Arbres aléatoires : asymptotique de fonctionnelles et limites locales

Michel Nassif

École doctorale Mathématiques et STIC (MSTIC)

THÈSE DE DOCTORAT

Spécialité : Mathématiques

Présentée par

Michel NASSIF

Thèse préparée au CERMICS, École des Ponts ParisTech

Arbres aléatoires : asymptotique de fonctionnelles et limites locales

---

Soutenance le 5 Décembre 2022 devant le jury composé de :

Romain ABRAHAM	Université d'Orléans	*Directeur de thèse*
Jean-François DELMAS	École des Ponts ParisTech	*Directeur de thèse*
Thomas DUQUESNE	Sorbonne Université	*Rapporteur*
Mihai GRADINARU	Université Rennes 1	*Président du jury*
Bénédicte HAAS	Université Sorbonne Paris Nord	*Examinatrice*
Pascal MAILLARD	Université Paul Sabatier	*Examinateur*
Amandine VÉBER	Université Paris Cité	*Examinatrice*

Après avis des rapporteurs :

Thomas DUQUESNE	Sorbonne Université
Jason SCHWEINSBERG	UC San Diego

بكرة بيخلص هالكابوس
وبدل الشمس بتضوي شموس.

Remerciements

Alors que le soleil se couche sur ce voyage de trois ans, je suis rempli de joie et de gratitude pour tout ce que j'ai vécu mais aussi d'excitation pour ce qui est à venir.
Je tiens tout d'abord à remercier mes deux directeurs de thèse, Romain Abraham et Jean-François Delmas, d'avoir accepté d'encadrer ma thèse et de m'avoir proposé des sujets riches et passionnants. Merci à eux de m'avoir transmis leurs connaissances et de m'avoir fait décourvrir la recherche avec leur pédagogie mais aussi leur intuition et leur élégance mathématique. Je suis reconnaissant pour tout le temps qu'ils m'ont consacré, pour leur investissement et leur patience. Ils ont su – malgré quelques difficultés – me guider durant ces trois années avec beaucoup d'humanité et d'exigence et me mener à ce point final.
Je suis honoré que Thomas Duquesne et Jason Schweinsberg aient accepté la lourde tâche de lire et d'évaluer ce travail. Je remercie Bénédicte Haas, Pascal Maillard et Amandine Véber d'avoir gentiment accepté de faire partie du jury. Je remercie chaleureusement Mihai Gradinaru de m'avoir donné le goût des probabilités et de m'avoir guidé en particulier dans la recherche d'une thèse. Je remercie également Jean Bertoin de m'avoir accueilli à Zurich pendant un mois. Merci pour toutes les discussions intéressantes tant sur le plan mathématique que sur le plan humain.
Mes trois années de thèse ont été particulièrement agréables grâce au cadre de travail du Cermics. Je remercie notamment Aurélien Alfonsi et Frédéric Meunier de m'avoir fait confiance pour enseigner leurs cours. Merci à Isabelle Simunic et Stéphanie Bonnel pour leur efficacité redoutable et leur présence rassurante tout le long.
Il est maintenant temps de remercier les doctorant·e·s du Cermics. Parce que la recherche peut être gratifiante, mais cela peut aussi donner un sentiment de solitude et d'isolement (et les confinements n'ont pas vraiment aidé). Pour cette raison, je suis sincèrement reconnaissant à toutes les personnes qui m'ont empêché de travailler, consciemment j'imagine. En essayant de saboter ma carrière, elles ont apporté des fous rires et des conversations stimulantes dans mon quotidien. Je ne saurais trop insister sur la joie de prendre le RER en sachant ce qui m'attendait : je pense aux midis à la cantine et aux gâteaux du goûter, mais aussi aux pauses continues et éternelles du deuxième et la bonne ambiance générale. J'ai la chance d'avoir rencontré plein de gens pleins de vie, avec qui on a rigolé, discuté, joué, appris des langues[1], chanté et dansé. Grâce à elles, j'ai grandi. J'adresse donc mes remerciements les plus chaleureux à Edoardo, pour son rire éclatant et son empathie réconfortante; Zoé, qui met de la joie dans la vie; Emanuele, la plus belle des lasagnes, le mec le plus intelligent (paraît-il) et le plus drôle (à

[1] زبالة

vérifier) qui est remonté dans mon classement le jour où il m'a fait un gâteau; Roberta, pour sa gentillesse et sa façon singulière de voir la vie, pour être toujours à l'écoute et pour m'avoir offert un paquet de chips à un moment où je n'allais pas bien, enfin parce que le labo est moins bien quand elle n'est pas là; Rémi[1], pour avoir été présent pendant toute cette aventure et pour nos discussions mathématiques souvent infructueuses; Hélène[1], ma bff dès son premier jour, pour toutes les fois où je suis venu à son bureau ou elle au mien, pour tous les moments passés ensemble à ne rien faire. Merci également à Nerea ma cinema mate, Clément mon demi-frère de thèse, Jean pour sa sagesse, Kexin pour sa générosité, Laurent pour sa théâtralité, Rutger pour son humour, Cyrille pour sa sympathie et ses chocolats, Maël pour toutes les fois où il m'a fait du pied, Gaspard et Inass pour les parties de Sushi Go! et de Kingdomino indispensables pour survivre le confinement, Éloïse pour ses appels à la pause toujours opportuns, Alfred pour son extravagance, mais aussi à Hervé, Régis, Léo, Julien, Camila, Louis, Kacem, Stefano, Renato et je n'oublie pas les anciens Sébastien et Thomas. Merci à toutes ces personnes et certainement d'autres que j'ai oubliées d'avoir partagé avec moi leur humour, leur joie de vivre et leur énergie.

De l'autre côté de l'avenue Blaise Pascal, je remercie mes amis du LAMA qui ont été d'un grand soutien. Merci à Elias pour son énergie, son empathie et ses histoires passionantes. Merci à Quentin, Josué et Ahmed, en particulier pour cette semaine qu'on a passée ensemble à Oléron, pour nos jeux et nos échanges enrichissants.

À l'autre bout du RER A, ces trois dernières années parisiennes ont été bonnes pour moi, et j'ai la chance d'avoir été entouré d'ami·e·s – certain·e·s je connaissais déjà, d'autres j'ai rencontré·e·s plus récemment – qui ont contribué à les rendre agréables et passionnantes. Ils et elles ont tou·te·s contribué à leur manière à ce voyage, et j'ai l'impression d'avoir beaucoup grandi et beaucoup appris. Je tiens à les remercier pour avoir fait de moi une version plus aboutie de moi-même. Merci à mes amis de Rennes : Adrian[2] le plus gentil, Roméo[3] le plus cool, Émeric et Lucien et nos week-ends nostalgiques. Merci à Rudy qui nous rassemble depuis six ans déjà, pour nos soirées de grand rassemblement, pour toutes ces pizzas (resp. bières) mangées (resp. bues) sur les quais avec Rémi. Merci à Corentin pour avoir toujours été présent malgré la distance, pour toutes les fois où il est passé me voir à Paris, pour nos discussions souvent drôles et parfois éclairantes[4] et pour toutes nos passions communes. Merci[5] à Shaza et Nathalie pour avoir été là, pour avoir été un espace safe et un soutien émotionnel sans égal. Je suis certain[6] que Paris n'aurait pas été aussi bien sans vous. Merci à Nathalie pour sa sagesse, sa bienveillance et son vocabulaire de grand-mère que personne ne comprend. Merci à Shaza notre extravertie parce qu'elle me comprend sans que j'aie besoin de parler, pour sa spontanéité et sa curiosité et pour toutes nos aventures parisiennes. Merci infiniment à Hélène la meilleure coloc, pour tout ce qu'on a vécu et partagé ensemble, nos brunchs et recettes pompettes, nos voyages au Liechtenstein et ailleurs. Merci pour ses histoires émouvantes et sa narration captivante, son honnêteté admirable, son sens de l'aventure et surtout de la vie.

Je termine en remerciant ma famille et mes proches pour leur soutien indéfectible et pour tous les moments précieux passés ensemble. Merci à ma mamie Laïla pour ses encouragements et son amour. Merci à Eva et à mes cousines Shery et Rola pour nos rassemblements ressourçants pleins de joie et d'énergie. Merci à mes parents pour leur soutien moral et financier depuis

[1] Voir plus bas aussi
[2] et Amandine
[3] et Camille
[4] eye-opening
[5] حب كتير
[6] et je n'ai pas envie d'essayer

toutes ces années. Grâce à eux, j'ai pu quitter mon Caire natal – il y a neuf ans déjà – et venir en France faire mes études supérieures. Je me sens privilégié d'avoir vécu cette expérience formatrice, non seulement sur le plan professionnel mais aussi humain. Enfin, merci à ma sœur Caroline[1] pour sa présence irremplaçable même aux moments les plus difficiles et son écoute toujours précieuse.

[1] كوكي

Abstract

This thesis is devoted to the study of some asymptotic properties of Bienaymé-Galton-Watson (BGW) trees and Lévy trees. BGW trees encode the genealogical structure of BGW processes which describe the evolution of a population whose individuals reproduce asexually and independently of each other. Lévy trees are the continuous analogues of BGW trees: they emerge as the scaling limits of the latter and encode the genealogical structure of continuous-state branching processes.

First, we study very general additive functionals of size-conditioned BGW trees whose offspring distribution is critical and lies in the domain of attraction of a stable law. We show that in the so-called global regime, when properly rescaled, they converge to functionals of a normalized stable Lévy tree. For functionals depending only on the size and height of the tree, we describe a phase transition using an integral test.

Next, we study the shape of normalized stable Lévy trees near their root. We show that, when zooming in at the root at the proper speed, we get the immortal tree which consists of an infinite branch onto which trees are grafted according to a Poisson point measure which does not depend on the initial normalization. We apply this result to study the asymptotic behavior of the aforementioned functionals of a normalized stable Lévy tree and we identify two regimes in which either the size or the height dominates the other.

Finally, we study the maximal degree of critical and subcritical Lévy trees. We establish a Poissonian decomposition of the tree along its large nodes and we determine the genealogical structure of those nodes. Furthermore, we make sense of the distribution of the Lévy tree conditioned to have a fixed maximal degree. We apply this to study the local limit of the Lévy tree conditioned on having large maximal degree. We show that a condensation phenomenon occurs in the subcritical case, whereas there is local convergence to the immortal tree in the critical case.

Keywords: Galton-Watson trees, Lévy trees, additive functionals, scaling limit, random measure, phase transition, immortal tree, maximal degree, local limit, condensation.

2020MSC classifications: 60J80, 05C05, 60F17, 60G52, 60G55, 60G57, 60J25.

Résumé

Les arbres aléatoires apparaissent dans des contextes variés : en informatique, en biologie ou encore en chimie organique. Cette thèse porte sur l'étude de quelques propriétés asymptotiques de deux modèles d'arbres aléatoires : les arbres de Bienaymé-Galton-Watson (BGW) ainsi que leurs limites d'échelle continues qui sont les arbres de Lévy. Les arbres de BGW encodent la structure généalogique des processus de BGW qui décrivent l'évolution de la taille d'une population dont les individus se reproduisent asexuellement et indépendamment les uns des autres. Plus précisement, on fixe une variable aléatoire ξ à valeurs dans $\mathbb{N}$ appelée la loi de reproduction. L'arbre de BGW de loi de reproduction ξ peut alors être construit de façon récursive comme suit. On commence avec un individu (la racine) qui donne naissance à un nombre aléatoire d'enfants distribué selon ξ. Aux générations suivantes, on réitère ce procédé : chaque individu donne naissance à un nombre aléatoire d'enfants de loi ξ indépendamment de tous les autres. Cela permet de définir un arbre aléatoire, l'arbre de BGW de loi de reproduction ξ, dont la taille (c'est-à-dire le nombre de nœuds) est aléatoire. On note alors τ^n l'arbre de BGW conditionné à avoir exactement n nœuds. Si l'on suppose que la loi de reproduction ξ est critique (c'est-à-dire $\mathbb{E}[\xi] = 1$) et qu'elle appartient au domaine d'attraction d'une loi stable d'indice $\gamma \in (1,2]$ (c'est-à-dire qu'il existe une suite $(b_n, n \geq 1)$ de réels positifs tels que $b_n^{-1}(\sum_{k=1}^n \xi_k - n)$ converge en loi vers une variable aléatoire X_1 de transformée de Laplace $\mathbb{E}\left[\exp(-\lambda X_1)\right] = \exp(\lambda^\gamma)$, où $(\xi_n, n \geq 1)$ est une suite de variables aléatoires indépendantes de loi ξ) , alors il est bien connu que l'arbre de BGW conditionné à avoir n nœuds, vu comme un espace métrique mesuré aléatoire muni de la distance de graphe et de la mesure de probabilité uniforme sur l'ensemble des nœuds, converge en loi après renormalisation vers l'arbre stable $\mathscr{T}$. Ce dernier est un espace métrique aléatoire (ou plus précisément un arbre réel aléatoire), muni d'un point distingué $\emptyset$ appelé la racine et d'une mesure de probabilité μ (la loi uniforme sur l'ensemble des feuilles).

Dans le premier chapitre, on étudie le comportement asymptotique des fonctionnelles additives sur les arbres de BGW conditionnés. Une fonctionnelle F définie sur l'ensemble des arbres est dite additive si elle est de la forme :

$$F(\mathbf{t}) = \sum_{w \in \mathbf{t}} f(\mathbf{t}_w),$$

où $\mathbf{t}_w$ est le sous-arbre de $\mathbf{t}$ au-dessus du nœud w et f est une certaine fonction de coût. Ces fonctionnelles apparaissent naturellement en informatique pour représenter le coût des algorithmes de type "diviser pour régner", en phylogénétique pour évaluer l'équilibre d'un arbre

ou encore en chimie comme un outil de prédiction de certaines propriétés chimiques d'une molécule. Des exemples de fonctionnelles additives incluent la longueur de cheminement total :

$$P(\mathbf{t}) = \sum_{v\in\mathbf{t}} d(\emptyset, v) = \sum_{w\in\mathbf{t}} |\mathbf{t}_w| - |\mathbf{t}|,$$

où $|\mathbf{t}_w|$ est le nombre de nœuds de $\mathbf{t}_w$, d est la distance de graphe et $\emptyset$ est la racine de $\mathbf{t}$, l'indice de Wiener :

$$W(\mathbf{t}) = \sum_{u,v\in\mathbf{t}} d(u,v) = 2|\mathbf{t}| \sum_{w\in\mathbf{t}} |\mathbf{t}_w| - 2\sum_{w\in\mathbf{t}} |\mathbf{t}_w|^2,$$

et l'indice B_1 de Shao et Sokal :

$$B_1(\mathbf{t}) = \sum_{\substack{w\in\mathbf{t}^\circ \\ w\neq\emptyset}} \frac{1}{\mathfrak{h}(\mathbf{t}_w)},$$

où $\mathbf{t}^\circ$ est l'ensemble des nœuds internes de $\mathbf{t}$ et $\mathfrak{h}(\mathbf{t}_w)$ est la hauteur du sous-arbre $\mathbf{t}_w$. L'idée, qui trouve son origine dans l'article [52], consiste à voir une fonctionnelle additive $\sum_{w\in\mathbf{t}} f(\mathbf{t}_w)$ comme l'intégrale de la fonction f par rapport à une certaine mesure associée à l'arbre $\mathbf{t}$. Plus précisément, on associe à tout arbre réel T, muni d'une distance d, d'une racine $\emptyset$ et d'une mesure finie μ, une mesure Ψ_T définie par :

$$\langle \Psi_T, f\rangle = \int_T \mu(\mathrm{d}x) \int_0^{d(\emptyset,x)} f(T_{r,x})\,\mathrm{d}r,$$

où $T_{r,x}$ est le sous-arbre de T au-dessus du niveau r qui contient le nœud x. On montre alors que $\langle \Psi_T, f\rangle$ est une approximation de la fonctionnelle additive associée à la fonction de coût f et que l'application $T \mapsto \Psi_T$ est continue. Ce résultat purement analytique permet alors de déduire la convergence des fonctionnelles additives sur les arbres de BGW conditionnés vers des fonctionnelles de l'arbre stable lorsque la fonction de coût f est suffisamment régulière. Notons que cette méthode est robuste puisqu'elle s'applique à toute suite d'arbres aléatoires qui possède une limite d'échelle continue. On cherche ensuite à améliorer ce résultat de convergence dans le cas particulier des arbres de BGW conditionnés en prenant des fonctions de coût singulières. On se concentre sur des fonctions de coût de la forme $f(\mathbf{t}) = |\mathbf{t}|^a \mathfrak{h}(\mathbf{t})^b$ où a et b sont deux réels quelconques et on montre le résultat suivant.

