingale Central Limit Theorem (Corollary 3.1 in [HH80]) to

$$M_n^* = X_n^* - \int_{\mathrm{T}_n^*} \theta(s)\ \ell(ds).$$

We recall this theorem below for convenience:

Theorem (Hall, Heyde [HH80]). *Let $(M_n,\ n \geq 1)$ be a zero-mean square-integrable $(\mathscr{G}_n)$-martingale, and let η^2 be an a.s. finite random variable. Suppose that, for some sequence a_n increasing to $+\infty$, we have*

1. (Asymptotic smallness) *For all $\varepsilon > 0$, we have the convergence in probability*

$$\lim_{n\to\infty} a_n^{-2} \sum_{k=1}^{n} \mathbf{E}\left[(M_k - M_{k-1})^2 \mathbf{1}_{\{|M_k - M_{k-1}| > \varepsilon a_k\}} \middle| \mathscr{G}_{k-1}\right] = 0$$

2. (Convergence of the conditional variance) *We have the convergence in probability*

$$\lim_{n\to\infty} a_n^{-2} \sum_{k=1}^{n} \mathbf{E}\left[(M_k - M_{k-1})^2 | \mathscr{G}_{k-1}\right] = \eta^2.$$

Then, the sequence $(a_n^{-1} M_n,\ n \geq 1)$ converges in distribution to a random variable Z with characteristic function $\mathbf{E}[\exp(-\eta^2 t^2/2)]$.

However, M_n^* is not a martingale in the filtration $(\mathscr{F}_n,\ n \geq 1)$, because the $(n+1)$st branch B_{n+1} might be connected to T_n through a vertex on $[\![\emptyset, s_{\emptyset,n}]\!]$. In that case, $M_{n+1}^* - M_n^*$ has a nonnegative $\mathscr{F}_n$-measurable part, corresponding to the atoms on $[\![s_{\emptyset,n+1}, s_{\emptyset,n}]\!]$. For this reason, we will consider

$$\widehat{M}_n = \sum_{s \in \mathrm{T}_n \setminus \mathrm{T}_1} \mathbf{1}_{\{\theta(s-) > \theta(s)\}} - \int_{\mathrm{T}_n \setminus \mathrm{T}_1} \theta(s)\ \ell(ds), \quad n \geq 2$$

and $\widehat{M}_1 = 0$. The process $(\widehat{M}_n, n \geq 1)$ is a $(\mathscr{F}_n)$-martingale. It is actually more convenient to introduce the filtration $(\mathscr{G}_n,\ n \geq 1)$, defined by:

$$\mathscr{G}_n = \sigma(\{(\mathrm{T}_m,\ m \geq 1),\ (\theta(s),\ s \in \mathrm{T}_n)\}),$$

Notice that the branching point $s_{n+1} = \mathrm{B}_{n+1} \cap \mathrm{T}_n$, as well as $\ell(\mathrm{B}_{n+1})$ and $\theta(s_{n+1})$ are all $\mathscr{G}_n$-measurable. In this filtration, $\widehat{M}$ is also a martingale. Indeed, it is obvious that $\widehat{M}$ is $\mathscr{G}$-adapted. Furthermore, we have

$$\widehat{M}_{n+1} - \widehat{M}_n = \sum_{s \in \mathrm{B}_{n+1}} \mathbf{1}_{\{\theta(s-) > \theta(s)\}} - \int_{\mathrm{B}_{n+1}} \theta(s)\ \ell(ds),$$

which is, conditionally on $\mathscr{G}_n$, distributed as $N_{\ell(\mathrm{B}_{n+1})}$, where N is the martingale from (4.4) for a linear record process started at $\theta(s_{n+1})$. Thus, $\mathbb{E}_\infty^{(1)}[\widehat{M}_{n+1} - \widehat{M}_n | \mathscr{G}_n] = 0$.

Convergence of the asymptotic variance

In order to get a convergence in distribution of $n^{-1/4}\widehat{M}_n$, we first need to compute the asymptotic variance of the martingale. This is done in the following proposition.

Proposition 4.8. *We have:*

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} \sum_{k=2}^{n} \mathbb{E}_\infty^{(1)}\left[\left(\widehat{M}_k - \widehat{M}_{k-1}\right)^2 \middle| \mathscr{G}_{k-1}\right] = \sqrt{2}\Theta, \tag{4.38}$$

in probability.

Proof. Using the martingale from (4.5), in the present case of a linear record process started at $\theta(s_k)$, we easily get that, for $k \geq 2$,

$$\mathbb{E}_\infty^{(1)}\left[\left(\widehat{M}_k - \widehat{M}_{k-1}\right)^2 \middle| \mathscr{G}_{k-1}\right] = \mathbb{E}_\infty^{(1)}\left[\int_{\mathrm{B}_k} \theta(s)\ \ell(ds) \middle| \mathscr{G}_{k-1}\right]. \tag{4.39}$$

A Law of Large Numbers argument will show that we have

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} \sum_{k=2}^{n} \mathbb{E}_\infty^{(1)}\left[\int_{\mathrm{B}_k} \theta(s)\ \ell(ds) \middle| \mathscr{G}_{k-1}\right] = \lim_{n\to\infty} \frac{1}{\sqrt{n}} \int_{\mathrm{T}_n^* \setminus [\![s_{\emptyset,n}, \mathrm{x}_1]\!]} \theta(s)\ \ell(ds). \tag{4.40}$$

We postpone the proof of this equality to the end of this section. Now, recall Proposition 6.3 in [AD11], which shows that a.s.

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} \int_{\mathrm{T}_n^*} \theta(s)\ \ell(ds) = \sqrt{2}\Theta.$$

Since $\mathrm{T}_n \setminus \mathrm{B}_1 = \mathrm{T}_n^* \setminus [\![s_{\emptyset,n}, \mathrm{x}_1]\!]$, the convergence (4.38) will follow if we manage to prove that

$$S_n = \frac{1}{\sqrt{n}} \int_{[\![s_{\emptyset,n}, \mathrm{x}_1]\!]} \theta(s)\ \ell(ds)$$

converges in probability to 0. We will simply compute the first moment:

$$\begin{aligned}\sqrt{n}\mathbb{E}_\infty^{(1)}[S_n] = \mathbb{E}_\infty^{(1)}\left[\int_{[\![s_{\emptyset,n}, \mathrm{x}_1]\!]} \theta(s)\ \ell(ds)\right] &= \mathbb{E}_\infty^{(1)}\left[\int_{h_{\emptyset,n}}^{L_1} \theta(s)\ ds\right] \\ &= \mathbb{E}_\infty^{(1)}\left[\int_0^{L_1 - h_{\emptyset,n}} \mathbf{E}_{\theta(s_{\emptyset,n})}[\theta(s)]\ ds\right],\end{aligned}$$

by the Markov property of θ at $h_{\emptyset,n}$. We can compute this expectation using (4.3):

$$\begin{aligned}&= \mathbb{E}_\infty^{(1)}\left[\int_0^{L_1 - h_{\emptyset,n}} \frac{1 - \mathrm{e}^{-s\theta(s_{\emptyset,n})}}{s}\ ds\right] \\ &\le \mathbb{E}_\infty^{(1)}\left[\int_0^{L_1} \frac{1}{s}(s\theta(s_{\emptyset,n}))^{1/4}\ ds\right] = 4\mathbb{E}_\infty^{(1)}\left[\theta(s_{\emptyset,n})^{1/4} L_1^{1/4}\right],\end{aligned}$$

by the elementary inequality $1 - \exp(-t) \le t^{1/4}$. The Cauchy-Schwarz inequality then gives the bound

$$\sqrt{n}\mathbb{E}_\infty^{(1)}[S_n] \le C \cdot \mathbb{E}_\infty^{(1)}\left[\theta(s_{\emptyset,n})^{1/2}\right]^{1/2}. \tag{4.41}$$

As $\theta(s_{\emptyset,n})$ is, conditionally on $\mathscr{T}$, exponentially distributed with parameter $h_{\emptyset,n}$, we get

$$\mathbb{E}_\infty^{(1)}[S_n] \le C \cdot n^{-1/2}\mathbb{E}_\infty^{(1)}[h_{\emptyset,n}^{-1/2}]^{1/2},$$

which converges to 0 as $n \to \infty$ by (4.9), which shows (4.38).

We still have to show (4.40) to end the proof. The process

$$\left(Q_n = \sum_{k=2}^n \int_{\mathrm{B}_k} \theta(s)\ell(ds) - \mathbb{E}_\infty^{(1)}\left[\int_{\mathrm{B}_k} \theta(s)\ell(ds)\middle|\mathscr{G}_{k-1}\right],\ n \ge 1\right) \tag{4.42}$$

is a $\mathscr{G}$-martingale. We will write

$$\langle Q\rangle_n = \sum_{k=1}^n \mathbb{E}_\infty^{(1)}\left[\left(\int_{\mathrm{B}_k} \theta(s)\ \ell(ds)\right)^2\middle|\mathscr{G}_{k-1}\right] - \mathbb{E}_\infty^{(1)}\left[\int_{\mathrm{B}_k} \theta(s)\ \ell(ds)\middle|\mathscr{G}_{k-1}\right]^2 \tag{4.43}$$

for its quadratic variation process. Conditionally on $\mathscr{G}_{k-1}$, the process $(\theta(s),\ s \in \mathrm{B}_k)$ is distributed as a linear record process started from $\theta(s_k)$. Hence, using (4.9) and (4.3), we get:

$$\mathbb{E}_\infty^{(1)}\left[\int_{\mathrm{B}_k} \theta(s)\ \ell(ds)\middle|\mathscr{G}_{k-1}\right] = \mathbf{E}_{\theta(s_k)}\left[\int_0^{\ell(\mathrm{B}_k)} \theta(s)\ ds\right] = \int_0^{\theta(s_k)\ell(\mathrm{B}_k)} \frac{1 - \mathrm{e}^{-u}}{u}\ du. \tag{4.44}$$

Similarly, we have:

$$\begin{aligned}\mathbb{E}_\infty^{(1)}\left[\left(\int_{\mathrm{B}_k} \theta(s)\ \ell(ds)\right)^2\middle|\mathscr{G}_{k-1}\right] &= \mathbf{E}_{\theta(s_k)}\left[\left(\int_0^{\ell(\mathrm{B}_k)} \theta(s)\ ds\right)^2\right] \\ &= 2 \cdot \mathbf{E}_{\theta(s_k)}\left[\int_0^{\ell(\mathrm{B}_k)} du \int_0^u dv\ \theta(u)\theta(v)\right].\end{aligned}$$

The latter can be computed by applying the Markov property at u, as well as (4.3), giving

$$\mathbb{E}_{\infty}^{(1)}\left[\left(\int_{\mathrm{B}_k}\theta(s)\ \ell(ds)\right)^2\Big|\mathcal{G}_{k-1}\right]=\frac{1}{\theta(s_k)}\int_0^{\theta(s_k)\ell(\mathrm{B}_k)}\frac{1-\mathrm{e}^{-s}}{s}-\mathrm{e}^{-s}\ ds$$
$$+2\int_0^{\theta(s_k)\ell(\mathrm{B}_k)}ds\int_0^s dt\frac{1}{s-t}\left(\frac{1-\mathrm{e}^{-t}}{t}-\frac{1-\mathrm{e}^{-s}}{s}\right). \quad (4.45)$$

Now, putting (4.44) and (4.45) together, compensations occur, so that we get, after tedious computations:

$$\begin{aligned}\langle Q\rangle_n &= \sum_{k=1}^n \mathbb{E}_{\infty}^{(1)}\left[\left(\int_{\mathrm{B}_k}\theta(s)\ \ell(ds)\right)^2\Big|\mathcal{G}_{k-1}\right]-\mathbb{E}_{\infty}^{(1)}\left[\int_{\mathrm{B}_k}\theta(s)\ \ell(ds)\Big|\mathcal{G}_{k-1}\right]^2\\ &= \sum_{k=1}^n\frac{2}{\theta(s_k)}\int_0^{\theta(s_k)\ell(\mathrm{B}_k)}\frac{1-\mathrm{e}^{-s}}{s}-\mathrm{e}^{-s}\ ds\\ &\quad +2\int_0^{\theta(s_k)\ell(\mathrm{B}_k)}ds\int_0^s dt\ \frac{s\mathrm{e}^{-s}-t\mathrm{e}^{-t}-(s-t)\mathrm{e}^{-(s+t)}}{st(s-t)}.\end{aligned}$$

The term $s\mathrm{e}^{-s}-t\mathrm{e}^{-t}-(s-t)\mathrm{e}^{-(s+t)}$ being negative for $t<s$, we get

$$\begin{aligned}0\le\langle Q\rangle_n &\le \sum_{k=1}^n\frac{2}{\theta(s_k)}\int_0^{\theta(s_k)\ell(\mathrm{B}_k)}\frac{1-\mathrm{e}^{-s}}{s}-\mathrm{e}^{-s}\ ds\\ &\le \sum_{k=1}^n\frac{2}{\theta(s_k)}\theta(s_k)\ell(\mathrm{B}_k)=2\sum_{k=1}^n\ell(\mathrm{B}_k),\end{aligned}$$

the second inequality coming from $(1-\mathrm{e}^{-s})/s-\mathrm{e}^{-s}\le 1$ if $s>0$. Then, recall that by definition, $\sum_{k=1}^n\ell(\mathrm{B}_k)\le L_n$, and that L_n is the square root of a $Gamma(n,1)$-distributed variable (Proposition 5.2 in [AD11]). Thus, for any $\gamma>1/2$, we have

$$\frac{1}{n^\gamma}\mathbb{E}_{\infty}^{(1)}[\langle Q\rangle_n]\le\frac{2}{n^\gamma}\mathbb{E}_{\infty}^{(1)}[L_n]\to 0 \quad (4.46)$$

Then, by the conditional Law of Large Numbers (Theorem 1.3.17 in [Duf97]), we get that $n^{-1/4-\varepsilon}Q_n$ converges a.s. to 0 for any $\varepsilon>0$, which implies (4.40), hence ends the proof. □

Asymptotic smallness

We now turn to the proof of the asymptotic smallness of the sequence $(\widehat{M}_n,\ n\ge 1)$. In order to prove this, we will use a Liapounov-type criterion, which is sufficient to prove asymptotic negligibility.

Proposition 4.9. *We have the following convergence in probability:*

$$\lim_{n\to\infty}\frac{1}{\sqrt{n}}\sum_{k=1}^n\mathbb{E}_{\infty}^{(1)}\left[(\widehat{M}_k-\widehat{M}_{k-1})^2\mathbf{1}_{\{|\widehat{M}_k-\widehat{M}_{k-1}|>\varepsilon n^{1/4}\}}\Big|\mathcal{G}_{k-1}\right]=0.$$

Proof. We use the standard inequality $\mathbf{1}_{\{|\widehat{M}_k - \widehat{M}_{k-1}| > \varepsilon n^{1/4}\}} \leq (\widehat{M}_k - \widehat{M}_{k-1})^2/\varepsilon^2\sqrt{n}$ to get that, for $\varepsilon > 0$:

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} \mathbb{E}_{\infty}^{(1)} \left[\left(\widehat{M}_k - \widehat{M}_{k-1}\right)^2 \mathbf{1}_{\{|\widehat{M}_k - \widehat{M}_{k-1}| > \varepsilon n^{1/4}\}} \middle| \mathcal{G}_{k-1} \right] \leq \frac{1}{\varepsilon^2 n} \sum_{k=1}^{n} \mathbb{E}_{\infty}^{(1)} \left[\left(\widehat{M}_k - \widehat{M}_{k-1}\right)^4 \middle| \mathcal{G}_{k-1} \right].$$

Using the martingale from (4.6), we find that:

$$\frac{1}{\varepsilon^2 n} \sum_{k=2}^{n} \mathbb{E}_{\infty}^{(1)} \left[\left(\widehat{M}_k - \widehat{M}_{k-1}\right)^4 \middle| \mathcal{G}_{k-1} \right] = \frac{3}{\varepsilon^2 n} \sum_{k=2}^{n} \mathbb{E}_{\infty}^{(1)} \left[\left(\int_{\mathrm{B}_k} \theta(s)\ \ell(ds) \right)^2 \middle| \mathcal{G}_{k-1} \right]$$
$$+ \frac{1}{\varepsilon^2 n} \sum_{k=2}^{n} \mathbb{E}_{\infty}^{(1)} \left[\int_{\mathrm{B}_k} \theta(s)\ \ell(ds) \middle| \mathcal{G}_{k-1} \right].$$

In this expression, the term $n^{-1} \sum_{k=2}^{n} \mathbb{E}[\int_{\mathrm{B}_k} \theta(s)\ell(ds) | \mathcal{G}_{k-1}]$ converges in probability to 0, according to (4.39) and Proposition 4.8. Furthermore, recall from (4.43) that

$$\frac{3}{\varepsilon^2 n} \sum_{k=1}^{n} \mathbb{E}_{\infty}^{(1)} \left[\left(\int_{\mathrm{B}_k} \theta(s)\ \ell(ds) \right)^2 \middle| \mathcal{G}_{k-1} \right] = \frac{3\langle Q \rangle_n}{\varepsilon^2 n} + \frac{3}{\varepsilon^2 n} \sum_{k=1}^{n} \mathbb{E}_{\infty}^{(1)} \left[\int_{\mathrm{B}_k} \theta(s)\ \ell(ds) \middle| \mathcal{G}_{k-1} \right]^2,$$

where Q is the martingale defined in (4.42). The quadratic variation process $\langle Q \rangle_n / n$ converges in probability to 0 by (4.46). Also, applying Lemma 4.13 to $a_k = \mathbb{E}_{\infty}^{(1)}[\int_{\mathrm{B}_k} \theta(s)\ell(ds) | \mathcal{G}_{k-1}]$, we find that

$$\frac{1}{n} \sum_{k=1}^{n} \mathbb{E}_{\infty}^{(1)} \left[\int_{\mathrm{B}_k} \theta(s)\ \ell(ds) \middle| \mathcal{G}_{k-1} \right]^2 = 0,$$

which ends the proof. □

Putting all the previous elements together, we can now prove Theorem 4.1.

Proof of Theorem 4.1. First, we write that

$$n^{1/4} \left(\frac{X_n^*}{\sqrt{2n}} - \Theta \right) = \frac{\widehat{M}_n}{\sqrt{2} n^{1/4}} + \frac{M_n^* - \widehat{M}_n}{\sqrt{2} n^{1/4}} + n^{1/4} \left(\frac{1}{\sqrt{2n}} \int_{\mathrm{T}_n^*} \theta(s)\ \ell(ds) - \Theta \right).$$

The convergence in distribution of $n^{-1/4}\widehat{M}_n$ towards a non-degenerate limit Z is a consequence of the Martingale Central Limit Theorem recalled at the beginning of this section with $a_n = n^{1/4}$, as well as the two Propositions 4.8 and 4.9. Furthermore, the limiting random variable Z is indeed distributed as announced:

$$\mathbb{E}_{\infty}^{(1)} \left[\mathrm{e}^{itZ} \right] = \mathbb{E}_{\infty}^{(1)} \left[\mathrm{e}^{-t^2 \sqrt{2}\Theta/2} \right].$$

The term $e_n = M_n^* - \widehat{M}_n$ can be expressed as

$$e_n = M_n^* - \widehat{M}_n = \sum_{s \in [\![s_{\emptyset,n}, \mathrm{x}_1]\!]} \mathbf{1}_{\{\theta(s-) > \theta(s)\}} - \int_{[\![s_{\emptyset,n}, \mathrm{x}_1]\!]} \theta(s)\ \ell(ds).$$

Using the martingale (4.5) to compute its second moment, we get

$$\mathbb{E}_{\infty}^{(1)}\left[e_n^2\right] = \mathbb{E}_{\infty}^{(1)}\left[\int_{[\![s_{\emptyset,n}, \mathrm{x}_1]\!]} \theta(s)\ \ell(ds)\right],$$

so that $n^{-1/4}(M_n^* - \widehat{M}_n)$ converges to 0 in L^2, hence in distribution as $n \to \infty$, by the previously used bound (4.41). Finally, Proposition 4.3 and Proposition 4.4 show that the term $((2n)^{-1/2}\int_{\mathrm{T}_n^*}\theta(s)\ \ell(ds) - \Theta)$ brings no contribution in the asymptotic $n^{1/4}$. This ends the proof. □

Remark 11. *Note that, under our assumptions, since $\Theta > 0$, $\mathbb{P}_{\infty}$-a.s., we can actually prove that the convergence in distribution of $n^{-1/4}\widehat{M}_n$ is* mixing *(see [AE78] for more details on mixing limit theorems). This implies in particular that we can obtain a standard normal limit by renormalizing by the random factor V_n, where V_n^2 is the conditional variance*

$$V_n^2 = \sum_{k=1}^{n} \mathbb{E}_{\infty}^{(1)}\left[\left(\widehat{M}_k - \widehat{M}_{k-1}\right)^2 \middle| \mathscr{G}_{k-1}\right],$$

instead of the deterministic renormalization $n^{1/4}$. Corollary 3.2 in [HH80] then shows that $V_n^{-1}\widehat{M}_n$ converges in distribution to a standard $\mathcal{N}(0,1)$ random variable.

Technical appendix

In this appendix, we shall state and prove several lemmas that are used throughout the paper. They are purely analytic in nature, and their proof is elementary, so we gather them here, for the reader's convenience. First, we prove some universal bounds on $F(q,t) = \mathbf{E}_q[(\int_0^t \theta(s)ds)^2]$.

Lemma 4.10. *There exists $C_1, C_2, C_3, C_4 > 0$ such that*

$$F(q,t) \le C_1(qt)^{3/2} + C_2 q t^2 \tag{4.47}$$

$$F(q,t) \le C_3 \log^2(qt) + C_4 q^{-1/2} t^{1/2} \tag{4.48}$$

Proof. First, we recall that, according to (4.45),

$$\begin{aligned} F(q,t) &= \mathbf{E}_q\left[\left(\int_0^t \theta(s)\ ds\right)^2\right] \\ &= \frac{1}{q}\int_0^{qt} \frac{1-\mathrm{e}^{-s}}{s} - \mathrm{e}^{-s}\ ds + \int_0^{qt} ds \int_0^s dt\ \frac{1}{s-t}\left(\frac{1-\mathrm{e}^{-t}}{t} - \frac{1-\mathrm{e}^{-s}}{s}\right) \\ &:= \tilde{F}(q,t) + G(ql). \end{aligned}$$

The two estimates (4.47) and (4.48) will come from an asymptotic analysis of

$$\tilde{F}(q,t) = \frac{1}{q}\int_0^{qt} \frac{1-\mathrm{e}^{-s}}{s} - \mathrm{e}^{-s}\ ds$$

and

$$G(qt) = \int_0^{qt} ds \int_0^s dt \, \frac{1}{s-t} \left(\frac{1-\mathrm{e}^{-t}}{t} - \frac{1-\mathrm{e}^{-s}}{s} \right).$$

Let us start with $\tilde{F}$. We have

$$\begin{aligned}
\tilde{F}(q,t) &= \frac{1}{q} \int_0^{qt} \frac{1-\mathrm{e}^{-s}}{s} - \mathrm{e}^{-s} \, ds \\
&= \frac{1}{q} \left(\gamma + \log(qt) + \int_{qt}^{\infty} \frac{\mathrm{e}^{-t}}{t} \, dt + \mathrm{e}^{-qt} - 1 \right).
\end{aligned}$$

It is elementary to check that the function $\gamma + \log(x) + \int_x^\infty \frac{\mathrm{e}^{-t}}{t} \, dt + \mathrm{e}^{-x} - 1$ is equivalent to $x^2/4$ when $x \to 0$, and equivalent to $\log(x) = o(\sqrt{x})$ when $x \to \infty$. Since $\sqrt{x} = o(x^2)$ in the neighbourhood of $+\infty$ and $x^2 = o(\sqrt{x})$ in the neighbourhood of 0, by continuity, we can find constants C_2 and C_4 such that $\tilde{F}(q,t) \le C_2 (qt)^2/q$ and such that $\tilde{F}(qt) \le C_4 (qt)^{1/2}/q$.