Théorème. *Supposons que la loi de reproduction ξ est critique et qu'elle appartient au domaine d'attraction d'une loi stable d'indice $\gamma \in (1,2]$, avec une suite renormalisante $(b_n, n\geq 1)$ qui satisfait $\underline{b} \leq n^{-1/\gamma} b_n \leq \overline{b}$ où $\underline{b}$ et $\overline{b}$ sont deux réels strictement positifs. Soit τ^n l'arbre de BGW de loi de reproduction ξ conditionné à avoir n nœuds.*

(i) Si $\gamma a + (\gamma-1)b > 1$, on a la convergence en loi :

$$\frac{b_n^{1+b}}{n^{1+a+b}} \sum_{w\in\tau^{n,\circ}} |\tau^n_w|^a \mathfrak{h}(\tau^n_w)^b \xrightarrow[n\to\infty]{(d)} \mathbf{Z}^\gamma_{a-1,b},$$

où la variable aléatoire limite $\mathbf{Z}^{\gamma}_{a,b}$ *admet la représentation suivante en terme de l'arbre stable* $\mathscr{T}$ *:*

$$\mathbf{Z}^{\gamma}_{a,b} = \int_{\mathscr{T}} \mu(\mathrm{d}x) \int_0^{d(\emptyset,x)} \mu(\mathscr{T}_{r,x})^a \mathfrak{h}(\mathscr{T}_{r,x})^b \,\mathrm{d}r,$$

où $\mathscr{T}_{r,x}$ *est le sous-arbre de* $\mathscr{T}$ *au-dessus du niveau* r *qui contient la feuille* x*,* $\mu(\mathscr{T}_{r,x})$ *est sa masse et* $\mathfrak{h}(\mathscr{T}_{r,x})$ *sa hauteur.*

(ii) Si $\gamma a + (\gamma - 1) b \leq 1$*, on a la convergence en loi :*

$$\frac{b_n^{1+b}}{n^{1+a+b}} \sum_{w \in \tau^{n,\circ}} |\tau^n_w|^a \mathfrak{h}(\tau^n_w)^b \xrightarrow[n\to\infty]{(d)} \infty.$$

En particulier, ce théorème met en évidence une transition de phase en $\gamma a + (\gamma - 1) b = 1$. Ce résultat est complété par l'étude de la variable aléatoire limite $\mathbf{Z}^{\gamma}_{a,b}$ pour laquelle on observe la même transition de phase.

Proposition. *Soit* $\gamma \in (1,2]$ *et soient* $a, b \in \mathbb{R}$*. On a l'alternative suivante :*

$$\gamma a + (\gamma-1) b > 1 - \gamma \iff \mathbf{Z}^{\gamma}_{a,b} < \infty \text{ a.s.} \iff \mathbb{E}\left[\mathbf{Z}^{\gamma}_{a,b}\right] < \infty,$$
$$\gamma a + (\gamma-1) b \leq 1 - \gamma \iff \mathbf{Z}^{\gamma}_{a,b} = \infty \text{ a.s.} \iff \mathbb{E}\left[\mathbf{Z}^{\gamma}_{a,b}\right] = \infty.$$

De plus, l'espérance de $\mathbf{Z}^{\gamma}_{a,b}$ admet une expression explicite en fonction des moments de la hauteur de l'arbre stable.
Dans le deuxième chapitre, on étudie le comportement asymptotique de la fonctionnelle $\mathbf{Z}^{\gamma}_{a,b}$ sur l'arbre stable lorsque $\max(a,b)$ tend vers l'infini. Cela nécessite de comprendre la géométrie locale de l'arbre stable au voisinage de sa racine. En effet, on montre qu'en se rapprochant de la racine à la bonne vitesse, on obtient un arbre de Kesten formé d'une branche infinie sur laquelle sont greffés des arbres stables (non normalisés) selon une mesure ponctuelle de Poisson. Plus précisément, on choisit une feuille U uniformément au hasard dans l'arbre stable $\mathscr{T}$ et on considère la branche reliant la feuille U à la racine $\emptyset$. On note $\mathscr{T}_i$, $i \in I_U$ les sous-arbres de $\mathscr{T}$ greffés sur cette branche, chacun à une hauteur h_i. On fixe une fonction $\mathfrak{f}\colon (0,\infty) \to (0,\infty)$ qui représente la vitesse à laquelle on se rapproche de la racine et on définit pour tout $\varepsilon > 0$ une mesure ponctuelle $\mathscr{M}^{\mathfrak{f}}_{\varepsilon}(U)$ par :

$$\mathscr{M}^{\mathfrak{f}}_{\varepsilon}(U) = \sum_{h_i \leq \mathfrak{f}(\varepsilon) d(\emptyset,U)} \delta_{(\varepsilon^{-1} h_i, R_\gamma(\mathscr{T}_i, \varepsilon^{-1}))},$$

où $R_\gamma(\mathscr{T}_i, \varepsilon^{-1})$ est une renormalisation de l'arbre $\mathscr{T}_i$. Ainsi, la mesure ponctuelle $\mathscr{M}^{\mathfrak{f}}_{\varepsilon}(U)$ décrit la lignée ancestrale de la feuille U à partir de la racine et jusqu'à une certaine hauteur $\mathfrak{f}(\varepsilon) d(\emptyset, U)$. En d'autres termes, cela revient à faire un zoom sur la racine à la vitesse $\mathfrak{f}(\varepsilon)$ tout en renormalisant en même temps.

Théorème. *Soit* $\mathscr{T}$ *l'arbre stable d'indice* $\gamma \in (1,2]$.

(i) *Si* $\lim_{\varepsilon\to 0} \varepsilon^{-1/2}\mathfrak{f}(\varepsilon) = 0$ *et* $\lim_{\varepsilon\to 0} \varepsilon^{-1}\mathfrak{f}(\varepsilon) = \infty$, *alors on a la convergence en loi :*

$$\mathcal{M}_\varepsilon^{\mathfrak{f}}(U) \xrightarrow[\varepsilon\to 0]{(d)} \sum_{s\geq 0} \delta_{(s,\mathsf{T}_s)},$$

au sens de la convergence vague des mesures, où $\sum_{s\geq 0}\delta_{(s,\mathsf{T}_s)}$ *est une mesure ponctuelle de Poisson indépendante de* $(\mathcal{T}, d(\emptyset, U))$.

(ii) *Si* $\mathfrak{f}(\varepsilon) = \varepsilon$, *alors on a la convergence en loi :*

$$\mathcal{M}_\varepsilon^{\mathfrak{f}}(U) \xrightarrow[\varepsilon\to 0]{(d)} \sum_{s\leq d(\emptyset,U)} \delta_{(s,\mathsf{T}_s)},$$

où $\sum_{s\geq 0}\delta_{(s,\mathsf{T}_s)}$ *est une mesure ponctuelle de Poisson indépendante de* $(\mathcal{T}, d(\emptyset, U))$.

En application de ce théorème, on peut décrire le comportement asymptotique de $\mathbf{Z}^\gamma_{a,b}$ lorsque $\max(a,b)$ tend vers l'infini. On distingue deux régimes selon la vitesse de $b/a^{1-1/\gamma}$.

Théorème. *Soit* $\mathcal{T}$ *l'arbre stable d'indice* $\gamma \in (1,2]$.

(i) *Supposons que* $a \to \infty$, $b \geq 0$ *et* $b/a^{1-1/\gamma} \to c \in \mathbb{R}_+$. *Alors, on a la convergence en loi :*

$$\lim_{a\to\infty} a^{1-1/\gamma}\mathfrak{h}(\mathcal{T})^{-b}\mathbf{Z}^\gamma_{a,b} = \int_0^\infty \mathrm{e}^{-S_t - ct/\mathfrak{h}(\mathcal{T})}\,\mathrm{d}t,$$

où $(S_t,\, t\geq 0)$ *est un subordinateur stable de transformée de Laplace* $\mathbb{E}\left[\exp(-\lambda S_t)\right] = \exp(-t\gamma\lambda^{1-1/\gamma})$, *indépendent de* $\mathcal{T}$.

(ii) *Supposons que* $b \to \infty$, $a \geq 0$ *et* $a^{1-1/\gamma}/b \to 0$. *Alors, on a la convergence en probabilité :*

$$\lim_{b\to\infty} b\mathfrak{h}(\mathcal{T})^{-b}\mathbf{Z}^\gamma_{a,b} = \mathfrak{h}(\mathcal{T}).$$

Dans le régime $b/a^{1-1/\gamma} \to c$ la masse prédomine, ce qui explique l'apparition du subordinateur à la limite. Au contraire, dans le régime $b/a^{1-1/\gamma} \to \infty$, la hauteur prédomine et il n'y a plus de subordinateur à la limite.

Dans le troisième chapitre, on étudie le conditionnement d'un arbre de Lévy par son degré maximal. Les arbres de Lévy forment une famille importante d'arbres réels aléatoires qui contient les arbres stables. Dans un certain sens, ils codent la structure généalogique des processus de branchement à espace d'état continu et constituent ainsi les limites d'échelle possibles des arbres de BGW. Tout comme les processus de branchement, la loi d'un arbre de Lévy est entièrement caractérisée par une fonction qui s'appelle le mécanisme de branchement et qui admet la forme de Lévy-Khintchine suivante :

$$\psi(\lambda) = \alpha\lambda + \beta\lambda^2 + \int_{(0,\infty)} (\mathrm{e}^{-\lambda r} - 1 + \lambda r)\,\pi(\mathrm{d}r), \quad \forall \lambda \in \mathbb{R}_+,$$

où $\alpha \in \mathbb{R}$, $\beta \in \mathbb{R}_+$ et π est une mesure σ-finie sur $(0,\infty)$ qui vérifie $\int_{(0,\infty)}(r\wedge r^2)\,\pi(\mathrm{d}r) < \infty$. On dit que l'arbre de Lévy est critique (resp. sous-critique) si $\alpha = 0$ (resp. $\alpha > 0$). La fonction ψ est

en particulier l'exposant de Laplace d'un certain processus de Lévy X sans sauts négatifs et c'est grâce à ce processus X que l'on peut construire l'arbre de Lévy. Il est bien connu qu'un point de branchement x de l'arbre de Lévy est soit binaire (cela est dû à la partie brownienne $\beta\lambda^2$), soit de degré infini (cela est dû à la mesure de Lévy π). Dans ce dernier cas, le nœud x correspond exactement à un saut du processus de Lévy sous-jacent X dont la taille Δ_x est prise comme définition du degré généralisé du nœud x. De façon plus intrinsèque à l'arbre, on peut montrer que Δ_x est la limite correctement renormalisée du nombre $n(x,\varepsilon)$ de sous-arbres au-dessus du nœud x dont la hauteur est plus grande que ε. Le degré maximal Δ de l'arbre de Lévy est alors défini comme le supremum du degré généralisé Δ_x lorsque x parcourt l'ensemble des nœuds de l'arbre. Dans la littérature, plusieurs décompositions de l'arbre de Lévy ont été obtenues : citons la décomposition de Bismut le long de la lignée ancestrale d'une feuille choisie uniformément au hasard [58], la décomposition de Williams le long de la lignée ancestrale de la feuille la plus haute [2] ou encore la décomposition le long du diamètre de l'arbre [60]. Ici, on établit une nouvelle décomposition de l'arbre de Lévy $\mathscr{T}$ le long de ses gros nœuds. Plus précisément, on fixe un seuil $\delta > 0$ et on s'intéresse à l'arbre réduit selon les nœuds de degré plus grand que δ. On montre que l'arbre élagué $\mathscr{T}^\delta$ obtenu à partir de l'arbre $\mathscr{T}$ en effaçant ces nœuds est à nouveau un arbre de Lévy dont le mécanisme de branchement est changé. De plus, on décrit comment obtenir l'arbre de départ $\mathscr{T}$ en greffant de façon poissonnienne des arbres de loi connue sur l'arbre élagué $\mathscr{T}^\delta$. Ensuite, cette décomposition sert à donner un sens à l'arbre de Lévy conditionné à un degré maximal fixé.

Théorème. *Supposons que $\delta > 0$ n'est pas un atome de la mesure de Lévy π. Alors, conditionnellement à $\Delta = \delta$, l'arbre de Lévy peut être construit de la façon suivante :*

(i) prendre un arbre de Lévy biaisé $\widetilde{\mathscr{T}}$ par la taille de degré maximal inférieur à δ,

(ii) choisir une feuille x au hasard dans l'arbre $\widetilde{\mathscr{T}}$,

(iii) à cette feuille x, greffer un arbre de Lévy de degré initial δ conditionné à ne pas avoir d'autres nœuds de degré supérieur à δ.

En application de ce théorème, on peut obtenir le comportement asymptotique de l'arbre de Lévy conditionné à avoir un degré maximal tendant vers l'infini. On voit apparaître deux comportements radicalement différents. Dans le cas sous-critique, on observe un phénomène de condensation : un nœud de degré infini apparaît à une hauteur finie. Cela constitue le premier résultat de convergence vers un arbre de condensation continu. Dans le cas critique, le nœud de degré infini part à l'infini et il y a alors convergence vers un arbre de Kesten formé d'une branche infinie sur laquelle sont greffés des arbres de Lévy selon une mesure ponctuelle de Poisson.

Contents

Introduction

The goal of this thesis is to study some functionals on two related models of random trees: Bienaymé-Galton-Watson (or BGW for short) trees and Lévy trees. It contains three papers.

- [7]: *Global regime for general additive functionals of conditioned Bienaymé-Galton-Watson trees,* with R. Abraham and J.-F. Delmas, Probab. Theory Related Fields **182** (2022), no. 1-2, 277-351.
- [128]: *Zooming in at the root of the stable tree,* Electron. J. Probab. **27**: 1-38 (2022).
- [6]: *Conditioning (sub)critical Lévy trees by their maximal degree: decomposition and local limit,* with R. Abraham and J.-F. Delmas, to be submitted.

This introductory chapter is divided into three parts. In Section 0.1, we give a historical overview (without any pretense of exhaustion) of BGW trees and Lévy trees to motivate our work. We also point out some connections to other fields of probability. In Section 0.2, we introduce the main mathematical objects we consider: in particular, we define the two models of random trees that we will be studying. Finally, Section 0.3 presents the main contributions of this thesis.