Turning to the function G, we can write

$$\begin{aligned}
G(x) &= \int_0^x ds \int_0^s dt \frac{1}{s-t} \left(\frac{1-\mathrm{e}^{-t}}{t} - \frac{1-\mathrm{e}^{-s}}{s} \right) \\
&= \int_0^1 du \int_0^u dv \frac{1}{u-v} \left(\frac{1-\mathrm{e}^{-xv}}{v} - \frac{1-\mathrm{e}^{-xu}}{u} \right),
\end{aligned}$$

so that

$$G'(x) = \int_0^1 du \int_0^u dv \frac{1}{u-v} (\mathrm{e}^{-xv} - \mathrm{e}^{-xu}),$$

and that

$$G''(x) = \int_0^1 du \int_0^u dv \frac{1}{u-v} (u\mathrm{e}^{-xu} - v\mathrm{e}^{-xv}).$$

Thus, we have $G(0) = G'(0) = 0$ and $G''(0) = 1$. Since G is smooth, we get that $G(x) \sim x^2/2$ when $x \to 0$.

As far as the asymptotic $x \to \infty$ is concerned, we can express $G'(x)$ in terms of the exponential integral[1] function $Ei(x) = \int_{-\infty}^x \exp(t)/t \, dt$:

$$\begin{aligned}
G'(x) &= \int_0^1 du \int_0^u \frac{dv}{u-v} (\mathrm{e}^{-xv} - \mathrm{e}^{-xu}) \\
&= \int_0^1 du \mathrm{e}^{-xu} \int_0^{xu} \frac{dv}{v} (\mathrm{e}^{v} - 1) \\
&= \int_0^1 du \mathrm{e}^{-xu} (Ei(xu) - \log(xu) - \gamma).
\end{aligned}$$

[1] Note that this integral is to be taken in the sense of Cauchy's principal value.

When $x \to \infty$, we get

$$\begin{aligned} G'(x) &\sim \int_0^1 du \mathrm{e}^{-xu} Ei(xu) = \frac{1}{x} \int_0^x du \mathrm{e}^{-t} Ei(t) \\ &\sim \frac{\log x}{x}. \end{aligned}$$

Integrating from 0 to x, we get $G(x) \sim \log^2 x = o(\sqrt{x})$ when $x \to \infty$. Again, $\sqrt{(}x) = o(x^2)$ in the neighbourhood of $+\infty$ and $x^2 = o(\sqrt{x})$ in the neighbourhood of 0, so that by continuity, there exist two constants C_1 and C_2 such that $G(x) \le C_1 x^2$ and such that $G(x) \le C_2 x^{1/2}$. Thus, we get the two dominations (4.47) and (4.48). ☐

We now turn to a useful estimation of the moments of the distribution $r_a(dx)$ introduced in (4.30):

$$r_a(dx) = (a+x)\mathrm{e}^{-x^2/2-ax} \mathbf{1}_{(0,\infty)}(x)\ dx.$$

Lemma 4.11. *Let $\lambda > 0$. Then, if $(a(n),\ n \ge 1)$ is some sequence in $\mathbf{R}_+$ increasing to $+\infty$, then, as $n \to \infty$, we have $\int_0^\infty r_{a(n)}(dx) x^\lambda = O(a(n)^{-\lambda})$.*

Proof. This is fairly easy: if $\lambda > 0$, we can write

$$\begin{aligned} \int_0^\infty r_{a(n)}(dx) x^\lambda &= \int_0^\infty x^\lambda (a(n)+x) \mathrm{e}^{-x^2/2-a(n)x}\ dx \\ &= \int_0^\infty \frac{u^\lambda}{a(n)^\lambda} \left(a(n) + \frac{u}{a(n)} \right) \mathrm{e}^{-u^2/(2a(n)^2)-u} \frac{du}{a(n)} \\ &\le \frac{1}{a(n)^\lambda} \int_0^\infty u^\lambda \mathrm{e}^{-u}\ du + \frac{1}{a(n)^{\lambda+2}} \int_0^\infty u^{\lambda+1} \mathrm{e}^{-u}\ du, \end{aligned}$$

which ends the proof. ☐

Lemma 4.12. *For any $0 < q < \infty$ and any $v \ge 0$, we have*

$$\mathbb{E}_q^{(v)}[\Theta] \le \sqrt{\pi/2} \min(qv, \sqrt{v}).$$

Proof. We will use formula (21) from [AD11], stating that, in our context, if Y is a Rayleigh-distributed variable, then

$$\mathbb{E}_q^{(v)}[\Theta] = \sqrt{v} \int_0^{q\sqrt{v}} \mathbb{E}\left[\mathrm{e}^{-tY}\right]\ dt.$$

We simply expand the Laplace transform, giving

$$\begin{aligned} \mathbb{E}_q^{(v)}[\Theta] &= \sqrt{v} \int_0^{q\sqrt{v}} \int_0^\infty x \mathrm{e}^{-x^2/2} \mathrm{e}^{-tx}\ dx\ dt \\ &= \sqrt{v} \int_0^\infty \mathrm{e}^{-x^2/2} \left(1 - \mathrm{e}^{-xq\sqrt{v}}\right)\ dx. \end{aligned}$$

Now, we use the obvious inequality $1 - \exp(-x) \le \min(x, 1)$, to get the desired domination, since $qv \int_0^\infty x \exp(-x^2/2) = qv$ and $\sqrt{v} \int_0^\infty \mathrm{e}^{-x^2/2} dx = \sqrt{\pi v/2}$. ☐

Finally, the next lemma is needed to prove the asymptotic smallness of the martingale $\widehat{M}_n$.

Lemma 4.13. *Let* $(a_n, n \geq 1)$ *be a nonnegative sequence such that*

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} \sum_{k=1}^{n} a_k < \infty.$$

Then, we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} a_k^2 = 0.$$

Proof. Let $s_n = n^{-1/2} \sum_{k=1}^{n} a_k$. Taking the difference $s_n - s_{n-1}$, we easily see that $n^{-1/2} a_n$ converges to 0. Then, if $\varepsilon > 0$, there exists $n_0 \geq 1$ such that for all $n \geq n_0$, $a_n < \varepsilon\sqrt{n}$. Thus, if $n \geq n_0$, we have

$$\begin{aligned}\sup_{k\leq n} a_k &\leq \sup_{k<n_0} a_k + \sup_{n_0\leq k\leq n} a_k \\ &\leq \sup_{k<n_0} a_k + \varepsilon,\end{aligned}$$

which proves that actually

$$\lim_{n\to\infty} \frac{\sup_{k\leq n} a_k}{\sqrt{n}} = 0.$$

Then, we simply write

$$\frac{1}{n} \sum_{k=1}^{n} a_k^2 \leq \left(\frac{\sup_{k\leq n} a_k}{\sqrt{n}}\right)\left(\frac{1}{\sqrt{n}} \sum_{k=1}^{n} a_k\right)$$

to conclude. □

APPENDIX A

Generalization of the special Markov Property of the pruning process

In this chapter, we will give a proof of the special Markov property stated in Theorem 3.25, in the subcritical case, when taking into account the pruning times. The proofs are adapted from the proofs in [ADV10]. To state and prove the theorem, we will use the formalism of the exploration process, which is a measure-valued Markov process describing the depth-first exploration of a Lévy tree. For more precise definitions, see [DL02]. Let us first rephrase the definitions of Section 3.2 in this context.

We will consider, for some (sub)critical branching mechanism ψ satisfying Assumption 1 the marked exploration process $((\rho_t, m_t),\ t \geq 0)$, where, for any $t \geq 0$, conditionally on ρ_t, m_t has two components, m_t^{ske} and m_t^{nod}, representing marks on the skeleton and marks on nodes respectively, which are such that

- m_t^{ske} is a σ-finite measure on $\mathbf{R}_+ \times \mathbf{R}_+$ such that, for any $\theta > 0$, $m_t^{\text{ske}}(\cdot \times [0,\theta])$ has support included in $\text{Supp}(\rho_t)$.
- m_t^{nod} is a σ-finite measure on $\mathbf{R}_+ \times \mathbf{R}_+$ such that, for any $\theta > 0$, $m_t^{\text{nod}}(\cdot \times [0,\theta])$ is absolutely continuous with respect to ρ_t.

We will denote by $\mathbb{S}$ the state-space of the marked exploration process. More precise definitions can be found in ([Voi10]) where the measures are defined using a Lévy snake with Poisson paths. Now, for some fixed $\theta > 0$, we will consider the process (ρ_t, m_t^θ), where

$$m_t^\theta(dx) = \left(m_t^{\text{ske}}(dx, [0,\theta]), m_t^{\text{nod}}(dx, [0,\theta])\right),$$

thus considering only the marks appearing between time 0 and time θ. If O^θ is the interior of the set $\{s \geq 0,\ m_s^\theta \neq 0\}$, let us write

$$O^\theta = \bigcup_{i \in \mathscr{I}} \left(\alpha_i^\theta, \beta_i^\theta\right)$$

and say that the intervals $(\alpha_i^\theta, \beta_i^\theta)_{i\in\mathscr{I}}$ are the excursions intervals of the marked exploration process $\mathscr{S}^\theta = (\rho, m^\theta)$ away from $\{s \geq 0,\ m_s^\theta = 0\}$. We also define the following continuous additive functional of the process $((\rho_t, m_t^\theta), t \geq 0)$:

$$A_t = \int_0^t \mathbf{1}_{\{m_s^\theta = 0\}}\, ds \quad \text{for } t \geq 0. \tag{A.1}$$

We set $C_t^\theta = \inf\{r > 0,\ A_r > t\}$, the right-continuous inverse of A, with the convention that $\inf \emptyset = \infty$. For every $i \in \mathscr{I}$, let us define the measure-valued process $\mathscr{S}^{i,\theta} = (\rho^{i,\theta}, m^{i,\theta})$. For every $f \in \mathrm{B}_+(\mathbf{R}_+)$, $t \geq 0$, we set, if $H_t = \langle \rho_t, 1\rangle$,

$$\begin{aligned} \left\langle \rho_t^{i,\theta}, f \right\rangle &= \int_{[H_{\alpha_i^\theta}, +\infty)} f(x - H_{\alpha_i^\theta}) \rho_{(\alpha_i^\theta + t)\wedge\beta_i^\theta}(dx) \\ \left\langle (m^{\mathrm{a}})_t^{i,\theta}, g \right\rangle &= \int_{(H_{\alpha_i^\theta}, +\infty)\times \mathbf{R}_+} f(x - H_{\alpha_i^\theta}, \theta') m^{\mathrm{a}}_{(\alpha_i^\theta + t)\wedge\beta_i^\theta}(dx, d\theta') \quad \text{with a } \in \{\mathrm{nod}, \mathrm{ske}\} \end{aligned} \tag{A.2}$$

and $m_t^{i,\theta} = ((m^{\mathrm{nod}})_t^{i,\theta}, (m^{\mathrm{ske}})_t^{i,\theta})$. Furthermore, if $i \in \mathscr{I}$, we define θ_i by:

$$\theta_i = \inf\left\{\theta \in \mathbf{R}_+, m^\theta_{\alpha_i^\theta}\left([0, H_{\alpha_i^\theta}] \times [0,\theta]\right) \neq 0\right\}. \tag{A.3}$$

Let $\mathscr{F}_\infty^\theta$ be the σ-field generated by $\mathscr{S}^\theta = ((\rho_{C_t^\theta}, m_{C_t^\theta}), t \geq 0)$. We will use the notation $\mathbb{P}^*_{\mu,\Pi}(d\mathscr{S})$ to denote the law of the marked exploration process $\mathscr{S}$ started at $(\mu, \Pi) \in \mathbb{S}$ and stopped when ρ reaches 0. For $\ell \in (0, +\infty)$, we will write $\mathbb{P}^*_\ell$ for $\mathbb{P}^*_{\ell\delta_0, 0}$.

Let us now state the Special Markov Property (Theorem 3.25):

Theorem 1 (Special Markov property). *Let ϕ be a non-negative measurable function defined on $\mathbf{R}_+ \times \mathscr{M}_f(\mathbf{R}_+) \times \mathbb{S} \times \mathbf{R}_+$. Then, we have $\mathbb{P}$-a.s.*

$$\begin{aligned} \mathbb{E}&\left[\exp\left(-\sum_{i\in\mathscr{I}} \phi(A_{\alpha_i^\theta}, \rho_{\alpha_i^{\theta-}}, \mathscr{S}^i, \theta^i)\right) \Big| \mathscr{F}_\infty^\theta\right] \\ &= \exp\left(-\int_0^\theta d\theta' \int_0^\infty du\, 2\beta\mathbb{N}\left[1 - \mathrm{e}^{-\phi(u,\mu,\cdot,\theta')}\right]_{|\mu=\rho_u^\theta}\right) \\ &\exp\left(-\int_0^\theta d\theta' \int_0^\infty du \int_{(0,\infty)} \ell \mathrm{e}^{-\theta'\ell}\pi(d\ell)\left(1 - \mathbb{E}^*_\ell[\mathrm{e}^{-\phi(u,\mu,\cdot,\theta')}]_{|\mu=\rho_u^\theta}\right)\right). \end{aligned} \tag{A.4}$$

In other words, the law under $\mathbb{P}$ of the excursion process $\sum_{i\in\mathscr{I}} \delta_{(A_{\alpha_i}, \rho_{\alpha_i -}, \mathscr{S}^i, \theta^i)}(du, d\mu, d\mathscr{S}, d\theta')$, given $\mathscr{F}_\infty^\theta$, is the law of a Poisson point measure with intensity

$$\mathbf{1}_{\{u\geq 0\}} du\, \mathbf{1}_{\{\theta' \in [0,\theta)\}} d\theta' \delta_{\tilde{\rho}_u}(d\mu) \left(2\beta\mathbb{N}(d\mathscr{S}) + \int_{(0,\infty)} \ell \mathrm{e}^{-\theta'\ell}\pi(d\ell)\mathbb{P}^*_\ell(d\mathscr{S})\right).$$

Remark 12. *The special Markov property proven here is nothing else but a disintegrated version of the special Markov property in [ADV10]. Indeed, when integrating over $[0,\theta]$, we get exactly the formulation from Theorem 16, in which we don't have access to the pruning times. However, this information is needed to prove (Chapter 3) that the recursively defined tree-growth process has the same distribution as the time-returned pruning process.*

We will sometimes need the $\mathcal{M}_f(\mathbf{R}_+)$-valued process $\eta = (\eta_t, t \geq 0)$ defined by

$$\eta_t(dr) = \beta \mathbf{1}_{[0,H_t]}(r)\, dr + \sum_{\substack{0<s\leq t \\ X_{s-}<I_t^s}} (X_s - I_t^s)\delta_{H_s}(dr).$$

The process η is the dual process of ρ under $\mathbb{N}$ (see Corollary 3.1.6 in [DL02]). Let the measures $\mathcal{N}_0(dx, d\ell, du)$, $\mathcal{N}_1(dx, d\ell, du, d\theta')$ and $\mathcal{N}_2(dx, d\theta')$ be independent Poisson point measures respectively on $[0,+\infty)^3$, $[0,+\infty)^4$ and $[0,+\infty)^2$ with respective intensity

$$dx\, \ell \pi(d\ell)\, \mathbf{1}_{[0,1]}(u) du, \quad dx\, \ell \mathrm{e}^{-\theta'\ell} \pi(d\ell)\, \mathbf{1}_{[0,1]}(u) du d\theta' \quad \text{and} \quad 2\beta dx d\theta.$$

For every $a > 0$, let us denote by $\mathbb{M}_a$ the law of the measures $(\mu, \nu, m^{\mathrm{nod}}, m^{\mathrm{ske}})$ defined by: for any $f \in \mathcal{B}_+(\mathbf{R}_+)$, $g \in \mathcal{B}_+(\mathbf{R}_+ \times \mathbf{R}_+)$,

$$\begin{aligned}
\langle \mu, f\rangle &= \int \big(\mathcal{N}_0(dx, d\ell, du) + \mathcal{N}_1(dx, d\ell, du, d\theta')\big)\, \mathbf{1}_{[0,a]}(x) u\ell f(x) + \beta \int_0^a f(r)\, dr,\\
\langle \nu, f\rangle &= \int \big(\mathcal{N}_0(dx, d\ell, du) + \mathcal{N}_1(dx, d\ell, du, d\theta')\big)\, \mathbf{1}_{[0,a]}(x) (1-u)\ell f(x) + \beta \int_0^a f(r)\, dr,\\
\langle m^{\mathrm{nod}}, g\rangle &= \int \mathcal{N}_1(dx, d\ell, du, d\theta')\, \mathbf{1}_{[0,a]}(x) u\ell g(x,\theta')\\
\langle m^{\mathrm{ske}}, g\rangle &= \int \mathcal{N}_2(dx, d\theta')\, \mathbf{1}_{[0,a]}(x) g(x,\theta').
\end{aligned}$$

We finally set $\mathbb{M} = \int_0^{+\infty} da\, \mathrm{e}^{-\alpha a} \mathbb{M}_a$. Using the construction of the snake, it is easy to deduce from Proposition 3.1.3 in [DL02], the following Poisson representation.

Proposition. *For every non-negative measurable function F on $\mathcal{M}_f(\mathbf{R}_+)^4$,*

$$\mathbb{N}\left[\int_0^\sigma F(\rho_t, \eta_t, m_t)\, dt\right] = \int \mathbb{M}(d\mu\, d\nu\, dm) F(\mu, \nu, m),$$

where $m = (m^{\mathrm{nod}}, m^{\mathrm{ske}})$ and $\sigma = \inf\{s > 0; \rho_s = 0\}$ denotes the length of the excursion.

Preliminaries

Fix $t > 0$ and $\eta > 0$. For $\mathcal{S} = (\mathcal{S}_s = (\rho_s, m_s), s \geq 0)$, we set

$$B = \{\sigma(\mathcal{S}) = +\infty\} \cup \left\{T_\eta(\mathcal{S}) = +\infty\right\} \cup \left\{L_\eta(\mathcal{S}) = -\infty\right\},$$

where $\sigma(\mathcal{S}) = \inf\{s > 0; \rho_s = 0\}$ is the length of the current excursion above 0, where $T_\eta(\mathcal{S}) = \inf\{s \geq 0; \langle \rho_s, 1\rangle \geq \eta\}$ and $L_\eta(\mathcal{S}) = \sup\{s \in [0, \sigma(\mathcal{S})]; \langle \eta_s, 1\rangle \geq \eta\}$, with the convention $\inf \emptyset = +\infty$ and $\sup \emptyset = -\infty$. We consider non-negative bounded functions ϕ satisfying the assumptions of the Special Markov Property and the four following conditions:

- (h_1) $\phi(u, \mu, \mathcal{S}, \theta) = 0$ for any $u \geq t$.
- (h_2) $(u, \theta) \mapsto \phi(u, \mu, \mathcal{S}, \theta)$ is uniformly Lipschitz (with a constant that does not depend on μ and $\mathcal{S}$).

(h_3) $\phi(u,\mu,\mathscr{S},\theta) = 0$ on B; and if $\mathscr{S} \in B^c$ then $\phi(u,\mu,\mathscr{S},\theta)$ depends on $\mathscr{S}$ only through $(\mathscr{S}_u, u \in [T_\eta, L_\eta])$.

(h_4) The function $\mu \mapsto \phi(u,\mu,\mathscr{S},\theta)$ is continuous with respect to the metric

$$D(\mu,\mu') = d_{\mathrm{Pr}}(\mu,\mu') + |\langle \mu, 1\rangle - \langle \mu', 1\rangle|$$

on $\mathscr{M}_f(\mathbf{R}_+)$, where d_{Pr} is the Prokhorov metric on $\mathscr{M}_f(\mathbf{R}_+)$ (recall that d_{Pr} metrizes weak convergence in $\mathscr{M}_f(\mathbf{R}_+)$).

Lemma 2. *Let ϕ be a function satisfying $(h_1 - h_3)$ and let w be defined on $(0,\infty) \times [0,\infty) \times \mathscr{M}_f(\mathbf{R}_+) \times \mathbf{R}_+$ by*

$$w(\ell,u,\mu,\theta) = \mathbb{E}^*_\ell \left[\mathrm{e}^{-\phi(u,\mu,\cdot,\theta)} \right].$$

Then, for $\mathbb{N}$-a.e. $\mu \in \mathscr{M}_f(\mathbf{R}_+)$, the function $(\ell,u,\theta) \mapsto w(\ell,u,\mu,\theta)$ is uniformly continuous on $(0,\infty) \times [0,\infty) \times \mathbf{R}_+$.

The proof of this lemma is exactly the same as in [ADV10], as the stronger hypothesis (h_2) we made on ϕ enables us to get uniform continuity in the θ variable as well.

Stopping times

Let $R(dt,du)$ be a Poisson point measure on $\mathbf{R}_+^2$ (defined on $(\mathbb{S},\mathscr{F})$) independent of $\mathscr{F}_\infty$ with intensity the Lebesgue measure. We denote by $\mathscr{G}_t$ the σ-field generated by $R(\cdot \cap [0,t] \times \mathbf{R}_+)$. For every $\varepsilon > 0$, the process $R^\varepsilon_t := R([0,t] \times [0,1/\varepsilon])$ is a Poisson process with intensity $1/\varepsilon$. We denote by $(e^\varepsilon_k, k \geq 1)$ the time intervals between the jumps of $(R^\varepsilon_t, t \geq 0)$. The random variables $(e^\varepsilon_k, k \geq 1)$ are i.i.d. exponential random variables with mean ε, and are independent of $\mathscr{F}_\infty$. They define a mesh of $\mathbf{R}_+$ which is finer and finer as ε decreases to 0.

For $\varepsilon > 0$, we consider $T^\varepsilon_0 = 0$, $M^\varepsilon_0 = 0$ and for $k \geq 0$,

$$\begin{aligned}
M^\varepsilon_{k+1} &= \inf\left\{ i > M^\varepsilon_k;\ m^\theta_{T^\varepsilon_k + \sum_{j=M^\varepsilon_k+1}^{i} e^\varepsilon_j} \neq 0 \right\}, \\
S^\varepsilon_{k+1} &= T^\varepsilon_k + \sum_{j=M^\varepsilon_k+1}^{M^\varepsilon_{k+1}} e^\varepsilon_j, \\
T^\varepsilon_{k+1} &= \inf\left\{ s > S^\varepsilon_{k+1};\ m^\theta_s = 0 \right\},
\end{aligned} \tag{A.5}$$

with the convention $\inf \emptyset = +\infty$. For every $t \geq 0$, we set $\tau^\varepsilon_t = \int_0^t ds\, \mathbf{1}_{\bigcup_{k\geq 1} [T^\varepsilon_k, S^\varepsilon_{k+1})}(s)$ and

$$\mathscr{F}^e_t = \sigma(\mathscr{F}_t \cup \mathscr{G}_{\tau^\varepsilon_t}). \tag{A.6}$$

Notice that T^ε_k and S^ε_k are $\mathscr{F}^e$-stopping times.

Now we introduce a notation for the process defined above the marks on the intervals $[S^\varepsilon_k, T^\varepsilon_k]$. We set, for $a \geq 0$, $\bar{H}_a$ the level of the first mark, ρ^-_a the restriction of ρ_a strictly below it, ρ^+_a the restriction of ρ_a above it and θ_a the index of the first mark:

$$\bar{H}_a = \sup\{t > 0, m_a([0,t]) = 0\}, \quad \rho^-_a = \rho_a(\cdot \cap [0,\bar{H}_a)) \tag{A.7}$$

and ρ_a^+ is defined by $\rho_a = [\rho_a^-, \rho_a^+]$, that is for any $f \in \mathrm{B}_+(\mathbf{R}_+)$,

$$\langle \rho_a^+, f\rangle = \int_{[\bar{H}_a,\infty)} f(r-\bar{H}_a)\,\rho_a(dr). \tag{A.8}$$

$$\theta_a = \inf\{\theta' \in [0,\theta],\ m_a([0,H_a]\times[0,\theta']) \neq 0\} \tag{A.9}$$

with the convention $\theta_a = \infty$ if there is no mark present at time a, that is, when $\rho_a^+ = 0$.