0.1 Historical context

0.1.1 The BGW process

In probability, a branching process is a stochastic process used to model a population which evolves randomly in time. As its name suggests, a key feature of such process is the so-called branching property, which roughly speaking means that two disjoint subpopulations evolve independently of each other. The simplest example of a branching process is the BGW process, which corresponds to the following population model: start with one individual; at generation n, each individual has a random number of children distributed according to some probability distribution (called the offspring distribution), independently of all others. The origin of this model goes back to the 19th century where it was introduced by Galton and Watson [75] to study the extinction of family names. While their method based on generating functions is correct, they concluded erroneously that the population will always go exctinct almost surely. Much later, in 1972, Heyde and Senata [89] discovered a note written by Bienaymé in 1845 where he correctly states that the extinction probability is equal to 1 if and only if the mean of the offspring distribution is at most 1. Bienaymé gave some explanations, but no proof was ever found. We refer the reader to Kendall [102, 103] and [25] for a historical background. Ever since, the long-time behavior of BGW processes has received a lot of attention, the classical reference on the subject being the monograph of Athreya and Ney [23]. In particular, a celebrated result by Lamperti [114] in 1967 identifies all possible scaling limits of BGW processes. It turns out that these are exactly the continuous-state branching (CB for short) processes introduced by Jiřina [97] in 1958. CB processes are the continuous (in space and time) analogues of BGW processes. Roughly speaking, they are $\mathbb{R}_+$-valued Markov processes that describe the evolution of a continuous population and that enjoy the branching property. Without too much foreshadowing, let us mention that Lamperti [113] provided a relationship between CB processes and spectrally positive Lévy processes: more specifically, there is a one-to-one correspondance via a random time change which is commonly referred to as the Lamperti transformation. We shall revisit this relationship between the two families through a different lens when we discuss Lévy trees.

0.1.2 Scaling limits of BGW trees

When dealing with population dynamics, it is natural to study the behavior of the whole genealogy rather than just a single generation at a time. The genealogy of a BGW process can be represented by a random plane (i.e. rooted and ordered) tree, called the BGW tree, which is finite if and only if the population goes extinct. The size of the tree (which corresponds then to the size of the population) is random. In applications, one usually possesses some information concerning the size of the population; different notions of size can be interesting such as the total number of individuals but also the number of "leaves". Therefore, it is often more natural to consider BGW trees conditioned by their size. This family of random trees is particularly rich as one can recover the uniform distribution on different classes of "combinatorial trees" simply by changing the offspring distribution. Of particular interest was the behavior of BGW trees conditioned to be large in some sense (for example large size, large height, etc). Up to the 1990's, this was done through a case by case study where only certain statistics of the tree were considered, see e.g. Takács [148]. It was Aldous [16, 17, 18] who first suggested to study the scaling limits of large random trees as a whole. More specifically, he showed that when the offspring distribution has mean 1 and finite variance, conditioned BGW trees, considered as (random) compact subsets of the space ℓ^1 of summable sequences, converge in distribution after rescaling toward a (random) compact subset called the Brownian tree. One central idea emerging from this work was the advantage of viewing random trees as random metric spaces and to study their scaling limits with respect to some variant of the Hausdorff topology on compact subsets of a metric space. Aldous also gave several constructions of the Brownian tree: via the coding by a normalized Brownian excursion, through its finite-dimensional marginals, or through a line-breaking construction. There was another major development in this field in 2003 when Evans, Pitman and Winter [70] suggested to use the formalism of real trees – introduced earlier for geometric and algebraic purposes – together with the Gromov-Hausdorff topology, see e.g. [133]. This point of view is inspired by Aldous' idea of considering random trees as random metric spaces; however, it is intrinsic as it does not require embedding the tree into ℓ^1. As such, it is now widely used in the field of random combinatorial structures and gives a powerful framework for studying scaling limits.

0.1.3 Universality of the Brownian tree

The importance of the Brownian tree stems from the fact that it is the scaling limit of a large class of random trees such as unordered binary trees [123], uniform unordered trees [82, 131], critical multitype BGW trees [126] and some random trees with a prescribed degree sequence [43]. It is also the scaling limit of several models of large random graphs which are not trees including random dissections [47], random quadrangulations with a large boundary [35], random outerplanar maps [44, 145], random bipartite maps with one macroscopic face [96] and subcritical random graphs [132]. In addition to combinatorial motivations, the Brownian tree has applications in statistical physics for its connection with random surfaces. Indeed, one important question to physicists is whether there is a "uniform" 2-dimensional surface; see e.g. Ambjørn, Durhuus and Jonsson [22]. One way to answer this question is to consider a discrete version of random surfaces, namely random planar maps, and to study their scaling limits. In the seminal work [45], Chassaing and Schaeffer established a bijection between random maps and labelled trees, the so-called Cori-Vauquelin-Schaeffer (CVS) bijection. Then, it makes sense that the "uniform" 2-dimensional surface is coded by the "uniform" labelled tree through the CVS bijection. This is the construction given in 2007 by Le Gall [117] of the so-called Brownian map in terms of the Brownian tree with Brownian labels. The Brownian map was then shown to be the scaling limit of a large class of random maps, see e.g. [127, 118, 36, 11, 12].

0.1.4 Superprocesses

A superprocess is a measure-valued Markov process which combines the branching structure of a CB process with a spatial motion given by some Markov process, see e.g. Dawson [48], Perkins [134] and references therein. In other words, individuals do not only reproduce but they also move in space independently according the same law. The superprocess at a given time t is then a random measure describing the positions of individuals alive at time t. It has been known for a long time that superprocesses are related to partial differential equations, see Dynkin [64, 65], and stochastic nonlinear partial differential equations, see the survey [49] by Dawson and Perkins. The super-Brownian motion is a special instance of a superprocess where the spatial motion is given by a linear Brownian motion and the genealogical structure is given by the Brownian tree. In [57], Duquesne and Le Gall make a connection between the Brownian tree and superprocesses. More precisely, they explain how to combine the genealogical structure of the Brownian tree with a Brownian motion to obtain a super-Brownian motion, see also Etheridge [67, Chapter 3].

0.1.5 Lévy trees

Lévy trees were introduced by Le Gall and Le Jan [119] and Duquesne and Le Gall [57] as a generalization of Aldous' Brownian tree. They are coded by the so-called height process, which is a local time functional of a spectrally positive Lévy process. Duquesne and Winkel [62] provided an alternative construction which does not make use of the height process: they obtain Lévy trees as the limit of a growing family of BGW trees with edge lengths which is consistent under Bernoulli percolation on leaves. Lévy trees constitute the possible scaling limits of BGW trees just like CB processes are the scaling limits of BGW processes. Furthermore, a generalization of the celebrated second Ray-Knight theorem states that the process describing the size of the population alive at a given "time" in a Lévy tree is in fact a CB process. This justifies that Lévy trees encode the genealogy of CB processes. Since their introduction, Lévy trees have received a lot of attention: Duquesne and Le Gall [58] studied their fractal properties and proved that they enjoy a branching property similar to that of BGW trees which states that the subtrees above a given level are independent and distributed as the original tree; they also showed in [59] that Lévy trees are invariant under re-rooting. Their importance is also due to the fact that they are the building blocks used to construct scaling limits of some models of random graphs, see e.g. Broutin, Duquesne and Wang [41, 42]. The success of Lévy trees has also led to a couple of interesting generalizations. We mention Duquesne's continuum random trees with immigration [56] which encode the genealogy of CB processes with immigration, as well as the recent work of Berestycki, Fittipaldi and Fontbona [28] and also Li, Pardoux and Wakolbinger [120] providing a description of the genealogy of a CB process with interaction which is motivated by biological applications.

0.1.6 Stable trees and coalescent processes

Stable trees constitute a one-parameter subfamily of Lévy trees indexed by $\gamma \in (1,2]$ that includes Aldous' Brownian tree (the latter corresponding to $\gamma = 2$). Duquesne [54] showed that a conditioned BGW tree whose offspring distribution belongs to the domain of attraction of a stable law with index γ converges to the γ-stable tree when the size goes to infinity. Stable trees enjoy some remarkable properties such as self-similarity. Curien and Haas [46] showed that stable trees are nested in the sense that inside the γ-stable tree one can find a rescaled version of the γ'-stable tree for every $1 < \gamma < \gamma' \leq 2$. In [77], Goldschmidt and Haas gave a line-breaking construction of the stable tree which generalizes Aldous' construction of the Brownian tree. Stable trees are also intimately related to coalescent and fragmentation processes, see Bertoin [33] for a general introduction on the subject. Coalescent processes go back to 1982 when Kingman [105, 106] introduced his coalescent; then Pitman [135] and Sagitov [142] generalized this model to the so-called Λ-coalescent. They are now widely used in population genetics to study the genealogy of branching processes backward in time: in other words, the idea is to trace back the ancestral lines of individuals (or species to be more exact) to see where they diverged, see e.g. Tavaré [149]. Berestycki, Berestycki and Schweinsberg [27] proved that the Beta-coalescent (which is a special case of the Λ-coalescent) can be embedded in the stable tree. Similarly, Berestycki and Berestycki [26] gave a construction of Kingman's coalescent in terms of the Brownian tree. In both cases, those embeddings were applied to obtain new results about the coalescents using fine properties of stable trees.

0.1.7 Fragmentation processes and trees

Time-reversing a coalescent process gives a fragmentation process: this is a Markov process that describes how an object with given total mass evolves as it breaks into several fragments randomly as time passes. Fragmentation processes were introduced by Bertoin [30, 31, 32]. The first connection between fragmentation processes and continuum random trees was established implicitly by Aldous and Pitman [19] who showed that one can obtain the 1/2-self-similar fragmentation process by splitting the Brownian tree in a Poissonian fashion along its skeleton; see also Bertoin [30] for a somewhat simpler construction of the Aldous-Pitman fragmentation process. Miermont [125] obtained a self-similar fragmentation process in a similar way from the stable tree. In this case, some care needs to be taken as there is a fundamental difference between the Brownian tree and the non-Brownian stable trees: indeed, the latter contain nodes with infinite degree (these are absent from the Brownian tree) that need to be removed. More generally, Abraham and Delmas [1] and Voisin [152] studied fragmentation processes associated with the Lévy tree by placing marks both on the skeleton and on infinite branching nodes.

Instead of splitting along certain nodes, one can also fragment the tree by simply discarding all the nodes located under a certain height t. This was first studied in the case of the Brownian tree by Bertoin [32] who obtained a striking connection with the Aldous-Pitman fragmentation: the only difference between the two is the speed at which fragments decay. This "duality" was later extended to the non-Brownian stable case by Miermont [124, 125]. Delmas [51] then studied this fragmentation at height in the general Lévy case with no Brownian part.

Conversely, there is a natural genealogical structure associated with any fragmentation process: the common ancestor of two fragments is simply the block that contained both of them for the last time, before a dislocation event had separated them. Haas and Miermont [81] showed that one can encode the genealogy of a self-similar fragmentation process with negative index by a continuum random tree, such that one recovers the original fragmentation process by splitting the tree at height. This provides another model of continuum random trees, called fragmentation trees, different from Lévy trees, with the intersection of those two models being exactly the stable trees. This competing model of continuum random trees has also received a lot of attention: Haas [80] studied the asymptotic behavior of a fragmentation tree when the distances between nodes converge to infinity (or, equivalently, a fragmentation process with initial mass converging to infinity); Haas, Miermont, Pitman and Winkel [83] showed that fragmentation trees arise as the scaling limits of discrete models of random trees which notably include Aldous' beta-splitting model and Ford's alpha model; Haas and Miermont [82] also proved that Markov branching trees converge after suitable rescaling to self-similar fragmentation trees.

0.2 BGW trees and Lévy trees

In this section, we shall introduce the main mathematical objects that we will handle, namely BGW trees and Lévy trees, and see how they relate to each other.

0.2.1 Discrete trees

A discrete tree is a connected acyclic graph. The trees that we will consider will be finite, rooted (i.e. they have a distinguished vertex called the root) and ordered (i.e. children of each vertex are ordered from left to right). Such trees are also called plane trees in the litterature. We shall use Neveu's formalism for discrete trees, see [129]. Let $\mathcal{U} = \cup_{n\geq 0}(\mathbb{N}^*)^n$ be the set of labels, with the convention $(\mathbb{N}^*)^0 = \{\emptyset\}$. In other words, an element of $\mathcal{U}$ is a (possibly empty) finite sequence of positive integers. If $v = (v^1, \dots, v^n) \in \mathcal{U}$, we denote by $H(v) = n$ its "height" or generation. By convention, we set $H(\emptyset) = 0$. If $v = (v^1, \dots, v^n)$, $w = (w^1, \dots, w^m)\ in \mathcal{U}$, we write $vw = (v^1, \dots, v^n, w^1, \dots, w^m)$ for the concatenation of v and w. In particular, we have $v\emptyset = \emptyset v = v$. We say that v is an ancestor of w and write $v \preccurlyeq w$ if there exists $u \in \mathcal{U}$ such that $w = vu$. Define a mapping $\mathrm{pr}\colon \mathcal{U} \setminus \{\emptyset\} \to \mathcal{U}$ by $\mathrm{pr}(v^1, \dots, v^n) = (v^1, \dots, v^{n-1})$ (i.e. $\mathrm{pr}(v)$ is the parent of v). Finally, a discrete tree is a *finite* subset of $\mathcal{U}$ satisfying the following conditions:

(i) $\emptyset \in \mathbf{t}$,

(ii) $v \in \mathbf{t} \setminus \{\emptyset\} \implies \mathrm{pr}(v) \in \mathbf{t}$,

(iii) for every $v \in \mathbf{t}$, there exists a finite integer $k_v(\mathbf{t}) \geq 0$ such that, for every $j \in \mathbb{N}^*$, $vj \in \mathbf{t}$ if and only if $1 \leq j \leq k_v(\mathbf{t})$.

The number $k_v(\mathbf{t})$ should be interpreted as the number of children of the vertex v in $\mathbf{t}$ and $\mathrm{pr}(v)$ is its parent. The vertex $\emptyset$ is called the root of $\mathbf{t}$. The vertex v is called a leaf if $k_v(\mathbf{t}) = 0$ and an internal vertex otherwise. We denote the set of leaves by $\mathrm{Lf}(\mathbf{t})$ and the set of internal vertices by $\mathbf{t}^\circ$. If $v \in \mathbf{t}$, we define the subtree $\mathbf{t}_v$ of $\mathbf{t}$ above v (rooted at v) by:

$$\mathbf{t}_v = \{w \in \mathcal{U}\colon\ vw \in \mathbf{t}\}.$$

We denote by $|\mathbf{t}| = \mathrm{Card}(\mathbf{t})$ the size of $\mathbf{t}$ and by $\mathfrak{h}(\mathbf{t}) = \max_{v\in\mathbf{t}} H(v)$ its height.

We now introduce a way of coding discrete trees. For a discrete tree $\mathbf{t}$, denote by $v_0 = \emptyset, v_1, \dots, v_{|\mathbf{t}|-1}$ the vertices of $\mathbf{t}$ listed in lexicographical order (which corresponds to the depthfirst search of the tree). The height function $H_{\mathbf{t}} = (H_{\mathbf{t}}(n) : 0 \le n < |\mathbf{t}|)$ is defined by:

$$H_{\mathbf{t}}(n) = H(v_n), \quad \forall 0 \le n < |\mathbf{t}|.$$

In words, the height function is simply the sequence of the generations of the individuals of $\mathbf{t}$

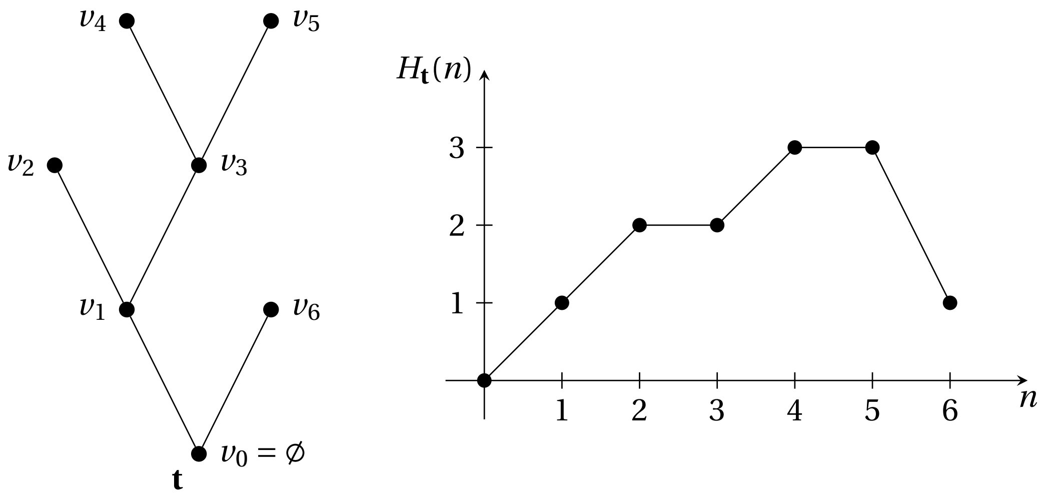

Figure 1 – Example of a discrete tree (left) and its associated height function (right).

listed in the lexicographical order, see Figure 1. It is easily checked that the height function $H_{\mathbf{t}}$ characterizes the tree $\mathbf{t}$.

0.2.2 BGW trees

A BGW tree is the genealogical tree of a BGW process. It corresponds to the evolution of a population in which each individual has, independently of the others, a random number of children distributed according to the same probability measure on $\mathbb{N}$.

Definition 0.2.1. *Let ξ be a non-constant $\mathbb{N}$-valued random variable with mean 1 (ξ is said to be critical). A random variable τ with values in the set of discrete trees is said to be a BGW tree with offspring distribution ξ (or BGW(ξ) tree for short) if:*

(i) $\mathbb{P}\left(k_{\varnothing}(\tau) = k\right) = \mathbb{P}(\xi = k)$ *for* $k \in \mathbb{N}$.

(ii) For every $k \ge 1$ with $\mathbb{P}(\xi = k) > 0$, conditionally on $k_{\varnothing}(\tau) = k$, the subtrees above the root $\tau_1, \dots, \tau_k$ are independent and distributed as τ.

We exclude the case $\xi = 1$: this corresponds to a (deterministic) population in which each individual has exactly one child. Equivalently, we always assume that $\mathbb{P}(\xi = 0) > 0$. Property (ii) is called the branching property of the BGW tree. One can show that this distribution exists and that it is uniquely determined by the offspring distribution. In fact, we have the following characterization: the random tree τ is a BGW(ξ) tree if and only if

$$\mathbb{P}(\tau = \mathbf{t}) = \prod_{v \in \mathbf{t}} \mathbb{P}(\xi = k_v(\mathbf{t})), \tag{0.2.1}$$

for every discrete tree $\mathbf{t}$.

Remark 0.2.2. The definition above still works when ξ is subcritical (i.e. $\mathbb{E}[\xi] < 1$). However, when ξ is supercritical (i.e. $\mathbb{E}[\xi] > 1$), some care needs to be taken as the BGW tree is infinite with positive probability. Let us mention that it is still possible to define the supercritical BGW tree in that case, but we do not enter such considerations as we will only deal with the critical case.

We now turn to the coding of BGW trees. Consider the height function H_τ of the BGW tree τ. In general, this random process is not Markovian, but Le Gall and Le Jan [119] noticed that it can be written as a simple functional of a random walk.

Proposition 0.2.3. *Let τ be a BGW(ξ) tree and let $(R_n, n \in \mathbb{N})$ be a random walk on $\mathbb{Z}$ with initial value $R_0 = 0$ and jump distribution $\mathbb{P}(R_1 = k) = \mathbb{P}(\xi = k+1)$ for every $k \geq -1$. Set $T = \inf\{n \geq 1 : R_n = -1\}$ and*

$$K_n = \operatorname{Card}\left\{0 \leq k \leq n-1 : R_k = \inf_{k \leq j \leq n} R_j\right\}. \tag{0.2.2}$$

Then, we have the following identity:

$$(H_\tau(n) : 0 \leq n < |\tau|) \overset{(d)}{=} (K_n : 0 \leq n < T). \tag{0.2.3}$$

This will serve as motivation for the coding of Lévy trees, see Section 0.2.6 below.

0.2.3 Size-conditioned BGW trees

The size of a BGW tree is random. More precisely, if we let $(R_n, n \in \mathbb{N})$ be a random walk starting from 0 with jump distribution $\mathbb{P}(R_1 = k) = \mathbb{P}(\xi = k+1)$, the well-known Otter-Dwass formula (see Otter [130], Dwass [63] and Pitman [137, Chapter 6]) writes:

$$\mathbb{P}(|\tau| = n) = \frac{1}{n}\mathbb{P}(R_n = -1). \tag{0.2.4}$$

Remark 0.2.4. In fact, the preceding identity can be recovered from (0.2.3), which implies that $|\tau|$ is distributed as the hitting time T of -1 by the random walk $(R_n, n \in \mathbb{N})$, together with Kemperman's formula [100, 101], which identifies the distribution of T.

Recall that the span of the integer-valued random variable ξ is the largest integer λ_0 such that $\mathbb{P}(\xi \in a + \lambda_0\mathbb{Z}) = 1$ for some $a \in \mathbb{Z}$. Since we assume that $\mathbb{P}(\xi = 0) > 0$, the span λ_0 is also the greatest common divisor of $\{k \geq 1 : \mathbb{P}(\xi = k) > 0\}$. In particular, using (0.2.4), one can show that $\mathbb{P}(|\tau| = n) = 0$ if $n \not\equiv 1 \pmod{\lambda_0}$ while $\mathbb{P}(|\tau| = n) > 0$ for all large n with $n \equiv 1 \pmod{\lambda_0}$, see Section 1.4 of Chapter 1.

For every n such that $\mathbb{P}(|\tau| = n) > 0$, we denote by τ^n a BGW(ξ) tree conditioned to have size n. In other words, the tree τ^n is distributed as τ conditionally on $|\tau| = n$. In what follows, we implicitly assume that n is such that $\mathbb{P}(|\tau| = n) > 0$.

Remark 0.2.5. For special choices of offspring distribution ξ, the size-conditioned BGW(ξ) tree is uniformly distributed on a class of "combinatorial trees" with n vertices, see Table 1. Let us give some examples. We say that ξ follows the geometric distribution with parameter $1/2$ and write $\xi \sim \mathcal{G}(1/2)$ if $\mathbb{P}(\xi = k) = 2^{-k-1}$ for every $k \in \mathbb{N}$. In that case, the conditioned BGW(ξ) tree τ^n is uniformly distributed over the set of all discrete (i.e. rooted ordered) trees with n vertices. If $\xi \sim \mathcal{P}(1)$ has a Poisson distribution with parameter 1 (i.e. $\mathbb{P}(\xi = k) = \mathrm{e}^{-1}/k!$), then τ^n is uniformly distributed over the set of all rooted Cayley (i.e. rooted unordered) trees with n vertices. Finally, if $\mathbb{P}(\xi = 0) = \mathbb{P}(\xi = 2) = 1/2$, then τ^n is uniformly distributed over the set of complete binary trees with n vertices.

Class of trees	Rooted ordered trees	Rooted Cayley trees	Complete binary trees
Offspring distribution	$\xi \sim \mathcal{G}(1/2)$	$\xi \sim \mathcal{P}(1)$	$\mathbb{P}(\xi = 0) = \mathbb{P}(\xi = 2) = 1/2$
Span	1	1	2
Variance	2	1	1

Table 1 – Uniform trees viewed as conditioned BGW trees

0.2.4 Real trees

Real trees (also known as $\mathbb{R}$-trees) are abstract metric spaces that generalize discrete trees in a continuous way. Take a discrete tree and embed it in the plane, allowing edges to have different lengths. Now, instead of thinking of the tree as just its finite set of vertices with a collection of distances between them, regard the edges as also part of the metric space. In other words, elements of the tree are not only the "vertices" of the discrete tree but also all the points lying in between. This space one obtains, formed by a union of line segments, should be a real tree. In fact, the definition of a real tree will allow for more erratic behavior.

Definition 0.2.6. *A quadruple $(T, d, \varnothing, \mu)$ is called a real tree (resp. a compact real tree) if (T, d) is a metric space (resp. a compact metric space) equipped with a distinguished vertex $\varnothing \in T$ called the root and a nonnegative finite measure μ on T and if the following two properties hold for every $x, y \in T$:*

(i) (Unique geodesics). There exists a unique isometric map $f_{x,y} \colon [0, d(x,y)] \to T$ such that $f_{x,y}(0) = x$ and $f_{x,y}(d(x,y)) = y$.

(ii) (Loop-free). If φ is a continuous injective map from $[0,1]$ *into* T *such that* $\varphi(0) = x$ *and* $\varphi(1) = y$*, then we have* $\varphi([0,1]) = f_{x,y}\big([0, d(x,y)]\big)$.

Property (ii) says that there is a unique path joining any two nodes in a real tree and this is what makes it "tree-like". For example, the space $\mathbb{R}^2$ – which is certainly not tree-like – satisfies property (i) but not (ii) since there are infinitely many paths between any two points in the plane.

Remark 0.2.7. In the litterature, a real tree is generally defined as a metric space (T,d) satisfying properties (i) and (ii) above without being equipped with a distinguished vertex $\emptyset$ or a measure μ. In this terminology, the quadruple $(T,\emptyset,d,\mu)$ is usually called a *rooted measured* real tree. However, we choose to include the root and the measure in the definition of a real tree since the trees that we will consider (viz. BGW trees and Lévy trees) are naturally equipped with this structure. In particular, the measure μ gives an intrisic way of choosing a vertex of T *uniformly* at random. In that sense, our definition is in the spirit of Aldous' definition of a continuum real tree; see [16].

Remark 0.2.8. We will need to view discrete trees as real trees. Let $a > 0$. We simply embed the discrete tree $\mathbf{t}$ in the plane, connect every vertex to its children in such a way that the distance between any two adjacent vertices (of the discrete tree) is a and equip this resulting metric space with the uniform probability measure on the set of vertices of the discrete tree. We denote this real tree by $a\mathbf{t}$.

If $(T,d,\emptyset,\mu)$ is a real tree, we define its set of leaves by:

$$\mathrm{Lf}(T) = \{x \in T \setminus \{\emptyset\} :\ T \setminus \{x\} \text{ is connected}\},$$

with the convention that $\mathrm{Lf}(T) = \{\emptyset\}$ if $T = \{\emptyset\}$ is the trivial tree. The height of the tree T is defined by $\mathfrak{h}(T) = \sup_{x\in T} H(x)$, where $H(x) = d(\emptyset, x)$ denotes the height of x. Note that if $(T,d,\emptyset,\mu)$ is a compact real tree, then $\mathfrak{h}(T) < \infty$. The range of the mapping $f_{x,y}$ described in (i) above is denoted by $[\![x,y]\!]$ (this is the line segment between x and y in the tree). In particular, $[\![\emptyset,x]\!]$ is the path going from the root to x, which we will interpret as the ancestral line of the node x. We define a partial order on the tree by setting $x \preccurlyeq y$ (x is an ancestor of y) if and only if $x \in [\![\emptyset,y]\!]$. If $x,y \in T$, there is a unique $z \in T$ such that $[\![\emptyset,x]\!] \cap [\![\emptyset,y]\!] = [\![\emptyset,z]\!]$. We write $z = x \wedge y$ and call it the most recent common ancestor of x and y. Let $x \in T$ be a node. Let $r \in [0, H(x)]$. We denote by $x_r \in T$ the unique ancestor of x with height $H(x_r) = r$. As in the discrete case, we also define the subtree T_x of T above x by:

$$T_x = \left\{y \in T :\ x \preccurlyeq y\right\},$$

and the subtree $T_{r,x}$ of T above level r containing x by:

$$T_{r,x} = \left\{y \in T :\ H(x \wedge y) \geq r\right\} = T_{x_r}.$$

Then T_x (resp. $T_{r,x}$) can be naturally viewed as a real tree, rooted at x (resp. at x_r) and endowed with the distance d and the measure $\mu_{|T_x} = \mu(\cdot \cap T_x)$ (resp. the measure $\mu_{|T_{r,x}}$). Note that $T_{0,x} = T$ and $T_{H(x),x} = T_x$.

Similarly to the discrete case, there is a way to code a real tree via excursion-type functions which we now present, see e.g. [69, Chapter 3, Example 3.14]. Let e be a positive excursion, that is $e\colon \mathbb{R}_+ \to \mathbb{R}_+$ is a continuous function such that $e(0) = 0$, $e(s) > 0$ for $0 < s < \sigma$ and $e(s) = 0$ for $s \geq \sigma$, where $\sigma = \inf\{s > 0\colon\ e(s) = 0\} \in (0, \infty)$ is the lifetime of the excursion. Set $d_e(t,s) = e(t) + e(s) - 2\inf_{[t\wedge s, t\vee s]} e$ for every $t, s \in [0,\sigma]$ and define an equivalence relation on $[0,\sigma]$ by letting $t \sim_e s$ if and only if $d_e(t,s) = 0$. Then the real tree T_e coded by the excursion e is defined as the quotient space $[0,\sigma]/\sim_e$, rooted at $p_e(0)$ where $p_e\colon [0,\sigma] \to T_e$ is the quotient map and equipped with the distance d_e and the pushforward measure $\mu_e = \lambda \circ p_e^{-1}$ where λ is the Lebesgue measure $[0,\sigma]$. Informally, this passage to the quotient corresponds to putting

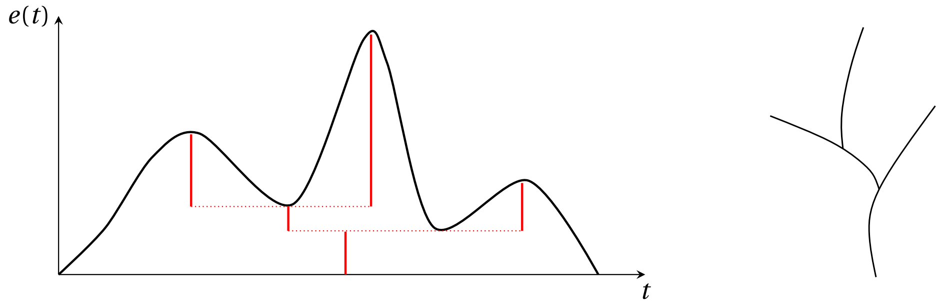

Figure 2 – Example of an excursion (left) and the real tree that it codes (right).

some glue on the bottom of the graph of e then folding it horizontally; see Figure 2. It can be checked that this defines a compact real tree. Notice that the total mass of T_e is given by $\mu(T_e) = \sigma$ and its height by $\mathfrak{h}(T_e) = \sup_{s\geq 0} e(s)$.

Remark 0.2.9. Conversely, every compact real tree $(T, \emptyset, d)$ (notice that we do not assume that T is equipped with a measure) can be coded by a continuous excursion. Furthermore, every compact (measured) real tree $(T, \emptyset, d, \mu)$ equipped with a linear order $\leq$ that satisfies some natural compatibility assumptions can be coded by a càglàd excursion. We refer the reader to Duquesne [55] for these results and for futher details on the coding of real trees.

0.2.5 Convergence of real trees

One of our goals is to study the convergence of (rescaled) discrete trees to continuous ones. To this end, it is necessary to have a notion of convergence for real trees. We shall use a variant of the Gromov-Hausdorff distance between compact metric spaces (which was introduced by Gromov [79] for geometric applications) which takes into account that the metric spaces we consider are marked and measured.