For $k \geq 1$ and $\varepsilon > 0$ fixed, we define $\mathscr{S}^{k,\varepsilon} = \left(\rho^{k,\varepsilon}, m^{k,\varepsilon}\right)$ in the following way: for $s > 0$ and $f \in \mathrm{B}_+(\mathbf{R}_+)$

$$\begin{aligned} \rho_s^{k,\varepsilon} &= \rho^+_{(S_k^\varepsilon+s)\wedge T_k^\varepsilon}, \\ \left\langle (m^{\mathrm{a}})_s^{k,\varepsilon}, f\right\rangle &= \int_{(\bar{H}_{S_k^\varepsilon},+\infty)} f(r-\bar{H}_{S_k^\varepsilon})\, m^{\mathrm{a}}_{(S_k^\varepsilon+s)\wedge T_k^\varepsilon}(dr), \quad \text{with a } \in \{\mathrm{nod}, \mathrm{ske}\}, \end{aligned}$$

and $m_s^{k,\varepsilon} = ((m^{\mathrm{nod}})_s^{k,\varepsilon}, (m^{\mathrm{ske}})_s^{k,\varepsilon})$. Notice that $\rho_s^{k,\varepsilon}(\{0\}) = \rho_{S_k^\varepsilon}(\{\bar{H}_{S_k^\varepsilon}\})$.

Finally, we define $\theta^{k,\varepsilon} = \theta_{S_k^\varepsilon}$ which is the index of the lowest mark present at time S_k^ε. For $k \geq 1$, we consider the σ-field $\mathscr{F}^{(\varepsilon),k}$ generated by the processes $(\mathscr{S}_{(T_\ell^\varepsilon+s)\wedge S_{\ell+1}^\varepsilon-},\ s > 0)$ for $\ell \in \{0,\dots,k-1\}$ Notice that for $k \in \mathbf{N}^*$

$$\mathscr{F}^{(\varepsilon),k} \subset \mathscr{F}^e_{S_k^\varepsilon}. \tag{A.10}$$

Approximation of the functional

Remember we want to compute the conditional distribution of $\sum_{i\in\mathscr{I}} \phi(A_{\alpha_i^\theta}, \rho_{\alpha_i^\theta-}, \mathscr{S}^i, \theta^i)$, where $\mathscr{I}$ is the index set of the excursions of the exploration process above the marks. Now, we decompose, for every $i \in \mathscr{I}$, the excursion $\mathscr{S}^i$ into the excursions $(\bar{\mathscr{S}}_j^i, j \in \mathscr{J})$ the exploration process makes above its minimum mass process. The last excursions, which correspond to excursions above 0, are assembled in one single excursion. If g is a functional, we will define

$$g^*(\mathscr{S}) = \sum_{j\in\mathscr{J}} g(\mathscr{S}_j),$$

where the sum is taken over all the excursions above the minimum mass process. The following lemma, which generalizes a result from [ADV10], enables us, when computing the conditional expectation we want, to use the approximation given by the ε-mesh defined above:

Lemma 3. $\mathbb{P}$*-a.s., we have, for* $\varepsilon > 0$ *small enough,*

$$\sum_{i\in\mathscr{I}} \phi\left(A_{\alpha_i}, \rho_{\alpha_i-}, \mathscr{S}^i, \theta^i\right) = \sum_{k=1}^\infty \phi\left(A_{S_k^\varepsilon}, \rho^-_{S_k^\varepsilon}, \mathscr{S}^{k,\varepsilon}, \theta^{k,\varepsilon}\right) = \sum_{k=1}^\infty \phi^*\left(A_{S_k^\varepsilon}, \rho^-_{S_k^\varepsilon}, \mathscr{S}^{k,\varepsilon}, \theta^{k,\varepsilon}\right), \tag{A.11}$$

where the sums have a finite number of non-zero terms.

The proof is exactly the same as in [ADV10], and relies on hypotheses (h_1) and (h_3) we made earlier.

Computation of the conditional expectation

What we want to prove is the following lemma. This is where we need to be careful in treating the pruning times, because major differences exist with [ADV10].

Lemma 4. *For every $\tilde{\mathscr{F}}_\infty$-measurable non-negative random variable Z, we have*

$$\mathbb{E}\left[Z\exp\left(-\sum_{k=1}^{\infty}\phi^*\left(A_{S_k^\varepsilon},\rho_{S_k^\varepsilon}^-,\mathscr{S}^{k,\varepsilon},\theta^{k,\varepsilon}\right)\right)\right]=\mathbb{E}\left[Z\prod_{k=1}^{\infty}K_\varepsilon\left(A_{S_k^\varepsilon},\rho_{S_k^\varepsilon}^-\right)\right],$$

where $\gamma=\psi^{-1}(1/\varepsilon)$ and

$$K_\varepsilon(r,\mu)=\frac{\psi(\gamma)}{\phi_\theta(\gamma)}\int_0^\theta d\theta'\frac{\gamma-v(r,\mu,\theta')}{\psi(\gamma)-\psi(v(r,\mu,\theta'))}\left[2\beta+\int_0^1du\int_{(0,\infty)}\ell^2\mathrm{e}^{-\theta'\ell}\pi(d\ell)\,w(u\ell,r,\mu,\theta')\mathrm{e}^{-\gamma(1-u)\ell}\right],\quad\text{(A.12)}$$

with

$$w(\ell,r,\mu,\theta')=\mathbb{E}_\ell^*\left[\mathrm{e}^{-\phi(r,\mu,\cdot,\theta')}\right]\ \textit{and}\ v(r,\mu,\theta')=\mathbb{N}\left[1-\mathrm{e}^{-\phi(r,\mu,\cdot,\theta')}\right].\quad\text{(A.13)}$$

Proof. Step 1. We introduce first a special form of the random variable Z.

Let $p\geq1$. Recall that $H_{t,t'}$ denotes the minimum of H between t and t' and that $\bar{H}_t$ defined by A.7 represents the height of the lowest mark. We set

$$\xi_d^{p-1}=\sup\left\{t>T_{p-1}^\varepsilon;\ H_t=H_{T_{p-1}^\varepsilon,S_p^\varepsilon}\right\},$$
$$\xi_g^p=\inf\left\{t>T_{p-1}^\varepsilon;\ H_t=\bar{H}_{S_p^\varepsilon}\text{ and }H_{t,S_p^\varepsilon}=H_t\right\}.$$

ξ_d^{p-1} is the time at which the height process reaches its minimum over $[T_{p-1}^\varepsilon,S_p^\varepsilon]$. By definition of T_{p-1}^ε, $m_{T_{p-1}^\varepsilon}=0$ (there is no mark on the linage of the corresponding individual). On the contrary, $m_{S_p^\varepsilon}\neq0$, $m_{S_p^\varepsilon}(\{\bar{H}_{S_p^\varepsilon}\})\neq0$ but $m_{S_p^\varepsilon}([0,\bar{H}_{S_p^\varepsilon}))=0$. In other words, at time S_p^ε, some mark exists and the lowest mark is situated at height $\bar{H}_{S_p^\varepsilon}$. Roughly speaking, ξ_g^p is the time at which this lowest mark appears. Let us recall that, by the snake property, $m_{\xi_d^{p-1}}=0$ and consequently, $\xi_d^{p-1}<\xi_g^p$ a.s.

We consider a bounded non-negative random variable Z of the form $Z=Z_0Z_1Z_2Z_3$, where the bounded, non-negative random variables $(Z_i,0\leq i\leq3)$ are such that $Z_0\in\mathscr{F}^{(\varepsilon),p-1}$, $Z_1\in\sigma(\mathscr{S}_u,T_{p-1}^\varepsilon\leq u<\xi_d^{p-1})$, $Z_2\in\sigma(\mathscr{S}_u,\xi_d^{p-1}\leq u<\xi_g^p)$ and $Z_3\in\sigma(\mathscr{S}_{(T_k^\varepsilon+s)\wedge S_{k+1}^\varepsilon-},s\geq0,k\geq p)$.

Step 2. Using the strong Markov property of the exploration process several times, we finally get

$$\mathbb{E}\left[Z\exp\left(-\sum_{k=1}^{p}\phi^*(A_{S_k^\varepsilon},\rho_{S_k^\varepsilon}^-,\mathscr{S}^{k,\varepsilon},\theta^{k,\varepsilon})\right)\right]$$
$$=\mathbb{E}\left[Z_0\exp\left(-\sum_{k=1}^{p-1}\phi^*(A_{S_k^\varepsilon},\rho_{S_k^\varepsilon}^-,\mathscr{S}^{k,\varepsilon},\theta^{k,\varepsilon})\right)\phi\left(A_{S_{p-1}^\varepsilon},\rho_{S_{p-1}^\varepsilon}^-,\rho_{T_{p-1}^\varepsilon}\right)\right],\quad\text{(A.14)}$$

with
$$\phi(b,\mu,\nu) = \mathbb{E}^*_\nu\left[Z_1 Z_2 \mathbb{E}^*_{\rho^+_{\tau'}}\left[\mathrm{e}^{-\phi^*(b+A_{\tau'},\mu,\cdot,\theta_{\tau'})}\right]\mathbb{E}^*_{\rho^-_{\tau'}}[Z_3]\right], \tag{A.15}$$
where τ' is distributed under $\mathbb{P}^*_\nu$ as S^ε_1.

Step 3. We compute the function ϕ given by A.15. To simplify the formulas, we set
$$F(b',\mu',\theta') = \mathbb{E}^*_{\mu'}\left[\mathrm{e}^{-\phi^*(b+b',\mu,\cdot,\theta')}\right] \quad ; \quad G(\mu') = \mathbb{E}^*_{\mu'}[Z_3]$$
(the dependence on b and μ of F is omitted) so that
$$\phi(b,\mu,\nu) = \mathbb{E}^*_\nu\left[Z_1 Z_2 F(A_{\tau'},\rho^+_{\tau'},\theta_{\tau'})G(\rho^-_{\tau'})\right]. \tag{A.16}$$

Lemma 5. *We set* $q(du,d\ell,d\theta') = 2\beta\delta_{(0,0)}(du,d\ell)d\theta' + du\,\ell^2\mathrm{e}^{-\theta'\ell}\pi(d\ell)d\theta'$ *and by convention* $\pi(\{0\}) = 0$. *We have:*
$$\phi(b,\mu,\nu) = \mathbb{E}_\nu\left[Z_1 Z_2\frac{\Gamma_F(A_{\tau'})}{\Gamma_1}G(\rho^-_{\tau'})\right], \tag{A.17}$$
where for a non-negative function f defined on $[0,\infty)\times\mathcal{M}_f(\mathbf{R}_+)\times\mathbf{R}_+$
$$\Gamma_f(a) = \int_{[0,1]\times[0,\infty)} q(du,d\ell,d\theta')\int \mathbb{M}(d\rho',d\eta',dm')\,\mathrm{e}^{-\gamma\langle\rho',1\rangle-\gamma u\ell} f(a,\eta'+(1-u)\ell\delta_0,\theta')$$
and for $f=1$, Γ_1 does not depend on a.

The proof of this lemma, although quite technical, can actually be given by adapting the proof in [ADV10] *mutatis mutandis*. For this reason, we will not repeat it here.