Let $(T,\emptyset,d,\mu)$ and $(T',\emptyset',d',\mu')$ be two compact real trees. We say that a mapping $\varphi\colon (T,\emptyset,d,\mu) \to (T',\emptyset',d',\mu')$ is a root-preserving and measure-preserving isometry if $\varphi\colon (T,d)\to(T',d')$ is an isometry such that $\varphi(\emptyset)=\emptyset'$ and $\mu\circ\varphi^{-1}=\mu'$. Denote by $\mathbb{T}$ the set of root-preserving and measure-preserving isometry classes of *compact* real trees. We will often identify a class with an element of this class, writing $(T,\emptyset,d,\mu)$ (or simply T when there is no ambiguity) for the class of $(T,\emptyset,d,\mu)$ in $\mathbb{T}$. We start by recalling the definition of the Hausdorff distance. Let (E,δ) be a metric space. Given a non-empty subset $A\subset E$ and $\varepsilon>0$, the ε-neighborhood of A is defined by $A^\varepsilon=\{x\in E\colon\ \delta(x,A)<\varepsilon\}$. The Hausdorff distance δ_{H} between two non-empty subsets $A,B\subset E$ is defined by:

$$\delta_{\mathrm{H}}(A,B)=\inf\left\{\varepsilon>0\colon\ A\subset B^\varepsilon \text{ and } B\subset A^\varepsilon\right\}.$$

Next, we give the definition of the Lévy-Prokhorov distance which metrizes the topology of weak convergence of finite (nonnegative) measures. The Lévy-Prokhorov distance between two finite measures μ,ν on E is defined by:

$$\delta_{\mathrm{P}}(\mu,\nu)=\inf\left\{\varepsilon>0\colon\ \mu(A)\le\nu(A^\varepsilon)+\varepsilon \text{ and } \nu(A)\le\mu(A^\varepsilon)+\varepsilon,\ \forall A\subset E \text{ measurable}\right\}.$$

We can now give the standard definition of the Gromov-Hausdorff-Prokhorov (or GHP for short) topology. Given two compact real trees $(T,\emptyset,d,\mu),(T',\emptyset',d',\mu')\in\mathbb{T}$, we set:

$$d^\circ_{\mathrm{GHP}}(T,T')=\inf\left\{\delta(\varphi(\emptyset),\varphi'(\emptyset'))\vee\delta_{\mathrm{H}}(\varphi(T),\varphi'(T'))\vee\delta_{\mathrm{P}}(\mu\circ\varphi^{-1},\mu'\circ\varphi'^{-1})\right\},$$

where the infimum is taken over all isometries $\varphi\colon T\to E$ and $\varphi'\colon T'\to E$ into a common metric space (E,δ). It can be checked that the space $(\mathbb{T},d^\circ_{\mathrm{GHP}})$ is a Polish metric space, see e.g. [5].

Remark 0.2.10. Other topologies on the set of compact real trees (and more generally, compact metric measure spaces) have appeared in the litterature. As mentioned by Villani [151], there are essentially two approaches. The first one – which we follow – consists in combining the Hausdorff metric with the Prokhorov metric to compare both geometric and measure-related properties of the spaces. The second approach is to emphasize the role of the measures. This idea originates from the work of Gromov [79] where it is induced by the so-called box-metrics. Greven, Pfaffelhuber and Winter [78] introduced the so-called Gromov-weak topology which can be defined by the Gromov-Prokhorov distance. Later, it was shown by Löhr [121] that Gromov's box metric and the Gromov-Prokhorov metric are bi-Lipschitz equivalent. In a sense, this topology is weaker than the GHP topology since we do not take into account the Hausdorff distance between the two spaces. When working with this metric, the features outside the support of the measure in the underlying space are discarded which leads to different isometry classes. Building on the Gromov-weak topology, Athreya, Löhr and Winter introduced the stronger Gromov-Hausdorff-weak topology which additionally compares the Hausdorff distance of the supports of the respective measures. This is more closely related to the GHP topology; we refer the reader to [5, 24, 104] for further discussion of the two topologies.

0.2.6 Lévy trees

Following Duquesne and Le Gall [57, 58], we will define Lévy trees as the compact real trees coded by the so-called height process that we now introduce. Just like CB processes, the distribution of the height process is characterized by a function $\psi \colon \mathbb{R}_+ \to \mathbb{R}_+$ called the *branching mechanism* and having the Lévy-Khintchine form:

$$\psi(\lambda) = \alpha\lambda + \beta\lambda^2 + \int_{(0,\infty)} (\mathrm{e}^{-\lambda r} - 1 + \lambda r)\,\pi(\mathrm{d}r), \quad \forall \lambda \in \mathbb{R}_+, \tag{0.2.5}$$

where $\alpha \in \mathbb{R}$, $\beta \in \mathbb{R}_+$ and π is a σ-finite measure on $(0,\infty)$ satisfying the integrability condition $\int_{(0,\infty)} (r \wedge r^2)\,\pi(\mathrm{d}r) < \infty$. We say that the branching mechanism ψ is critical (resp. subcritical, resp. supercritical) if $\alpha = 0$ (resp. $\alpha > 0$, resp. $\alpha < 0$). We will restrict ourselves to the critical and subcritical cases. We also assume that the Grey condition holds:

$$\int^{\infty} \frac{\mathrm{d}\lambda}{\psi(\lambda)} < \infty. \tag{0.2.6}$$

Our starting point is a spectrally positive (i.e. without negative jumps) Lévy process $X = (X_t,\, t \geq 0)$ with Laplace exponent ψ starting from 0. Namely, we have:

$$\mathbb{E}\left[\exp(-\lambda X_t)\right] = \exp(t\psi(\lambda)), \quad \forall t, \lambda \in \mathbb{R}_+.$$

Note that (0.2.5) entails that X does not drift to ∞, see e.g. Bertoin [29, Chapter VII]. Furthermore, the Grey condition (0.2.6) implies that

$$\beta > 0 \quad \text{or} \quad \int_{(0,1)} r\,\pi(\mathrm{d}r) = \infty, \tag{0.2.7}$$

which is equivalent to the paths of X being a.s. of infinite variation, see Bertoin [29]. In analogy with (0.2.2), the idea is to define the height process as the "measure" of the set

$$\left\{ s \leq t \colon\ X_s = \inf_{[s,t]} X \right\}$$

in a local time sense. Notice that here the Lévy process X plays the role of the random walk R. In this direction, one can show that there exists a process $H = (H_t,\, t \geq 0)$ with continuous paths such that the following convergence holds in probability for every $t \geq 0$:

$$H_t = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_0^t \mathbf{1}_{\{I_t^s < X_s < I_t^s + \varepsilon\}}\,\mathrm{d}s, \tag{0.2.8}$$

where $I_t^s = \inf_{[s,t]} X$ is the past infimum of X. The process H is called the ψ-height process. In the Brownian case $\psi(\lambda) = \lambda^2$, it can be checked that H is distributed as a reflected Brownian motion. Heuristically, the height process codes a Lévy forest. In order to construct the Lévy tree, we first need to introduce the excursion measure of the height process.

To that end, denote by $I = (I_t,\ t \ge 0)$ the infimum process of X defined by $I_t = \inf_{[0,t]} X$. Basic results on Lévy processes, see e.g. [29, Chapter VII], ensure that $X - I$ is a strong Markov process with values in $\mathbb{R}_+$ and that the point 0 is regular. Furthermore, the process $-I$ is a local time at 0 for $X - I$. Denote by $\mathbf{N}^\psi$ the associated excursion measure of the process $X - I$ away from 0. In the Brownian case, $\mathbf{N}^\psi$ is the Itô positive excursion measure up to a scaling factor. Now it is not difficult to derive from (0.2.8) that the value H_t of the height process at time t only depends on the excursion of $X - I$ away from 0 which straddles time t. Furthermore, the excursion intervals of H away from 0 coincide with those of $X - I$ away from 0. It follows that the excursion measure of H away from 0 is the "distribution" of H under $\mathbf{N}^\psi$. We still denote it by $\mathbf{N}^\psi$ and we let

$$\sigma = \inf\{t > 0 :\ H_t = 0\} \tag{0.2.9}$$

be the lifetime of H under $\mathbf{N}^\psi$ (which coincides with the lifetime of $X - I$). We also let

$$\mathfrak{h} = \sup_{0 \le t \le \sigma} H_t. \tag{0.2.10}$$

Definition 0.2.11. *The ψ-Lévy tree $\mathcal{T} = T_H$ is the compact real tree coded by the ψ-height process H under $\mathbf{N}^\psi$.*

With a slight abuse of notation, we denote by $\mathbf{N}^\psi(\mathrm{d}\mathcal{T})$ the "distribution" of the ψ-Lévy tree under $\mathbf{N}^\psi(\mathrm{d}H)$. Observe that under $\mathbf{N}^\psi$, the Lévy tree $\mathcal{T}$ has total mass σ and height $\mathfrak{h}$.

Remark 0.2.12. The notation H_t for the height process is consistent with the height (i.e. distance to the root) in the Lévy tree $\mathcal{T}$. Indeed, recall from Section 0.2.4 that $p_H : [0,\sigma] \to \mathcal{T}$ is the canonical projection. Then, if we denote by $x = p_H(t)$ the node of $\mathcal{T}$ associated with t, it is easy to see that the height of x satisfies $H(x) = H_t$.

Remark 0.2.13. There are some other ways to define Lévy trees. Most notably, Duquesne and Le Gall [57, Chapter 2] showed that Lévy trees arise as scaling limits of BGW trees. Alternatively, Duquesne and Winkel [62] constructed Lévy trees as limits of growing families of BGW trees with exponential edge lengths, consistent under Bernoulli percolation on leaves.

We gather some properties of the Lévy tree under its excursion measure $\mathbf{N}^\psi$.

(i) **Height.** For $\mathbf{N}^\psi$-almost every $\mathcal{T}$, there is a unique leaf $x^* \in \mathcal{T}$ realizing the height, i.e. such that $H(x^*) = \mathfrak{h}$. Furthermore, the height $\mathfrak{h}$ of $\mathcal{T}$ satisfies:

$$\mathbf{N}^\psi[\mathfrak{h} > a] = v(a), \tag{0.2.11}$$

where the function $v : (0,\infty) \to (0,\infty)$ is the unique nonnegative solution of the equation:

$$\int_{v(a)}^\infty \frac{\mathrm{d}\lambda}{\psi(\lambda)} = a. \tag{0.2.12}$$

(ii) **Mass measure.** For $\mathbf{N}^\psi$-almost every $\mathscr{T}$, the mass measure μ is supported on the set of leaves $\mathrm{Lf}(\mathscr{T})$. Furthermore, the total mass $\sigma = \mu(\mathscr{T})$ satisfies:

$$\mathbf{N}^\psi\left[1 - \mathrm{e}^{-\lambda\sigma}\right] = \psi^{-1}(\lambda), \tag{0.2.13}$$

where ψ^{-1} is the right-continuous inverse of ψ.

(iii) **Local times.** For $\mathbf{N}^\psi$-almost every $\mathscr{T}$, there exists a process $(L^a,\, a \geq 0)$ with values in the space of finite measures on $\mathscr{T}$ which is càdlàg for the topology of weak convergence and such that

$$\mu(\mathrm{d}x) = \int_0^\infty \mathrm{d}a\, L^a(\mathrm{d}x). \tag{0.2.14}$$

For every $a \geq 0$, the measure L^a is supported on $\mathscr{T}(a) := \{x \in \mathscr{T} :\ H(x) = a\}$ the set of nodes at height a. Furthermore, the real-valued process $\left(L^a_\sigma := \langle L^a, 1\rangle,\, a \geq 0\right)$ is a CB process with branching mechanism ψ under its canonical measure.

(iv) **Branching property.** For every $a \geq 0$, let $(\mathscr{T}^i,\, i \in I_a)$ be the subtrees of $\mathscr{T}$ originating from level a and let x_i be their respective roots. Then, under $\mathbf{N}^\psi$ and conditionally on $r_a(\mathscr{T}) := \{x \in \mathscr{T} :\ H(x) \leq a\}$, the measure $\sum_{i \in I_a} \delta_{(x_i, \mathscr{T}^i)}$ is a Poisson point measure on $\mathscr{T}(a) \times \mathbb{T}$ with intensity $L^a(\mathrm{d}x)\,\mathbf{N}^\psi(\mathrm{d}\mathscr{T}')$.

(v) **Branching points.** For $\mathbf{N}^\psi$-almost every $\mathscr{T}$, the branching points of $\mathscr{T}$ are either binary or of infinite degree. The set of binary branching points is empty if $\beta = 0$ and is a countable dense subset of $\mathscr{T}$ if $\beta > 0$. The set

$$\mathrm{Br}_\infty(\mathscr{T}) := \{x \in \mathscr{T} :\ \mathscr{T} \setminus \{x\} \text{ has infinitely many connected components}\}$$

of infinite branching points is nonempty with $\mathbf{N}^\psi$-positive measure if and only if $\pi \neq 0$. If $\langle \pi, 1\rangle = \infty$, the set $\mathrm{Br}_\infty(\mathscr{T})$ is countable and dense in $\mathscr{T}$ for $\mathbf{N}^\psi$-almost every $\mathscr{T}$. Furthermore, the set $\{H(x),\, x \in \mathrm{Br}_\infty(\mathscr{T})\}$ coincides with the set of discontinuity times of the mapping $a \mapsto L^a$. For every such discontinuity time a, there is a unique $x_a \in \mathrm{Br}_\infty(\mathscr{T}) \cap \mathscr{T}(a)$ and $\Delta_a > 0$ such that

$$L^a = L^{a-} + \Delta_a \delta_{x_a}.$$

For convenience, we define Δ_a for every $a \geq 0$ by setting $\Delta_a = 0$ if $L^a = L^{a-}$. In particular, we have $L^a_\sigma = L^{a-}_\sigma + \Delta_a$, that is Δ_a is exactly the size of the jump of the associated CB process at time a. We will call Δ_a the degree (or the mass) of the node x_a. This is an abuse of language since a node $x_a \in \mathrm{Br}_\infty(\mathscr{T})$ has infinite degree by definition.

We end this section with a useful decomposition of the Lévy tree. Bismut [39] gave a description of the Itô positive excursion measure at a time chosen uniformly at random. The next theorem extends this result to the excursion measure $\mathbf{N}^\psi$. First, let us introduce some notations. If $T \in \mathbb{T}$ is a compact real tree and $x \in T$, we denote by $(T^i,\, i \in I_x)$ the subtrees originating from the branch $[\![\emptyset, x]\!]$ and let h_i be the height at which T^i is grafted. Define a point measure on

$\mathbb{R}_+ \times \mathbb{T}$ by:

$$\mathcal{M}_x^T = \sum_{i \in I_x} \delta_{(h_i, T^i)}.$$

We will also need the probability measure $\mathbb{P}_r^\psi(\mathrm{d}\mathcal{T})$ defined on $\mathbb{T}$ which is the distribution of a "Lévy forest" with initial degree r, see e.g. Abraham and Delmas [3] for a precise definition.

Definition 0.2.14. *Let $\sum_{i\in I} \delta_{\mathcal{T}_i}$ be a Poisson point measure on $\mathbb{T}$ with intensity $r\,\mathbf{N}^\psi$. Then $\mathbb{P}_r^\psi$ is defined as the distribution of the random tree $\mathcal{T}$ obtained by gluing together the trees $\mathcal{T}_i$ at their root.*

Finally, define:

$$\mathbf{N}_{\mathrm{B}}^\psi(\mathrm{d}\mathcal{T}) = 2\beta \mathbf{N}^\psi(\mathrm{d}\mathcal{T}) + \int_{(0,\infty)} r\,\pi(\mathrm{d}r)\,\mathbb{P}_r^\psi(\mathrm{d}\mathcal{T}). \tag{0.2.15}$$

The following theorem gives a decomposition of the Lévy tree along the ancestral line of a leaf chosen uniformly at random; see Duquesne and Le Gall [58, Theorem 4.5] or Abraham and Delmas [3, Theorem 2.1].