We now use the particular form of F to compute Γ_F. Using the Poissonian structure of the excursions of the exploration above its minimum, we get
$$\begin{aligned}F(a,\mu',\theta') &= \mathbb{E}^*_{\mu'}\left[\mathrm{e}^{-\phi^*(b+a,\mu,\cdot,\theta')}\right]\\ &= \mathbb{E}^*_{\mu'(\{0\})}\left[\mathrm{e}^{-\phi(b+a,\mu,\cdot,\theta')}\right]\mathrm{e}^{-\mu'((0,\infty))\mathbb{N}\left[1-\mathrm{e}^{-\phi(b+a,\mu,\cdot,\theta')}\right]}.\end{aligned}$$
Using w and v defined in A.13, we get
$$\begin{aligned}&\mathbb{M}_s\left[\mathrm{e}^{-\gamma\langle\rho,1\rangle-\gamma u\ell}F(a,\eta+(1-u)\ell\delta_0,\theta')\right]\\ &\qquad= w((1-u)\ell,b+a,\mu,\theta')\mathrm{e}^{-\gamma u\ell}\mathbb{M}_s\left[\mathrm{e}^{-\gamma\langle\rho,1\rangle}\mathrm{e}^{-v(b+a,\mu,\theta')\langle\eta,1\rangle}\right]\\ &\qquad= w((1-u)\ell,b+a,\mu,\theta')\mathrm{e}^{-\gamma u\ell}\exp\left(-s\left(\frac{\psi(\gamma)-\psi(v(b+a,\mu,\theta'))}{\gamma-v(b+a,\mu,\theta')}-\alpha\right)\right).\end{aligned}$$
We deduce that
$$\begin{aligned}\Gamma_F(a) = \int_0^\theta d\theta'\frac{\gamma-v(b+a,\mu,\theta')}{\psi(\gamma)-\psi(v(b+a,\mu,\theta'))}&\\ \left(2\beta+\int_0^1 du\int_{(0,\infty)}\ell^2\exp(-\theta'\ell)\pi(d\ell)\,w(u\ell,b+a,\mu,\theta')\mathrm{e}^{-\gamma(1-u)\ell}\right)&,\end{aligned}$$

and with $F = 1$, $\Gamma_1 = \frac{\gamma}{\psi(\gamma)} \frac{\phi_\theta(\gamma)}{\gamma} = \frac{\phi_\theta(\gamma)}{\psi(\gamma)}$.

Finally, plugging this formula in A.17 and using the function K_ε introduced in A.12, we have

$$\phi(b,\mu,\nu) = \mathbb{E}_\nu[Z_1 Z_2 K_\varepsilon(b + A_{\tau'},\mu) G(\rho^-_{\tau'})]. \tag{A.18}$$

Step 4. Induction.

Plugging the expression A.18 for ϕ in A.14, and using the arguments backward from A.14 we get

$$\mathbb{E}\left[Z \exp\left(-\sum_{k=1}^{p} \phi^*(A_{S_k^\varepsilon}, \rho^-_{S_k^\varepsilon}, \mathscr{S}^{k,\varepsilon}, \theta^{k\varepsilon}) \right)\right]$$
$$= \mathbb{E}\left[Z \exp\left(-\sum_{k=1}^{p-1} \phi^*(A_{S_k^\varepsilon}, \rho^-_{S_k^\varepsilon}, \mathscr{S}^{k,\varepsilon}, \theta^{k,\varepsilon}) \right) K_\varepsilon(A_{S_p^\varepsilon}, \rho^-_{S_p^\varepsilon}) \right].$$

In particular, using a monotone class argument, we see that this equality holds for any non-negative Z measurable w.r.t. the σ-field $\bar{\mathscr{F}}^\varepsilon_\infty = \sigma((\mathscr{S}_{C_t}, t \in [A_{T_k^\varepsilon}, A_{S_{k+1}^\varepsilon}]), k \geq 0)$. Notice that $K_\varepsilon(A_{S_p^\varepsilon}, \rho^-_{S_p^\varepsilon})$ is measurable w.r.t. $\bar{\mathscr{F}}_\infty$. So, we may iterate the previous argument and let p go to infinity to finally get that for any non-negative random variable $Z \in \bar{\mathscr{F}}_\infty$, we have

$$\mathbb{E}\left[Z \exp\left(-\sum_{k=1}^{\infty} \phi^*(A_{S_k^\varepsilon}, \rho^-_{S_k^\varepsilon}, \mathscr{S}^{k,\varepsilon}) \right)\right] = \mathbb{E}\left[Z \prod_{k=1}^{\infty} K_\varepsilon(A_{S_k^\varepsilon}, \rho^-_{S_k^\varepsilon}) \right].$$

Intuitively, $\bar{\mathscr{F}}^\varepsilon_\infty$ is the σ-field generated by $\tilde{\mathscr{F}}_\infty$ and the mesh $([A_{T_k^\varepsilon}, A_{S_{k+1}^\varepsilon}], k \geq 0)$. As $\bar{\mathscr{F}}^\varepsilon_\infty$ contains $\tilde{\mathscr{F}}_\infty$, the Lemma is proved. □

Computation of the limit

Again, the analysis in [ADV10] remains valid. We will only recall the following result, which we will of course apply to K_ε in order to pass to the limit in our conditional expectation:

Corollary 6. *There exists a sub-sequence $(\varepsilon_j, j \in \mathbf{N})$ decreasing to 0, s.t. $\mathbb{P}$-a.s. for any $t_0 \geq 0$ and any continuous function h defined on $\mathbf{R}_+ \times \mathscr{M}_f(\mathbf{R}_+)$ such that $h(u,\mu) = 0$ for $u \geq t_0$, we have, with $\gamma_j = \psi^{-1}(1/\varepsilon_j)$,*

$$\lim_{j\to\infty} \phi_1(\gamma_j)^{-1} \sum_{k=1}^{\infty} h(A_{S_k^{\varepsilon_j}}, \rho^-_{S_k^{\varepsilon_j}}) = \int_0^\infty h(u, \rho_u^\theta)\, du.$$

The next lemma ensures that K_ε converges uniformly to the expected limit:

Lemma 7. *There exists a deterministic function $\mathscr{R}$ s.t. $\lim_{\varepsilon\to 0} \mathscr{R}(\varepsilon) = 0$ and for all $\varepsilon > 0$ and $\mu \in \mathscr{M}_f(\mathbf{R}_+)$, we have:*

$$\sup_{r\geq 0} \left| \phi_\theta(\gamma) \log(K_\varepsilon(r,\mu)) - \int_0^\theta \Big(2\beta v(r,\mu,\theta') \right.$$
$$\left. - \int_{(0,\infty)} \ell e^{-\theta'\ell} \pi(d\ell)\, \big(1 - w(\ell,r,\mu,\theta')\big)\, d\theta' \Big) \right| \leq \mathscr{R}(\varepsilon). \tag{A.19}$$

Proof. The same transformations as in [ADV10] show that:

$$K_\varepsilon(r,\mu) = \int_0^\theta d\theta' \left(\frac{1 - v(r,\mu,\theta')/\gamma}{1 - \psi(v(r,\mu,\theta'))/\psi(\gamma)} \right)$$
$$\frac{1}{\phi_\theta(\gamma)} \left(\frac{\partial \phi_{\theta'}}{\partial \theta'}(\gamma) - \int_{(0,\infty)} \pi_{\theta'}(d\ell) \int_0^{\gamma\ell} \mathrm{e}^{-s} ds (1 - w(\ell - s/\gamma, r, \mu, \theta')) \right), \quad \text{(A.20)}$$

with $\pi_{\theta'}(d\ell) = \mathrm{e}^{-\theta'\ell}\pi(d\ell)$. Now, writing $1/(1-\psi(v)/\psi(\gamma)) = 1 + \psi(v)/\psi(\gamma) + R_1(r,\mu,\theta',\gamma)$ and

$$\int_{(0,\infty)} \pi_{\theta'}(d\ell) \int_0^{\gamma\ell} \mathrm{e}^{-s} ds (1 - w(\ell - s/\gamma, r, \mu, \theta'))$$
$$= \int_{(0,\infty)} \pi_{\theta'}(d\ell) \int_0^{\gamma\ell} \mathrm{e}^{-s} ds (1 - w(\ell, r, \mu, \theta'))$$
$$- \int_{(0,\infty)} \pi_{\theta'}(d\ell) \int_0^{\gamma\ell} \mathrm{e}^{-s} ds (w(\ell - s/\gamma, r, \mu, \theta') - w(\ell, r, \mu, \theta')) \quad \text{(A.21)}$$

we get the following exact expression for $\phi_\theta(\gamma)\log K_\varepsilon(r,\mu)$:

$$\phi_\theta(\gamma)\log\left(1 + \int_0^\theta d\theta' \left(-\frac{1}{\phi_\theta(\gamma)} \frac{\partial\phi_{\theta'}}{\partial\theta'}(\gamma) \frac{v(r,\mu,\theta')}{\gamma} + \frac{1}{\phi_\theta(\gamma)} \frac{\partial\phi_{\theta'}}{\partial\theta'}(\gamma) \frac{\psi(v(r,\mu,\theta'))}{\psi(\gamma)} \right.\right.$$
$$\left.\left. + \frac{1}{\phi_\theta(\gamma)} \int_{(0,\infty)} \pi_{\theta'}(d\ell)(1 - \mathrm{e}^{-\gamma\ell})(1 - w(\ell, r, \mu, \theta')) + R_2(r,\mu,\theta',\gamma) \right)\right), \quad \text{(A.22)}$$

where $R_2(r,\mu,\theta',\gamma)$ is the sum of 14 terms. The analysis in Section A shows that there exists $C_1 > 0$ such that for all $r \geq 0$ and $\theta' \in [0,\theta]$,

$$\left| \int_{(0,\infty)} \pi_{\theta'}(d\ell) \int_0^{\gamma\ell} \mathrm{e}^{-s} ds \big(w(\ell - s/\gamma, r, \mu, \theta') - w(\ell, r, \mu, \theta') \big) \right| \leq \frac{C_1}{\gamma} \quad \text{(A.23)}$$

In a similar way, using hypothesis (h_3), we see that there exists $C_2 > 0$ such that when γ is sufficiently large, for all $r \geq 0$ and $\theta' \in [0,\theta]$,

$$\left| R_1(r,\mu,\theta',\gamma) \right| \leq \frac{C_2}{\psi(\gamma)^2} \quad \text{(A.24)}$$

A detailed computation of $R_2(r,\mu,\theta',\gamma)$ then shows that, when γ is sufficiently large, there exists $C_3 > 0$ such that, for every $r \geq 0$ and $\theta' \in [0,\theta]$,

$$\left| R_2(r,\mu,\theta',\gamma) \right| \leq \frac{C_3}{\psi(\gamma)} \quad \text{(A.25)}$$

Thus, when developing the logarithm, we get:

$$
\begin{aligned}
&\left|\phi_\theta(\gamma)\log(K_\varepsilon(r,\mu)) - \int_0^\theta 2\beta v(r,\mu,\theta') - \int_{(0,\infty)} \ell \mathrm{e}^{-\theta'\ell}\pi(d\ell)\,\left(1-w(\ell,r,\mu,\theta')\right)d\theta'\right| \\
&\qquad = \left|\int_0^\theta d\theta'\left(v(r,\mu,\theta')\left(\frac{\partial\phi_{\theta'}}{\partial\theta'}(\gamma)/\gamma - 2\beta\right) - \int_{(0,\infty)} \pi_{\theta'}(d\ell)\mathrm{e}^{-\gamma\ell}(1-w(\ell,r,\mu,\theta'))\right.\right. \\
&\qquad\qquad \left.\left. + \frac{\partial\phi_{\theta'}}{\partial\theta'}(\gamma)\frac{\psi(v(r,\mu,\theta'))}{\psi(\gamma)} + R_3(r,\mu,\theta',\gamma)\right)\right| \qquad \text{(A.26)}
\end{aligned}
$$

where R_3 takes into account the error term R_2 as well as the error terms arising from the approximation $\log(1-u) \simeq u$. As such, considering (A.25) and considering that, for every $r \geq 0$,

$$
\left|\int_0^\theta \left(\left(\frac{\partial\phi_{\theta'}}{\partial\theta'}(\gamma) - 2\beta\right) v(r,\mu,\theta')\right) d\theta'\right| \leq c_1|\phi_\theta(\gamma)/\gamma - 2\beta| \qquad \text{(A.27)}
$$

$$
\left|\int_0^\theta \frac{\partial\phi_{\theta'}}{\partial\theta'}(\gamma)\frac{\psi(v(r,\mu,\theta'))}{\psi(\gamma)} d\theta'\right| \leq c_2|\phi_\theta(\gamma)/\psi(\gamma)| \qquad \text{(A.28)}
$$

with suitable constants c_1 and c_2, we easily get the result. □

Thus, using Corollary 6 along with the previous Lemma, we get the following convergence:

Lemma 8. *Let ϕ satisfying condition (h_1)–(h_3). There exists a sub-sequence $(\varepsilon_j, j \in \mathbf{N})$ decreasing to* 0*, s.t. $\mathbb{P}$-a.s.*

$$
\lim_{j\to\infty}\prod_{k=1}^\infty K_{\varepsilon_j}(A_{S_k^{\varepsilon_j}}, \rho_{S_k^{\varepsilon}}^-) = \exp - \int_0^\theta d\theta' \int_0^\infty du\,\left(2\beta v(u,\rho_u^\theta,\theta') + \int_{(0,\infty)} \ell\mathrm{e}^{-\theta'\ell}\pi(d\ell)\,(1-w(\ell,u,\rho_u^\theta,\theta'))\right).
$$

End of the proof

Let $Z \in \tilde{\mathscr{F}}_\infty$ non-negative such that $\mathbb{E}[Z] < \infty$. Let ϕ satisfying hypothesis of the Special Markov Property, along with (h_1)–(h_3). We have, using notation of the previous sections

$$
\begin{aligned}
&\mathbb{E}\left[Z\exp\left(-\sum_{i\in I}\phi(A_{\alpha_i}, \rho_{\alpha_i -}, \mathscr{S}^i, \theta^i)\right)\right] \\
&= \lim_{j\to\infty}\mathbb{E}\left[Z\exp\left(-\sum_{k=1}^\infty \phi^*\left(A_{S_k^{\varepsilon_j}}, \rho_{S_k^{\varepsilon_j}}^-, \mathscr{S}^{k,\varepsilon_j}, \theta^{k,\varepsilon_j}\right)\right)\right] \\
&= \lim_{j\to\infty}\mathbb{E}\left[Z\prod_{k=1}^\infty K_{\varepsilon_j}(A_{S_k^{\varepsilon_j}}, \rho_{S_k^{\varepsilon_j}}^-)\right] \\
&= \mathbb{E}\left[Z\exp\left(-\int_0^\infty du\int_0^\theta d\theta'\left(2\beta v(u,\rho_u^\theta,\theta') + \int_{(0,\infty)} \ell\mathrm{e}^{-\theta'\ell}\pi(d\ell)\,(1-w(\ell,u,\rho_u^\theta,\theta'))\right)\right)\right],
\end{aligned}
$$

where we used Lemma 3 and dominated convergence for the first equality, Lemma 4 for the second equality, Lemma 8 and dominated convergence for the last equality. By using a monotone class argument and by monotonicity, we can remove hypotheses (h_1)–(h_3). To end the proof of the first part, notice that

$$\int_0^\infty du \int_0^\theta d\theta' \left(2\beta v(u, \rho_u^\theta, \theta') + \int_{(0,\infty)} \ell \mathrm{e}^{-\theta' \ell} \pi(d\ell)\, (1 - w(\ell, u, \rho_u^\theta, \theta'))\right)$$

is $\tilde{\mathscr{F}}_\infty$-measurable and so this is $\mathbb{P}$-a.s. equal to the conditional expectation (i.e. the left hand side term of (A.4)).

Bibliography

[ABBH11] Louigi Addario-Berry, Nicolas Broutin, and Cecilia Holmgren, *Cutting down trees with a Markov chainsaw*, Arxiv preprint arXiv:1110.6455 (2011), 1–23.

[AD07] Romain Abraham and Jean-François Delmas, *Fragmentation associated with Lévy processes using snake*, Probability Theory and Related Fields **141** (2007), 113–154.

[AD09a] ———, *Changing the branching mechanism of a continuous state branching process using immigration*, Annales de l'Institut Henri Poincaré (B) Probability and Statistics **45** (2009), no. 1, 226–238.

[AD09b] ———, *Williams' decomposition of the Lévy continuum random tree and simultaneous extinction probability for populations with neutral mutations*, Stochastic Processes and their Applications **119** (2009), no. 4, 1124–1143.

[AD11] ———, *Record process on the Continuum Random Tree*, Arxiv preprint arXiv:1107.3657 (2011), 1–25.

[AD12a] ———, *A continuum-tree-valued Markov process*, The Annals of Probability **40** (2012), no. 3, 1167–1211.

[AD12b] ———, *Record processes on Lévy trees*, In preparation (2012).

[ADH12a] Romain Abraham, Jean-François Delmas, and Hui He, *Pruning Galton–Watson trees and tree-valued Markov processes*, Annales de l'Institut Henri Poincaré (B) Probability and Statistics **48** (2012), no. 3, 688–705.

[ADH12b] Romain Abraham, Jean-François Delmas, and Patrick Hoscheit, *A note on Gromov-Hausdorff-Prokhorov distance between (locally) compact measure spaces*, Arxiv preprint arXiv:1202.5464 (2012), 1–19.

[ADH12c] ———, *Exit times for an increasing Lévy tree-valued process*, Arxiv preprint arXiv:1202.5463 (2012), 1–33.

[ADV10] Romain Abraham, Jean-François Delmas, and Guillaume Voisin, *Pruning a Lévy continuum random tree*, Electronic journal of probability **15** (2010), 1429–1473.

[AE78] David Aldous and G. K. Eagleson, *On Mixing and Stability of Limit Theorems*, The Annals of Probability **6** (1978), no. 2, 325–331.

[AEW12] Siva R. Athreya, Michael Eckhoff, and Anita Winter, *Brownian motion on* **R**-*trees*, To appear in Transactions of the AMS (2012), 1–39.

[Ald91a] David Aldous, *The Continuum Random Tree I*, The Annals of Probability **19** (1991), no. 1, 1–28.

[Ald91b] ———, *The continuum random tree II: an overview*, Stochastic Analysis (M.T. Barlow and N.H. Bingham, eds.), Cambridge University Press, 1991, pp. 23–70.

[Ald93] ———, *The Continuum Random Tree III*, The Annals of Probability **21** (1993), no. 1, 248–289.

[AP98a] David Aldous and Jim Pitman, *The standard additive coalescent*, The Annals of Probability **26** (1998), no. 4, 1703–1726.

[AP98b] ———, *Tree-valued Markov chains derived from Galton-Watson processes*, Annales de l'Institut Henri Poincaré (B) Probability and Statistics **34** (1998), no. 5, 637–686.

[AS02] Romain Abraham and Laurent Serlet, *Poisson snake and fragmentation*, Electronic Journal of Probability **7** (2002), 1–15.

[BBI01] Dmitri Burago, Youri Dmitrievitch Burago, and Sergeï Ivanov, *A course in metric geometry*, AMS Providence, 2001.

[BdFG04] Jérémie Bouttier, Philippe di Francesco, and Emmanuel Guitter, *Planar maps as labeled mobiles*, Electronic Journal of Combinatorics **11** (2004), 1–27.

[Ber01] Jean Bertoin, *Homogeneous fragmentation processes*, Probability Theory and Related Fields **121** (2001), no. 3, 301–318.

[Ber02] ———, *Self-similar fragmentations*, Annales de l'Institut Henri Poincaré (B) Probability and Statistics **3** (2002), 319–340.

[Ber06] ———, *Random Fragmentation and Coagulation Processes*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2006.

[Bes02] Mladen Bestvina, *R-trees in topology, geometry, and group theory*, Handbook of geometrical topology, North Holland, 2002, pp. 55–91.

[Bet10] Jérémie Bettinelli, *Scaling Limits for Random Quadrangulations of Positive Genus*, Electronic Journal of Probability **15** (2010), 1594–1644.

[Bis85] Jean-Michel Bismut, *Last exit decompositions and regularity at the boundary of transition probabilities*, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **69** (1985), no. 1, 65–98.

[BKMS11] Julien Berestycki, Andreas Kyprianou, and Antonio Murillo-Salas, *The prolific backbone for supercritical superdiffusions*, Stochastic Processes and their Applications **121** (2011), no. 6, 1315–1331.

[BM12] Jean Bertoin and Grégory Miermont, *The cut-tree of large Galton-Watson trees and the Brownian CRT*, Arxiv preprint arXiv:1201.4081 (2012), 1–25.

[CH10] David A Croydon and Benjamin M Hambly, *Spectral Asymptotics for Stable Trees*, Electronic Journal of Probability **15** (2010).

[CH12] Nicolas Curien and Bénédicte Haas, *The stable trees are nested*, Arxiv preprint arXiv:1207.5418 (2012), 1–31.

[Cha97] Loïc Chaumont, *Excursion normalisée, méandre et pont pour les processus de Lévy stables*, Bull. sci. math. **121** (1997), no. 5, 377–403.

[Chi01] Ian Chiswell, *Introduction to* Λ*-trees*, World Scientific Publishing Co. Inc., 2001.

[Cro07] David Croydon, *Volume growth and heat kernel estimates for the continuum random tree*, Probability Theory and Related Fields **140** (2007), no. 1-2, 207–238.

[Cro09] ______, *Hausdorff measure of arcs and Brownian motion on Brownian spatial trees*, The Annals of Probability **37** (2009), no. 3, 946–978.

[CS04] Philippe Chassaing and Gilles Schaeffer, *Random planar lattices and integrated superBrownian excursion*, Probability Theory and Related Fields **128** (2004), 161–212.

[DGM06] Rui Dong, Christina Goldschmidt, and James B. Martin, *Coagulation–fragmentation duality, Poisson–Dirichlet distributions and random recursive trees*, The Annals of Applied Probability **16** (2006), no. 4, 1733–1750.

[DGP11] Andrej Depperschmidt, Andreas Greven, and Peter Pfaffelhuber, *Tree-valued Fleming-Viot dynamics with mutation and selection*, The Annals of Probability (2011), 58.

[DH11] Jean-François Delmas and Olivier Hénard, *A Williams' Decomposition for Spatially Dependent Superprocesses*, Arxiv preprint arXiv:1106.3710 (2011), 1–43.

[DIMR08] Michael Drmota, Alex Iksanov, Martin Möhle, and Uwe Roesler, *A Limiting Distribution for the Number of Cuts Needed to Isolate the Root of a Random Recursive Tree*, Random Structures & Algorithms (2008), 319–336.

[DL02] Thomas Duquesne and Jean-François Le Gall, *Random trees, Lévy processes and spatial branching processes*, Astérisque, vol. 281, SMF, 2002.

[DL05] ______, *Probabilistic and fractal aspects of Lévy trees*, Probability Theory and Related Fields **131** (2005), no. 4, 553–603.

[DL06] ______, *The Hausdorff measure of stable trees*, Alea **1** (2006), 393–415.

[Dre84] Andreas W.M Dress, *Trees, tight extensions of metric spaces, and the cohomological dimension of certain groups: A note on combinatorial properties of metric spaces*, Advances in Mathematics **53** (1984), no. 3, 321–402.

[Drm09] Michael Drmota, *Random trees*, Springer, 2009.

[Duf97] Marie Duflo, *Random iterative models*, Applications of Mathematics, Springer-Verlag, 1997.

[Duq03] Thomas Duquesne, *A limit theorem for the contour process of condidtioned Galton–Watson trees*, The Annals of Probability **31** (2003), no. 2, 996–1027.

[Duq06] ______, *The coding of compact real trees by real valued functions*, Arxiv preprint arXiv:0604106 (2006), 50.

[Duq10] ______, *Packing and Hausdorff measures of stable trees*, Lévy Matters I (Ole Barndorff-Nielsen, Jean Bertoin, Jean Jacod, and Claudia Klüppleberg, eds.), Springer, January 2010, pp. 93–136.

[Duq12] ______, *The exact packing measure of Lévy trees*, Stochastic Processes and their Applications **122** (2012), no. 3, 968–1002.

[DVJ03] Daryl Daley and David Vere-Jones, *An Introduction to the Theory of Point Processes*, Probability and its Applications, Springer-Verlag, New York, 2003.

[DW07] Thomas Duquesne and Matthias Winkel, *Growth of Lévy trees*, Probability Theory and Related Fields **139** (2007), no. 3-4, 313–371.

[DW12] ______, *Hereditary tree growth and Levy forests*, Arxiv preprint arXiv:1211.2179 (2012), 44.

[Dyn91] E. B. Dynkin, *A probabilistic approach to one class of nonlinear differential equations*, Probability Theory and Related Fields **89** (1991), no. 1, 89–115.