Theorem 0.2.15. *For every measurable function $\Phi\colon \mathbb{R}_+ \times \mathbb{T} \to \mathbb{R}_+$ and for every $\lambda \geq 0$, we have:*

$$\mathbf{N}^\psi\left[\int_{\mathcal{T}} L^a(\mathrm{d}x)\,\mathrm{e}^{-\lambda H(x) - \langle \mathcal{M}_x^{\mathcal{T}}, \Phi\rangle}\right] = \mathrm{e}^{-\lambda a} \exp\left(-\int_0^a g(u)\,\mathrm{d}u\right), \tag{0.2.16}$$

where we set:

$$g(u) = \alpha + \mathbf{N}_{\mathrm{B}}^\psi\left[1 - \mathrm{e}^{-\Phi(u,\mathcal{T})}\right]. \tag{0.2.17}$$

In particular, we have:

$$\mathbf{N}^\psi\left[\int_{\mathcal{T}} \mu(\mathrm{d}x)\,\mathrm{e}^{-\lambda H(x) - \langle \mathcal{M}_x^{\mathcal{T}}, \Phi\rangle}\right] = \int_0^\infty \mathrm{d}a\,\mathrm{e}^{-\lambda a} \exp\left(-\int_0^a g(u)\,\mathrm{d}u\right). \tag{0.2.18}$$

In other words, under $\mathbf{N}^\psi[\sigma\,\mathrm{d}\mathcal{T}]$, if we choose a leaf U uniformly at random (i.e. according to the normalized mass measure $\sigma^{-1}\mu$), its height $H(U)$ has distribution $\mathrm{e}^{-\alpha a}\,\mathrm{d}a$ and, conditionally on $H(U) = a$, the point measure $\mathcal{M}_U^{\mathcal{T}}$ is a Poisson point measure on $[0, a]$ with intensity $\mathbf{N}_{\mathrm{B}}^\psi$. Let us mention that there are other known decompositions for the Lévy tree: a Williams' decomposition along the ancestral line of the heighest leaf (see Abraham and Delmas [2]) and a decomposition along the diameter of the Lévy tree (see Duquesne and Wang [60]).

0.2.7 Stable trees

Stable trees are special instances of Lévy trees that enjoy remarkable scaling properties. We say that ψ is a stable branching mechanism if it is of the form (up to a multiplicative constant):

$$\psi(\lambda) = \lambda^\gamma, \quad \forall \lambda \in \mathbb{R}_+, \tag{0.2.19}$$

where $\gamma \in (1,2]$. We will write $\mathbf{N}^\gamma$ for the excursion measure of the stable tree instead of $\mathbf{N}^\psi$. Observe that the Brownian case $\psi(\lambda) = \lambda^2$ corresponds to $\alpha = 0$, $\beta = 1$ and $\pi = 0$ while the non-Brownian stable case $\psi(\lambda) = \lambda^\gamma$ with $\gamma \in (1,2)$ corresponds to $\alpha = \beta = 0$ and

$$\pi(\mathrm{d}r) = \frac{\gamma(\gamma-1)}{\Gamma(2-\gamma)} \frac{\mathrm{d}r}{r^{1+\gamma}}, \tag{0.2.20}$$

where Γ is Euler's gamma function.

When the branching mechanism is stable, the Lévy process X is self-similar. More precisely, the process $(a^{-1/\gamma} X_{at},\, t \geq 0)$ is distributed as X for every $a > 0$. This implies the following identity in distribution for the stable tree:

$$R_\gamma(\mathscr{T}, a) \quad \text{under } \mathbf{N}^\gamma \quad \overset{(d)}{=} \quad \mathscr{T} \quad \text{under } a^{1/(\gamma-1)}\mathbf{N}^\gamma, \tag{0.2.21}$$

where the rescaling map $R_\gamma \colon \mathbb{T} \times (0,\infty) \to \mathbb{T}$ is defined by:

$$R_\gamma\big((T, \emptyset, d, \mu), a\big) = \big(T, \emptyset, ad, a^{\gamma/(\gamma-1)}\mu\big). \tag{0.2.22}$$

In words, $R_\gamma(T,a)$ is the compact real tree obtained from T by multiplying the distance by a and the measure by $a^{\gamma/(\gamma-1)}$. Using the scaling property 0.2.21, one can make sense of the stable tree conditioned by its total mass. More precisely, there exists a regular conditional probability $\mathbf{N}^\gamma_{(a)} = \mathbf{N}^\gamma[\cdot|\sigma = a]$ such that $\mathbf{N}^\gamma_{(a)}$-a.s. $\sigma = a$ and the following disintegration formula holds:

$$\mathbf{N}^\gamma[\mathrm{d}\mathscr{T}] = \frac{1}{\gamma\Gamma(1-1/\gamma)} \int_0^\infty \frac{\mathrm{d}a}{a^{1+1/\gamma}} \mathbf{N}^\gamma_{(a)}[\mathrm{d}\mathscr{T}]. \tag{0.2.23}$$

In particular, the distribution of the lifetime σ under $\mathbf{N}^\gamma$ is given by:

$$\mathbf{N}^\gamma[\sigma \in \mathrm{d}a] = \frac{1}{\gamma\Gamma(1-1/\gamma)} \frac{\mathrm{d}a}{a^{1+1/\gamma}}. \tag{0.2.24}$$

Informally, the probability measure $\mathbf{N}^\gamma_{(a)}$ can be seen as the distribution of the stable tree $\mathscr{T}$ with total mass a. Moreover, we have the following scaling property for the stable tree conditioned by its total mass:

$$\mathscr{T} \quad \text{under } \mathbf{N}^\gamma_{(a)} \quad \overset{(d)}{=} \quad R_\gamma(\mathscr{T}, a^{1-1/\gamma}) \quad \text{under } \mathbf{N}^\gamma_{(1)}. \tag{0.2.25}$$

Definition 0.2.16. *We call the stable tree with branching mechanism $\psi(\lambda) = \lambda^\gamma$ under $\mathbf{N}^\gamma_{(1)}$ the* normalized γ-stable tree.

We end this section by stating the convergence of size-conditioned BGW trees towards normalized stable trees. Let ξ be a $\mathbb{N}$-valued random variable which is critical (i.e. $\mathbb{E}[\xi] = 1$) and satisfies the following condition

(H_γ) ξ belongs to the domain of attraction of a stable law with index $\gamma \in (1,2]$, i.e. $\mathbb{E}\left[\xi^2 \mathbf{1}_{\{\xi \le n\}}\right] = n^{2-\gamma} L(n)$, where $L\colon \mathbb{R}_+ \to \mathbb{R}_+$ is a slowly varying function.

By [72, Theorem XVII.5.2], assumption (H_γ) is equivalent to the existence of a positive sequence $(b_n, n \ge 1)$ such that $R_n / b_n \xrightarrow{(d)} X_1$ where R is the random walk defined in Proposition 0.2.3 and X is the stable Lévy process with Laplace exponent $\psi(\lambda) = \lambda^\gamma$. The following result is due to Aldous [16] in the Brownian case $\gamma = 2$ and to Duquesne [54] in the non-Brownian stable case $\gamma \in (1,2)$, see also Kortchemski [110].

Theorem 0.2.17. *Assume that the offspring distribution ξ is critical and satisfies (H_γ). Let τ^n be a BGW(ξ) tree conditioned to have n vertices. The following convergence in distribution holds in the space $\mathbb{T}$:*

$$\frac{b_n}{n} \tau^n \xrightarrow[n\to\infty]{(d)} \mathscr{T}, \tag{0.2.26}$$

with respect to the GHP topology, where $\mathscr{T}$ is the normalized γ-stable tree.

Remark 0.2.18. Another natural notion of size for discrete trees is the number of leaves. Kortchemski [109] showed that a BGW tree with critical offspring distribution belonging to the domain of attraction of a stable law conditioned to have n leaves converges in distribution to the normalized stable tree after rescaling.

Remark 0.2.19. More generally, a stable branching mechanism is of the form $\psi(\lambda) = \kappa \lambda^\gamma$ with $\kappa > 0$, and the associated tree enjoys the same self-similarity properties as above. We choose the normalization $\kappa = 1$: notice that one can always modify the sequence b_n so that the convergence (0.2.26) holds with $\mathscr{T}$ being the normalized γ-stable tree with branching mechanism $\psi(\lambda) = \lambda^\gamma$.

0.3 Main contributions

0.3.1 Additive functionals on BGW trees

In this section, we present the main results of [7] which corresponds to Chapter 1.

Knowing that conditioned BGW trees converge in distribution to stable trees after rescaling (see Theorem 0.2.17), one can wonder if certain functionals of those trees converge. Obviously, the answer is positive for any functional which is continuous with respect to the GHP topology, for example the height of the tree. However, the GHP topology is not strong enough to deal with all interesting functionals. A typical example is the height profile of the tree, that is the number of vertices at any given level. As Theorem 0.2.17 suggests, this turns out to converge towards the local time process of the Brownian tree in the finite variance case; see [53, 136]. Here, we will be be interested in a class of functionals called additive.

Definition 0.3.1. *A functional F defined on the set of discrete trees is said to be additive if it satisfies the recursion:*

$$F(\mathbf{t}) = \sum_{i=1}^{k_\varnothing(\mathbf{t})} F(\mathbf{t}_i) + f(\mathbf{t}), \tag{0.3.1}$$

where $\mathbf{t}_1, \dots, \mathbf{t}_{k_\varnothing(\mathbf{t})}$ are the subtrees rooted at the children of the root of the tree $\mathbf{t}$, and f is a given toll function.

Notice that an additive functional F can also be written as:

$$F(\mathbf{t}) = \sum_{w \in \mathbf{t}} f(\mathbf{t}_w), \tag{0.3.2}$$

where $\mathbf{t}_w$ is the subtree of $\mathbf{t}$ above vertex w. Such functionals are naturally encountered in computer science where they represent the cost of divide-and-conquer type algorithms, in phylogenetics where they are used as a rough measure of the imbalance of trees or in chemical graph theory where they are employed as a predictive tool for some chemical properties. Among these, we mention the total path length defined by:

$$P(\mathbf{t}) = \sum_{v \in \mathbf{t}} H(v) = \sum_{w \in \mathbf{t}} |\mathbf{t}_w| - |\mathbf{t}|, \tag{0.3.3}$$

the Wiener index [147] defined by:

$$W(\mathbf{t}) = \sum_{u,v \in \mathbf{t}} d(u,v) = 2|\mathbf{t}| \sum_{w \in \mathbf{t}} |\mathbf{t}_w| - 2 \sum_{w \in \mathbf{t}} |\mathbf{t}_w|^2, \tag{0.3.4}$$

the shape functional, the Sackin index, the Colless index and the cophenetic index, see [144] for their definitions and also [52] for their representation using additive functionals, and references therein. We also mention the Shao and Sokal B_1 index [15, 144] defined by:

$$B_1(\mathbf{t}) = \sum_{\substack{w \in \mathbf{t}^\circ \\ w \neq \varnothing}} \frac{1}{\mathfrak{h}(\mathbf{t}_w)}. \tag{0.3.5}$$

It is used for assessing the balance of phylogenetic trees, see e.g. [71, 91, 107, 138, 143].

We shall consider size-conditioned BGW trees whose offspring distribution is critical and lies in the domain of attraction of a stable law. Roughly speaking, one can distinguish between two regimes: the local and the global regime. In the local regime, the toll function is small when the subtree is large so that the main contribution to the additive functional comes from small subtrees. These are almost independent which intuitively explains the asymptotic normality; see [95, 139, 153]. In the global regime, the toll function is large when the subtree is large and the main contribution comes from large subtrees which are strongly dependent. Thus the global shape of the BGW tree comes into play, which is why we expect the limit to be non-Gaussian but to depend on the normalized stable tree. We shall focus on the global regime for toll functions depending both on the size and height. Let us briefly review known results in the litterature. Fill and Kapur [74] studied the case of additive functionals associated with the toll function $f(\mathbf{t}) = |\mathbf{t}|^a$ for uniform binary trees (which is a special case of conditioned BGW trees, see Table 1); they observed a phase transition at $a = 1/2$ between the local and the global regime. In the global regime, corresponding to $a > 1/2$, they proved convergence in distribution after rescaling to a random variable characterized by its moments. This was very recently generalized by Fill and Janson [73] to BGW trees with critical offspring distribution with finite variance (this implies that $\gamma = 2$ in our notations below). When the offspring distribution has infinite variance but lies in the domain of attraction of a stable distribution with index $\gamma \in (1,2]$, Delmas, Dhersin and Sciauveau [52] proved convergence in distribution for $a \geq 1$ and conjectured a phase transition at $a = 1/\gamma$. We shall prove this conjecture as a particular case of our main result.

Let us compare the results of Fill and Janson [73] to ours which were obtained at the same time and independently. In that paper, the authors only consider offspring distributions with finite variance and toll functions of the form $f(\mathbf{t}) = |\mathbf{t}|^a$. They view the additive functional as a function of a; thus, on conditioned BGW trees it becomes a random analytic function. Using powerful techniques from complex analysis, they are able to show convergence for $\Re a > 1/2$ but also to extend their results to the so-called the intermediate regime which corresponds to $0 < \Re a < 1/2$. They also obtain interesting results on the line $\Re a = 1/2$ under an additional moment assumption on the offspring distribution. We use a very different approach which is intrinsic to trees, treating additive functionals on BGW trees as random measures. The problem is then reduced to the study of random measures, with the goal of showing their convergence in some appropriate space of measures which integrate functions with possible blow-up. This approach allows us to drop the finite variance assumption and also to consider more general toll functions, in particular of the form $f(\mathbf{t}) = |\mathbf{t}|^a \mathfrak{h}(\mathbf{t})^b$. One advantage of our approach is that it applies (at least partially) to other families of random trees that have scaling limits, for example random Pólya trees.

By letting f vary in (0.3.2), we obtain a measure associated with every tree. More precisely, for every discrete tree $\mathbf{t}$ and every $a > 0$, we define a (finite) measure $\mathscr{A}^\circ_{\mathbf{t},a}$ on the space $\mathbb{T} \times \mathbb{R}_+$ by:

$$\langle \mathscr{A}^\circ_{\mathbf{t},a}, f \rangle = \frac{a}{|\mathbf{t}|} \sum_{w \in \mathbf{t}^\circ} |\mathbf{t}_w| f(a\mathbf{t}_w, aH(w)), \tag{0.3.6}$$

where $a\mathbf{t}_w$ is the compact real tree obtained from $\mathbf{t}_w$ by multiplying all distances by a, equipped with $|\mathbf{t}|^{-1}$ times the counting measure. The measure $\mathscr{A}^\circ_{\mathbf{t},a}$ was already considered in [52] in a less general form. In Chapter 1, we show that $\mathscr{A}^\circ_{\mathbf{t},a}$ is a discretization of another (finite) measure $\Psi_{a\mathbf{t}}$, where for every compact real tree $(T, \emptyset, d, \mu)$, we set:

$$\langle \Psi_T, f \rangle = \int_T \mu(\mathrm{d}x) \int_0^{H(x)} f(T_{r,x}, r)\, \mathrm{d}r. \tag{0.3.7}$$

Then, we prove that the mapping $T \mapsto \Psi_T$ is continuous for the GHP topology on $\mathbb{T}$ and the topology of weak convergence on the space of measures on $\mathbb{R}_+ \times \mathbb{T}$. In conjunction with the scaling limit of conditioned BGW trees (see Theorem 0.2.17), this immediately yields the following result.

Proposition 0.3.2. *Let τ^n be a BGW(ξ) tree conditioned to have n vertices, with a critical offspring distribution ξ satisfying (H_γ) for some $\gamma \in (1,2]$. We have the following convergence in distribution and of all positive moments:*

$$\frac{b_n}{n^2} \sum_{w \in \tau^{n,\circ}} |\tau^n_w| f\left(\frac{b_n}{n}\tau^n_w, \frac{b_n}{n} H(w)\right) \xrightarrow[n\to\infty]{(d)} \langle \Psi_{\mathscr{T}}, f\rangle, \tag{0.3.8}$$

for every continuous and bounded function $f\colon \mathbb{T} \times \mathbb{R}_+ \to \mathbb{R}$, where $\mathscr{T}$ is the normalized γ-stable tree.

Clearly, the above reasoning is not specific to BGW trees but is very general: for any sequence of random discrete tree $(\tau^n, n \in \mathbb{N})$ such that $a_n \tau^n$ converges in distribution to some random compact real tree $\mathscr{T}$ for the GHP topology where $(a_n, n \in \mathbb{N})$ is a sequence of positive numbers converging to 0 and such that $(a_n \mathbb{E}[\mathfrak{h}(\tau^n)], n \in \mathbb{N})$ is bounded, one has the convergence in distribution of the random measure $\mathscr{A}^\circ_{\tau^n, a_n}$ to $\Psi_{\mathscr{T}}$. As an example, this applies to random Pólya trees which are known to converge to the Brownian tree; see [82, 131].

As a consequence of Proposition 0.3.2, the random variable

$$\sum_{w \in \tau^{n,\circ}} |\tau^n_w|^a \mathfrak{h}(\tau^n_w)^b \tag{0.3.9}$$

converges after rescaling for $a \geq 1$ and $b \geq 0$. Observe that when $b = 0$, this is not sufficient to get the phase transition conjectured in [52] at $a = 1/\gamma < 1$. In the specific case of conditioned BGW trees, we can improve this result as follows. We make the stronger assumption on the offspring distribution ξ:

(H'_γ) $\mathbb{E}\left[\xi^2 \mathbf{1}_{\{\xi \leq n\}}\right] = n^{2-\gamma} L(n)$ where $L\colon \mathbb{R}_+ \to \mathbb{R}_+$ is a slowly varying function which is bounded away from zero and infinity.

This is equivalent to the existence of a positive sequence $(b_n, n \geq 1)$ such that $b_n / n^{1/\gamma}$ is bounded away from zero and infinity and $R_n / b_n \xrightarrow{(d)} X_1$ where R is the random walk defined in Proposition 0.2.3 and X is the stable Lévy process with Laplace exponent $\psi(\lambda) = \lambda^\gamma$.

Theorem 0.3.3. *Let τ^n be a BGW(ξ) tree conditioned to have n vertices, with a critical offspring distribution ξ satisfying (H'_γ) for some $\gamma \in (1,2]$.*

(i) If $\gamma a + (\gamma - 1) b > 1$, we have the convergence in distribution and of the first moment:

$$\frac{b_n^{1+b}}{n^{1+a+b}} \sum_{w \in \tau^{n,\circ}} |\tau^n_w|^a \mathfrak{h}(\tau^n_w)^b \xrightarrow[n\to\infty]{(d)} \mathbf{Z}^\gamma_{a-1,b}. \tag{0.3.10}$$

(ii) If $\gamma a + (\gamma - 1)b \leq 1$, we have the convergence in distribution and of the first moment:

$$\frac{b_n^{1+b}}{n^{1+a+b}} \sum_{w \in \tau^{n,\circ}} |\tau_w^n|^a \mathfrak{h}(\tau_w^n)^b \xrightarrow[n\to\infty]{(d)} \infty. \tag{0.3.11}$$

The limiting random variable $\mathbf{Z}_{a,b}^\gamma$ has the following representation in terms of the normalized γ-stable random tree $\mathcal{T}$:

$$\mathbf{Z}_{a,b}^\gamma = \int_{\mathcal{T}} \mu(\mathrm{d}x) \int_0^{H(x)} \mu(\mathcal{T}_{r,x})^a \mathfrak{h}(\mathcal{T}_{r,x})^b \,\mathrm{d}r, \tag{0.3.12}$$

where $\mathcal{T}_{r,x}$ is the subtree of $\mathcal{T}$ above level r containing x.

Thus, we obtain a phase change for functionals of the size and height at $\gamma a + (\gamma - 1)b = 1$. Heuristically, the condition on a and b is due to the fact that the height of a (unnormalized) stable tree scales as its total mass to the power $(\gamma - 1)/\gamma$. In particular, we recover the result of Fill and Janson [73] by taking $b = 0$ in the finite variance case (so that $\gamma = 2$ and $\mathbf{Z}_{a,b}^2$ is a functional of the Brownian tree); in that case one can take $b_n = \sqrt{\sigma_\xi n}/2$ where σ_ξ^2 is the variance of ξ. Taking $b = 0$ also confirms the conjecture of Delmas, Dhersin and Sciauveau [52] as we observe a phase transition at $a = 1/\gamma$. If we fix $a = 0$ and let b vary, the phase transition occurs at $b = 1/(\gamma - 1) \geq 1$. In particular, Shao and Sokal's B_1 index is not covered by our result. Let us mention that in (0.3.10) and (0.3.11), we only sum over the set of internal vertices $\tau^{n,\circ}$ instead of the whole tree τ^n to avoid problems with dividing by 0 (indeed, if w is a leaf then $\mathfrak{h}(\tau_w^n) = 0$) but one can add the leaves as long as $b \geq 0$. More generally, instead of the toll function $f(\mathbf{t}) = |\mathbf{t}|^a \mathfrak{h}(\mathbf{t})^b$, our result applies to toll functions of the form $f(\mathbf{t}) = |\mathbf{t}|^a g(\mathfrak{h}(\mathbf{t}))$ or $f(\mathbf{t}) = h(|\mathbf{t}|)\mathfrak{h}(\mathbf{t})^\beta$, with the condition on a, b being replaced by an integral test; see Chapter 1 for further details.

Remark 0.3.4. Theorem 0.3.3-(ii) means that when $\gamma a + (\gamma - 1)b \leq 1$, the normalization b_n^{1+b}/n^{1+a+b} is not the correct one to get a nontrivial convergence result. Let us restrict ourselves to the finite variance case so that $\gamma = 2$. Then the main result of Janson [95] implies that under the condition $2a + b < 0$, the random variable in (0.3.9) converges in distribution after recentering and rescaling by $\sqrt{n}$ to a normal distribution, see Remark 1.1.2. This leaves a gap for $0 \leq 2a + b \leq 1$. When $b = 0$ and $0 < a < 1/2$, the situation is pretty well understood. Fill and Janson [73] proved that there is convergence in distribution after recentering at the same speed as our global regime, namely $n^{a+1/2}$. However, they do not have an explicit expression of the limit in terms of the Brownian tree, see Remark 1.27 therein in this direction. The nontrivial asymptotic behavior of (0.3.9) in the regime $\gamma a + (\gamma - 1)b < 1$ in the non-Brownian stable case $\gamma \in (1, 2)$ remains an open problem.

Problem 0.3.5. In the finite variance case, find an expression of the limit of (0.3.9) in terms of the Brownian tree for $0 < a < 1 + 2$ and $b = 0$. More generally, determine the asymptotic behavior of (0.3.9) for $b \neq 0$ and express the limit in terms of the Brownian tree.

Problem 0.3.6. In the non-Brownian stable case $\gamma \in (1, 2)$, find the asymptotic behavior of (0.3.9) in the regime $\gamma a + (\gamma - 1)b < 1$.

Next, we move on to the study of the limiting random variable $\mathbf{Z}_{a,b}^{\gamma}$. We give a complete description of its finiteness with the same phase transition at $\gamma a + (\gamma-1)b = 1-\gamma$ (notice that the random variable considered in Theorem 0.3.3 is $\mathbf{Z}_{a-1,b}^{\gamma}$) and we also compute its first moment.

Proposition 0.3.7. *Fix $\gamma \in (1,2]$ and let $a, b \in \mathbb{R}$. We have:*

$$\gamma a + (\gamma-1)b > 1-\gamma \iff \mathbf{Z}_{a,b}^{\gamma} < \infty \text{ a.s.} \iff \mathbb{E}\left[\mathbf{Z}_{a,b}^{\gamma}\right] < \infty, \tag{0.3.13}$$

$$\gamma a + (\gamma-1)b \le 1-\gamma \iff \mathbf{Z}_{a,b}^{\gamma} = \infty \text{ a.s.} \iff \mathbb{E}\left[\mathbf{Z}_{a,b}^{\gamma}\right] = \infty. \tag{0.3.14}$$

Furthermore, for every $a, b \in \mathbb{R}$ such that $\gamma a + (\gamma-1)b > 1-\gamma$, we have:

$$\mathbb{E}\left[\mathbf{Z}_{a,b}^{\gamma}\right] = \frac{1}{|\Gamma(-1/\gamma)|} \mathrm{B}\big(a + (b+1)(1-1/\gamma), 1-1/\gamma\big) \mathbb{E}\left[\mathfrak{h}(\mathscr{T})^{b}\right], \tag{0.3.15}$$

where $\mathscr{T}$ is the normalized γ-stable tree, Γ is the gamma function and B is the beta function.

Thanks to Duquesne and Wang [60], the height of the normalized γ-stable tree has finite moments of any order. In the Brownian case $\gamma = 2$, the moments of the height of the normalized Brownian tree are known explicitly:

$$\mathbb{E}\left[\mathbf{Z}_{a,b}^{2}\right] = \pi^{(b-1)/2} \Xi(b) \mathrm{B}\left(a + (b+1)/2, 1/2\right), \tag{0.3.16}$$

where Ξ is the Riemann xi function defined by $\Xi(s) = \frac{1}{2} s(s-1)\pi^{-s/2}\Gamma(s/2)\zeta(s)$ for every $s \in \mathbb{C}$ and ζ is the Riemann zeta function.

More generally, we can compute the intensity measure of the random measure $\Psi_{\mathscr{T}}$ where $\mathscr{T}$ is the normalized γ-stable tree. The next result is taken from [128].

Proposition 0.3.8. *Let $\mathscr{T}$ be the normalized γ-stable tree with $\gamma \in (1,2]$. Let f and g be non-negative measurable functions defined on $\mathbb{T}$ and $\mathbb{R}_+$ respectively. We have:*

$$\mathbb{E}\left[\int_{\mathscr{T}} \mu(\mathrm{d}x) \int_0^{H(x)} f(\mathscr{T}_{r,x}) g(r)\,\mathrm{d}r\right] = \mathbf{N}^{\gamma}\left[\sigma(1-\sigma)^{-1/\gamma} \mathbf{1}_{\{\sigma<1\}} f(\mathscr{T}) G(1-\sigma)\right], \tag{0.3.17}$$

where we set:

$$G(a) = \mathbb{E}\left[\int_{\mathscr{T}} \mu(\mathrm{d}x) g\left(a^{1-1/\gamma} H(x)\right)\right], \quad \forall a > 0. \tag{0.3.18}$$

In particular, we have:

$$\mathbb{E}\left[\int_{\mathscr{T}} \mu(\mathrm{d}x) \int_0^{H(x)} f(\mathscr{T}_{r,x})\,\mathrm{d}r\right] = \mathbf{N}^{\gamma}\left[\sigma(1-\sigma)^{-1/\gamma} \mathbf{1}_{\{\sigma<1\}} f(\mathscr{T})\right]. \tag{0.3.19}$$

This is a key result in the proof of Proposition 0.3.7 and can be interpreted as giving the distribution of the subtree $\mathscr{T}_{r,x}$ of the normalized stable tree, when x is a leaf chosen uniformly at random and r is chosen “uniformly” in $[0, H(x)]$.

Recall that Proposition 0.3.2 applies to other families of random trees having a scaling limit such as random Pólya trees. In particular, this implies the convergence of

$$\sum_{w \in \tau^{n,\circ}} |\tau^n_w|^a \mathfrak{h}(\tau^n_w)^b,$$

for $a \geq 1$ and $b \geq 0$, where τ^n is a uniform Pólya tree with n vertices. In [131], Panagiotou and Stufler showed that the shape of a uniform Pólya tree is essentially given by a large BGW tree (with finite variance) that it contains. This begs the question of whether additive functionals on uniform Pólya trees exhibit the same phase transition as conditioned BGW trees with finite variance.

Problem 0.3.9. Determine the phase transition for functionals of the size and height on random Pólya trees.

0.3.2 Zooming in at the root of the stable tree

In this section, we present the main results of [128] which corresponds to Chapter 2.

We study the shape of the normalized stable tree $\mathscr{T}$ near its root. More precisely, we show that, when zooming in at the root and rescaling, one gets the immortal tree, that is an infinite branch on which subtrees are grafted according to a Poisson point process. In particular, the rescaled subtrees near the root of $\mathscr{T}$ are independent and distributed as the unnormalized stable tree. In other words, we forget the conditioning for the total mass to be equal to 1 when zooming in. This idea to zoom in at the root of the stable tree is closely related to the small time asymptotics of the self-similar fragmentation process obtained from the stable tree by removing vertices located under height t, see Miermont [124] and Haas [80]. On the other end, Goldschmidt and Haas [76] studied the behavior of the aforementioned fragmentation process near its extinction time, which amounts to zooming in at the heighest leaf of $\mathscr{T}$.

Before stating our results, we first introduce some notations. Let $\mathscr{T}$ be the normalized γ-stable tree and, conditionally on $\mathscr{T}$, let U be a uniformly chosen leaf, that is U is a $\mathscr{T}$-valued random variable with distribution μ. Denote by $\mathscr{T}_i$, $i \in I_U$ the trees grafted on the branch $[\![\varnothing, U]\!]$ joining the root $\varnothing$ to the leaf U, each one at height h_i and with total mass $\sigma_i = \mu(\mathscr{T}_i)$, see Figure 3. Fix $\mathfrak{f}\colon (0,\infty) \to (0,\infty)$ (this represents the speed at which we zoom in) and define for every $\varepsilon > 0$ a point measure on $\mathbb{R}_+ \times \mathbb{T}$ by:

$$\mathscr{M}^{\mathfrak{f}}_\varepsilon(U) = \sum_{h_i \leq \mathfrak{f}(\varepsilon) H(U)} \delta_{(\varepsilon^{-1} h_i, R_\gamma(\mathscr{T}_i, \varepsilon^{-1}))}. \tag{0.3.20}$$

Observe that $\mathscr{M}^{\mathfrak{f}}_\varepsilon(U)$ describes the (rescaled) ancestral line of the leaf U starting from the root and up to a certain height given by $\mathfrak{f}(\varepsilon) H(U)$. In words, this amounts to zooming in at the root and rescaling everything at the same time. Finally, for any metric space E, we denote by $\mathscr{M}_p(E)$ the space of point measures on E equipped with the topology of vague convergence.

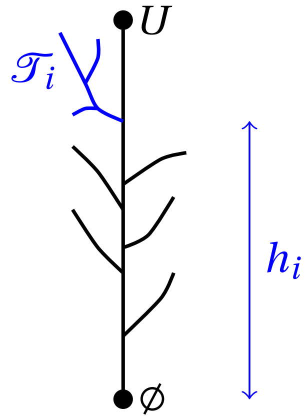

Figure 3 – The subtrees $\mathscr{T}_i$ grafted on the branch $[\![\varnothing, U]\!]$ at height h_i.

Our first main result states that the measure $\mathcal{M}_\varepsilon^{\mathfrak{f}}(U)$ converges to a Poisson point measure which is independent of the underlying tree $\mathscr{T}$ and of $H(U)$. Recall from (0.2.22) the definition of the rescaling map R_γ and from (0.2.15) the definition of the measure $\mathbf{N}_{\mathrm{B}}^{\psi}$ on $\mathbb{T}$. For $\gamma \in (1,2]$, denote by $\mathbf{N}_{\mathrm{B}}^{\gamma}$ the measure $\mathbf{N}_{\mathrm{B}}^{\psi}$ associated with a stable branching mechanism $\psi(\lambda) = \lambda^\gamma$.