[Dyn93] ______, *Superprocesses and Partial Differential Equations*, The Annals of Probability **21** (1993), no. 3, 1185–1262.

[EPW05] Steven N. Evans, Jim Pitman, and Anita Winter, *Rayleigh processes, real trees, and root growth with re-grafting*, Probability Theory and Related Fields **134** (2005), no. 1, 81–126.

[Eva08] Steven N. Evans, *Probability and real trees*, Lecture Notes in Mathematics, vol. 1920, Springer, 2008.

[EW06] Steven N. Evans and Anita Winter, *Subtree prune and regraft: A reversible real tree-valued Markov process*, The Annals of Probability **34** (2006), no. 3, 918–961.

[FO80] Philippe Flajolet and Andrew Odlyzko, *Exploring binary trees and other simple trees*, 21st Annual Symposium on Foundations of Computer Science, IEEE, October 1980, pp. 207–216.

[Fuk87] Kenji Fukaya, *Collapsing of Riemannian manifolds and eigenvalues of Laplace operator*, Inventiones Mathematicae **87** (1987), no. 3, 517–547.

[GPW08] Andreas Greven, Peter Pfaffelhuber, and Anita Winter, *Convergence in distribution of random metric measure spaces (Λ-coalescent measure trees)*, Probability Theory and Related Fields **145** (2008), no. 1-2, 285–322.

[GPW12] ———, *Tree-valued resampling dynamics Martingale problems and applications*, Probability Theory and Related Fields (2012).

[Gro81] Mikhail Gromov, *Groups of Polynomial growth and Expanding Maps*, Publications mathématiques de l'IHÉS **53** (1981), 53–78.

[Gro07] ———, *Metric Structures for Riemannian and Non-Riemannian Spaces*, Modern Birkhäuser Classics, Birkhäuser, 2007.

[HH80] P. Hall and C. C. Heyde, *Martingale limit theory and its applications*, Academic Press, Inc., 1980.

[HM04] Bénédicte Haas and Grégory Miermont, *The Genealogy of Self-similar Fragmentations with Negative Index as a Continuum Random Tree*, Electronic journal of probability **9** (2004), 57–97.

[HM12] ———, *Scaling limits of Markov branching trees, with applications to Galton-Watson and random unordered trees*, To appear in The Annals of Probability (2012), 1–58.

[HMPW08] Bénédicte Haas, Grégory Miermont, Jim Pitman, and Matthias Winkel, *Continuum tree asymptotics of discrete fragmentations and applications to phylogenetic models*, The Annals of Probability **36** (2008), no. 5, 1790–1837.

[Hol11] Cecilia Holmgren, *A weakly 1-stable distribution for the number of random records and cuttings in split trees*, Advances in Applied Probability **43** (2011), no. 1, 151–177.

[Hos12] Patrick Hoscheit, *Fluctuations for the number of records on random binary trees*, Submitted (2012).

[HPW09] Bénédicte Haas, Jim Pitman, and Matthias Winkel, *Spinal partitions and invariance under re-rooting of continuum random trees*, The Annals of Probability **37** (2009), no. 4, 1381–1411.

[Jan06] Svante Janson, *Random cutting and records in deterministic and random trees*, Random Structures & Algorithms **29** (2006), no. 2, 139–179.

[Jan12] ______, *Simply generated trees, conditioned Galton–Watson trees, random allocations and condensation*, Probability Surveys **9** (2012), 103–252.

[Kor12a] Igor Kortchemski, *Invariance principles for Galton-Watson trees conditioned on the number of leaves*, To appear in Stochastic Processes and Applications (2012), 1–46.

[Kor12b] ______, *Limit theorems for conditioned non-generic Galton-Watson trees*, Arxiv preprint arXiv:1205.3145 (2012), 1–27.

[Kyp06] Andreas Kyprianou, *Introductory Lectures on Fluctuations of Lévy Processes with Applications*, Universitext, Springer, 2006.

[LÏ2] Wolfgang Löhr, *Equivalence of Gromov-Prohorov- and Gromov's Box-1-Metric on the Space of Metric Measure Spaces*, Arxiv preprint arXiv:1111.5837 (2012), 1–8.

[Lam08] Amaury Lambert, *Population Dynamics and Random Genealogies*, Stochastic Models **24** (2008), 45–163.

[Le 91] Jean-François Le Gall, *Brownian Excursions, Trees and Measure-Valued Branching Processes*, The Annals of Probability **19** (1991), no. 4, 1399–1439.

[Le 93a] ______, *A class of path-valued Markov processes and its applications to superprocesses*, Probability Theory and Related Fields **95** (1993), no. 1, 25–46.

[Le 93b] ______, *The uniform random tree in a Brownian excursion*, Probability Theory and Related Fields **96** (1993), no. 3, 369–383.

[Le 94] ______, *Hitting probabilities and potential theory for the Brownian path-valued process*, Annales de l'Institut Fourier **44** (1994), 277–306.

[Le 95] ______, *The Brownian snake and solutions of $\Delta u = u^2$ in a domain*, Probability Theory and Related Fields **102** (1995), no. 3, 393–432.

[Le 05] ______, *Random trees and applications*, Probability Surveys **2** (2005), 245–311.

[Le 06] ______, *Random real trees*, Annales de la faculté des Sciences de Toulouse - Mathématiques **15** (2006), no. 1, 35.

[Le 12] ______, *Uniqueness and universality of the Brownian map*, To appear in The Annals of Probability (2012), 1–88.

[Li11] Zenghu Li, *Measure-Valued Branching Markov Processes*, Probability and Its Applications, Springer, 2011.

[Li12] ______, *Path-valued branching processes and nonlocal branching superprocesses*, To appear in The Annals of Probability (2012), 1–37.

[LL98a] Jean-François Le Gall and Yves Le Jan, *Branching processes in Lévy processes: Laplace functionals of snakes and superprocesses*, The Annals of Probability **26** (1998), no. 4, 1407–1432.

[LL98b] ———, *Branching processes in Lévy processes: the exploration process*, The Annals of Probability **26** (1998), no. 1, 213–252.

[LM11] Jean-François Le Gall and Grégory Miermont, *Scaling limits of random planar maps with large faces*, The Annals of Probability **39** (2011), no. 1, 1–69.

[LM12] ———, *Scaling limits of random trees and planar maps*, To appear in Clay Mathematics Proceedings **15** (2012), 1–76.

[LV09] John Lott and Cédric Villani, *Ricci curvature for metric-measure spaces via optimal transport*, Annals of Mathematics **169** (2009), no. 3, 903–991.

[Mar08] Philippe Marchal, *A note on the fragmentation of a stable tree*, Discrete Mathematics And Theoretical Computer Science (2008), 489–500.

[Mie03] Grégory Miermont, *Self-similar fragmentations derived from the stable tree I*, Probability Theory and Related Fields **127** (2003), no. 3, 423–454.

[Mie04] ———, *Self-similar fragmentations derived from the stable tree II: splitting at nodes*, Probability Theory and Related Fields **131** (2004), no. 3, 341–375.

[Mie09] ———, *Tessellations of random maps of arbitrary genus*, Annales scientifiques de l'École Normale supérieure **42** (2009), 725–781.

[Mie12] ———, *The Brownian map is the scaling limit of uniform random plane quadrangulations*, To appear in Acta Mathematica (2012), 1–76.

[MM70] A. Meir and J. W. Moon, *Cutting down random trees*, Journal of the Australian Mathematical Society **11** (1970), no. 03, 313–324.

[Nev86] Jacques Neveu, *Arbres et processus de Galton-Watson*, Annales de l'Institut Henri Poincaré (B) Probability and Statistics **22** (1986), no. 2, 199–207.

[Pan06] Alois Panholzer, *Cutting down very simple trees*, Quaestiones Mathematicae **29** (2006), no. 2, 211–227.

[Pit06] Jim Pitman, *Combinatorial Stochastic Processes*, Lecture Notes in Mathematics, vol. 1875, Springer-Verlag, Berlin/Heidelberg, 2006.

[PW06] Peter Pfaffelhuber and Anton Wakolbinger, *The process of most recent common ancestors in an evolving coalescent*, Stochastic Processes and their Applications **116** (2006), 1836–1859.

[PWW10] Peter Pfaffelhuber, Anton Wakolbinger, and Heinz Weisshaupt, *The tree length of an evolving coalescent*, Probability Theory and Related Fields **151** (2010), no. 3-4, 529–557.

[Riz11] Douglas Rizzolo, *Scaling limits of Markov branching trees and Galton-Watson trees conditioned on the number of vertices with out-degree in a given set*, Arxiv preprint arXiv:1105.2528 (2011), 1–23.

[RY05] Daniel Revuz and Marc Yor, *Continuous Martingales and Brownian Motion*, 3rd ed., Grundlehren der mathematischen Wissenschaften, Springer, 2005.

[Stu06a] Karl-Theodor Sturm, *On the geometry of metric measure spaces I.*, Acta Mathematica **196** (2006), no. 1, 65–131.

[Stu06b] ———, *On the geometry of metric measure spaces. II*, Acta Mathematica **196** (2006), no. 1, 133–177.

[Vil09] Cédric Villani, *Optimal Transport*, Grundlehren der mathematischen Wissenschaften, vol. 338, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.

[Voi10] Guillaume Voisin, *Dislocation measure of the fragmentation of a general Lévy tree*, ESAIM: Probability and Statistics (2010).

[Wei07] Mathilde Weill, *Regenerative real trees*, The Annals of Probability **35** (2007), no. 6, 2091–2121.

[Wil74] David Williams, *Path Decomposition and Continuity of Local Time for One-Dimensional Diffusions, I*, Proceedings of the London Mathematical Society **s3-28** (1974), no. 4, 738–768.

Résumé. Cette thèse est consacrée à l'étude de certains processus aléatoires à valeurs dans les arbres continus. Nous définissons d'abord un cadre conceptuel pour cette étude, en construisant une topologie polonaise sur l'espace des **R**-arbres localement compacts, complets et munis d'une mesure borélienne localement finie. Cette topologie, dite de Gromov-Hausdorff-Prokhorov, permet alors la définition de processus de Markov à valeurs arbre. Nous donnons ensuite une nouvelle construction du processus d'élagage d'Abraham-Delmas-Voisin, qui est un exemple de processus qui prend ses valeurs dans les arbres de Lévy. Notre construction, qui dévoile une nouvelle structure généalogique des arbres de Lévy, est trajectorielle, et permet d'identifier explicitement les transitions du processus d'élagage. Nous appliquons cette description à l'étude de certains temps d'arrêt, comme le premier temps auquel le processus franchit une hauteur donnée. Nous décrivons le processus à cet instant grâce à une nouvelle décomposition de type spinal. Enfin, nous nous intéressons à la fragmentation d'Aldous-Pitman de l'arbre brownien d'Aldous. En particulier, nous étudions, à la suite d'Abraham et Delmas, l'effet de cette fragmentation sur les sous-arbres discrets de l'arbre brownien. Le nombre de coupures nécessaires avant d'isoler la racine, convenablement renormalisé, converge vers une variable aléatoire de Rayleigh ; nous donnons un théorème central limite qui précise les fluctuations autour de cette limite.

Abstract. In this thesis, we study continuum tree-valued processes. First, we define an abstract framework for these processes, by constructing a metric on the space of locally compact, complete **R**-trees, endowed with a locally finite Borel measure. This topology, called Gromov-Hausdorff-Prokhorov topology, allows for the definition of tree-valued Markov processes. We then give a new construction of the pruning process of Abraham-Delmas-Voisin, which is an example of a Lévy tree-valued process. Our construction reveals a new genealogical structure of Lévy trees. Furthermore, it is a pathwise construction, which describes the transitions of the process explicitly. We apply this description to the study of certain stopping times, such as the first moment the process crosses a given height. We describe the process at that time through a new spinal decomposition. Finally, we focus on the Aldous-Pitman fragmentation of Aldous's Brownian tree. Following Abraham and Delmas, we study the effect of the fragmentation on discrete subtrees of the Brownian tree. The number of cuts needed to isolate the root, suitably renormalized, converges towards a Rayleigh-distributed random variable; we prove a Central Limit Theorem describing the fluctuations around this limit.