Theorem 0.3.10. *Let $\mathscr{T}$ be the normalized γ-stable tree with $\gamma \in (1,2]$. Conditionally on $\mathscr{T}$, let U be a $\mathscr{T}$-valued random variable with distribution μ. Let $(\mathsf{T}_s, s \geq 0)$ be a $\mathbb{T}$-valued Poisson point process with intensity $\mathbf{N}_{\mathrm{B}}^{\gamma}$, independent of $(\mathscr{T}, H(U))$.*

(i) If $\lim_{\varepsilon\to 0} \varepsilon^{-1/2}\mathfrak{f}(\varepsilon) = 0$ and $\lim_{\varepsilon\to 0}\varepsilon^{-1}\mathfrak{f}(\varepsilon) = \infty$, then we have the following convergence in distribution in the space $\mathbb{T} \times \mathbb{R}_+ \times \mathcal{M}_p(\mathbb{R}_+ \times \mathbb{T})$:

$$\left(\mathscr{T}, H(U), \mathcal{M}_\varepsilon^{\mathfrak{f}}(U)\right) \xrightarrow[\varepsilon\to 0]{(d)} \left(\mathscr{T}, H(U), \sum_{s\geq 0} \delta_{(s,\mathsf{T}_s)}\right). \tag{0.3.21}$$

(ii) If $\mathfrak{f}(\varepsilon) = \varepsilon$, then we have the following convergence in distribution in the space $\mathbb{T} \times \mathbb{R}_+ \times \mathcal{M}_p(\mathbb{R}_+ \times \mathbb{T})$:

$$\left(\mathscr{T}, H(U), \mathcal{M}_\varepsilon^{\mathfrak{f}}(U)\right) \xrightarrow[\varepsilon\to 0]{(d)} \left(\mathscr{T}, H(U), \sum_{s\leq H(U)} \delta_{(s,\mathsf{T}_s)}\right). \tag{0.3.22}$$

In other words, zooming in at the speed $\mathfrak{f}(\varepsilon) = \varepsilon$ gives a *finite* branch on which subtrees are grafted in a Poissonian manner, whereas zooming in at a slower speed gives an *infinite* branch at the limit.

As an application of this result, we study the asymptotic behavior as $\max(a,b) \to \infty$ of additive functionals $\mathbf{Z}_{a,b}^{\gamma}$ on the normalized stable tree. Fill and Janson [73] considered the case $\gamma = 2$ and $b = 0$ (i.e. functionals of the mass on the Brownian tree) and proved that there is convergence in distribution as $a \to \infty$ of $\mathbf{Z}^\gamma_{a,0}$ properly normalized to

$$\int_0^\infty \mathrm{e}^{-S_t}\,\mathrm{d}t,$$

where $(S_t,\ t \geq 0)$ is a 1/2-stable subordinator. Their proof relies on the connection between the normalized Brownian excursion which codes the Brownian tree and the three-dimensional Bessel bridge. Our aim is twofold: we extend their result to the non-Brownian stable case $\gamma \in (1,2)$ while also considering polynomial functionals depending on both the mass and the height. We use a different approach – intrinsic to trees – which relies on the Bismut decomposition of the stable tree.

We distinguish two regimes according to the behavior of $b/a^{1-1/\gamma}$. The regime $b/a^{1-1/\gamma} \to c \in \mathbb{R}_+$ is related to Theorem 0.3.10 and can be stated as follows.

Theorem 0.3.11. *Assume that $a \to \infty$, $b \geq 0$ and $b/a^{1-1/\gamma} \to c \in \mathbb{R}_+$. Fix $\gamma \in (1,2]$ and let $\mathscr{T}$ be the normalized γ-stable tree. Then we have the following convergence in distribution:*

$$\lim_{a\to\infty} a^{1-1/\gamma}\mathfrak{h}(\mathscr{T})^{-b}\mathbf{Z}^\gamma_{a,b} = \int_0^\infty \mathrm{e}^{-S_t - ct/\mathfrak{h}(\mathscr{T})}\,\mathrm{d}t, \tag{0.3.23}$$

where $(S_t,\ t\geq 0)$ is a stable subordinator with Laplace exponent $\varphi(\lambda) = \gamma\lambda^{1-1/\gamma}$, independent of $\mathscr{T}$.

Let us briefly explain why we get a subordinator S at the limit. It is well known that μ is supported on the set of leaves of $\mathscr{T}$. Let $x \in \mathscr{T}$ be a leaf and recall that $\mathscr{T}_{r,x}$ is the subtree above level r containing x. Since the total mass of the stable tree is 1, the main contribution to $\mathbf{Z}^\gamma_{a,b}$ as $a \to \infty$ comes from large subtrees $\mathscr{T}_{r,x}$ with r close to 0. The height $\mathfrak{h}(\mathscr{T}_{r,x})$ of such subtrees is approximately $\mathfrak{h} - r$. On the other hand, their mass is equal to 1 minus the mass we discarded from the subtrees grafted on the branch $[\![\emptyset, x]\!]$ at height less than r. By Theorem 0.3.10, subtrees are grafted on $[\![\emptyset, x]\!]$ according to a point process which is approximately Poissonian, at least close to the root. Thus the mass $\mu(\mathscr{T}_{r,x})$ is approximately $1 - S_r$.

Notice that as long as the exponent b of the height does not grow too quickly, viz. $b/a^{1-1/\gamma} \to 0$, the additional dependence on the height makes no contribution at the limit. On the other hand, in the regime $b/a^{1-1/\gamma} \to \infty$, the height $\mathfrak{h}(\mathscr{T}_{r,x})^b$ dominates the mass $\mu(\mathscr{T}_{r,x})^a$ so we get the convergence in probability of $\mathbf{Z}^\gamma_{a,b}$ with a different scaling and there is no longer a subordinator at the limit.

Theorem 0.3.12. *Assume that $b \to \infty$, $a \geq 0$ and $a^{1-1/\gamma}/b \to 0$ and fix $\gamma \in (1,2]$. Then we have the following convergence in probability:*

$$\lim_{b\to\infty} b\mathfrak{h}(\mathscr{T})^{-b}\mathbf{Z}^\gamma_{a,b} = \mathfrak{h}(\mathscr{T}). \tag{0.3.24}$$

Recall from Proposition 0.3.7 that the random variable $\mathbf{Z}^{\gamma}_{a,b}$ is finite if and only if $\gamma a + (\gamma - 1)b > 1 - \gamma$, see Figure 4. Observe that we have only investigated the asymptotic behavior of $\mathbf{Z}^{\gamma}_{a,b}$ in the first quadrant $a, b \in \mathbb{R}_+$.

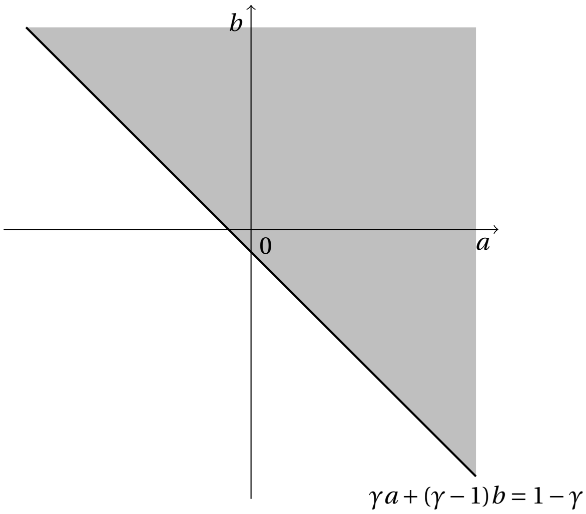

Figure 4 – Phase transition for $\mathbf{Z}^{\gamma}_{a,b}$: the random variable is finite in the shaded area and infinite otherwise.

Problem 0.3.13. Find the asymptotic behavior of $\mathbf{Z}^{\gamma}_{a,b}$ in the rest of the region $\gamma a + (\gamma - 1)b > 1 - \gamma$ as $\max(a, b) \to \infty$.

As mentioned above, the behavior of $\mathbf{Z}^{\gamma}_{a,b}$ depends on which of the mass or height dominates the other. One can then ask for a finer comparison between the two quantities near the root.

Problem 0.3.14. Determine the behavior of

$$\frac{\mu(\mathcal{T}_{r,U})}{\mathfrak{h}(\mathcal{T}_{r,U})^{\gamma/(\gamma-1)}}$$

as $r \to 0$, where $\mathcal{T}$ is the normalized γ-stable tree, $U \in \mathcal{T}$ is a uniform leaf and $\mathcal{T}_{r,U}$ is the subtree of $\mathcal{T}$ above level r containing U.

We believe the exponent $\gamma/(\gamma - 1)$ is the correct one due to the scaling properties of the stable tree under its excursion measure.

0.3.3 Maximal degree of Lévy trees

In this section, we present the main results of [6] which corresponds to Chapter 3.

We study the maximal degree of critical and subcritical Lévy trees. We establish a Poissonian decomposition of the tree along its large nodes and we determine the genealogical structure of those nodes. Furthermore, we make sense of the distribution of the Lévy tree conditioned to have a fixed maximal degree. Finally, we study the local limit of the Lévy tree conditioned on having large maximal degree. We show that a condensation phenomenon occurs in the subcritical case, whereas there is local convergence to the immortal tree in the critical case.

Questions regarding the maximal degree in the context of random graphs have been thoroughly investigated in the case of BGW trees. The first results in this direction were obtained by Jonsson and Stefánsson [98] who showed that a condensation phenomenon appears for a certain class of subcritical BGW trees conditioned to have a large size, in the sense that with high probability there exists a unique node with degree proportional to the size. Furthermore, the tree converges locally to a condensation tree consisting of a finite spine with random geometric length onto which independent and identically distributed BGW trees are grafted. This was later generalized by Janson [94], with further results in Kortchemski [111], Abraham and Delmas [4], Stufler [146]. On the other hand, He [85] shows that BGW trees conditioned on having large maximal degree converge locally to Kesten's tree (which consists of an infinite spine onto which independent and identically distributed BGW trees are grafted) in the critical case and to a condensation tree in the subcritical case.

In the continuum setting, Bertoin [34] determined the distribution of the maximal degree of a stable Lévy tree (his result is formulated in terms of Lévy processes). Using the formalism of CB processes, He and Li [88] treated the case of a general branching mechanism (in fact their result is more general as they considered CB processes with immigration). In [87], they also studied the local limit of a CB process conditioned to have large maximal degree (i.e. large maximal jump). In the critical case, they showed that it converges locally to a CB process with immigration. Later , He [86] extended the local convergence result to the whole genealogy: more precisely, he showed that a critical Lévy tree conditioned on having large maximal degree converges locally to an immortal tree (which is the continuous counterpart of Kesten's tree, consisting of an infinite spine onto which trees are grafted according to a Poisson point process). We improve these results by considering the density version of the conditioning instead of the tail version: more explicitly, we study the asymptotic behavior of critical Lévy trees conditioned to have maximal degree equal to (and not larger than) a given value. Density versions are finer than their tail counterparts and are usually more difficult to prove.

The existing litterature in the subcritical case is less developped. He and Li [87] showed that a subcritical CB process conditioned to have large maximal degree converges locally to a CB process with immigration which is killed (i.e. sent to infinity) at an independent exponential time, thus underlining a condensation phenomenon. We improve this result in several directions. Again we consider the density version of the conditioning instead of the tail version. We also extend the convergence result to the whole genealogical structure instead of the population size at a given time: this gives more information and, as an example, allows us to see that only one large node emerges. Finally, we are also able to describe precisely what happens above the condensation node.

Before stating these results, we need to introduce some notations. We consider a general branching mechanism ψ of the form (0.2.5) satisfying the Grey condition (0.2.6). We assume that $\pi \neq 0$ as otherwise the Lévy tree has no infinite branching points. Finally, for the sake of simplicity, we will assume that the Lévy measure π is diffuse. We refer the reader to Chapter 3 for results when this condition is not fulfilled.

Denote by $\bar{\pi}$ the tail of the Lévy measure π defined by:

$$\bar{\pi}(\delta) = \pi\left((\delta,\infty)\right), \quad \forall \delta \in \mathbb{R}_+. \tag{0.3.25}$$

Furthermore, for every $\delta > 0$, define the Laplace exponent ψ_δ by:

$$\psi_\delta(\lambda) = \left(\alpha + \int_{(\delta,\infty)} r\,\pi(\mathrm{d}r)\right)\lambda + \beta\lambda^2 + \int_{(0,\delta]} \left(\mathrm{e}^{-\lambda r} - 1 + \lambda r\right)\pi(\mathrm{d}r), \quad \forall \lambda \in \mathbb{R}_+, \tag{0.3.26}$$

which, in terms of the associated Lévy process, amounts to removing all the jumps with size (strictly) larger than δ. If the Lévy measure π is finite, we also define:

$$\psi_0(\lambda) = \left(\alpha + \int_{(0,\infty)} r\,\pi(\mathrm{d}r)\right)\lambda + \beta\lambda^2, \quad \forall \lambda \in \mathbb{R}_+. \tag{0.3.27}$$

Recall from Section 0.2.6 that every infinite branching point x_a (at height a) of the Lévy tree has a "degree" $\Delta_a > 0$. We denote by Δ the maximal degree of the Lévy tree defined by:

$$\Delta = \sup_{a \geq 0} \Delta_a. \tag{0.3.28}$$

The distribution of the maximal degree Δ under the excursion measure $\mathbf{N}^\psi$ has already been determined by He and Li [88] in the context of CB processes. However, the next result giving the joint distribution of the maximal degree Δ and the lifetime σ is new.

Proposition 0.3.15. *For every $\delta > 0$ and $\lambda \geq 0$, we have:*

$$\mathbf{N}^\psi\left[1 - \mathrm{e}^{-\lambda\sigma}\mathbf{1}_{\{\Delta \leq \delta\}}\right] = \psi_\delta^{-1}(\bar{\pi}(\delta) + \lambda). \tag{0.3.29}$$

Furthermore, if the Lévy measure π is finite, we have:

$$\mathbf{N}^\psi\left[1 - \mathrm{e}^{-\lambda\sigma}\mathbf{1}_{\{\Delta = 0\}}\right] = \psi_0^{-1}(\langle \pi, 1\rangle + \lambda). \tag{0.3.30}$$

Next, we give a decomposition of the Lévy tree along its large nodes, see Figure 5. More precisely, we show that the pruned Lévy tree (i.e. the subtree obtained by removing all nodes with large degree and the subtrees above them) is again a Lévy tree. Furthermore, one can recover the original tree from the pruned one (in distribution) by simply grafting Lévy forests at uniformly chosen leaves in a Poissonian manner. Before stating the result, we first need some notations. Recall from Definition 0.2.14 that $\mathbb{P}_r^\psi$ denotes the distribution of the Lévy forest with initial degree r. For every $\delta > 0$ such that $\bar{\pi}(\delta) > 0$, we let $\mathbb{Q}_\delta^\psi$ be the distribution of a Lévy forest with random initial degree with distribution π conditioned on being larger than δ. More formally, we set:

$$\mathbb{Q}_\delta^\psi(\mathrm{d}\mathscr{T}) = \frac{1}{\bar{\pi}(\delta)} \int_{(\delta,\infty)} \pi(\mathrm{d}r)\, \mathbb{P}_r^\psi(\mathrm{d}\mathscr{T}). \tag{0.3.31}$$

We also need to define the grafting procedure. Given a real tree $T \in \mathbb{T}$ and a finite or countable family $((x_i, T_i),\ i \in I)$ of elements of $T \times \mathbb{T}$, we denote by

$$T \circledast_{i\in I} (x_i, T_i)$$

the real tree obtained by grafting T_i on T at the node x_i. For a precise definition, we refer the reader to [3, Section 2.4].

Theorem 0.3.16. *Let $\delta \geq 0$ such that $\bar{\pi}(\delta) < \infty$. Under $\mathbf{N}^{\psi_\delta}$ and conditionally on $(\mathscr{T}, \emptyset, d, \mu)$, let $((x_i, \mathscr{T}_i),\ i \in I)$ be the atoms of a Poisson point measure with intensity $\bar{\pi}(\delta)\, \mu(\mathrm{d}x)\, \mathbb{Q}_\delta^\psi(\mathrm{d}\mathscr{T}')$. Then, under $\mathbf{N}^{\psi_\delta}$, the random tree $\mathscr{T} \circledast_{i\in I} (x_i, \mathscr{T}_i)$ has distribution $\mathbf{N}^\psi$.*

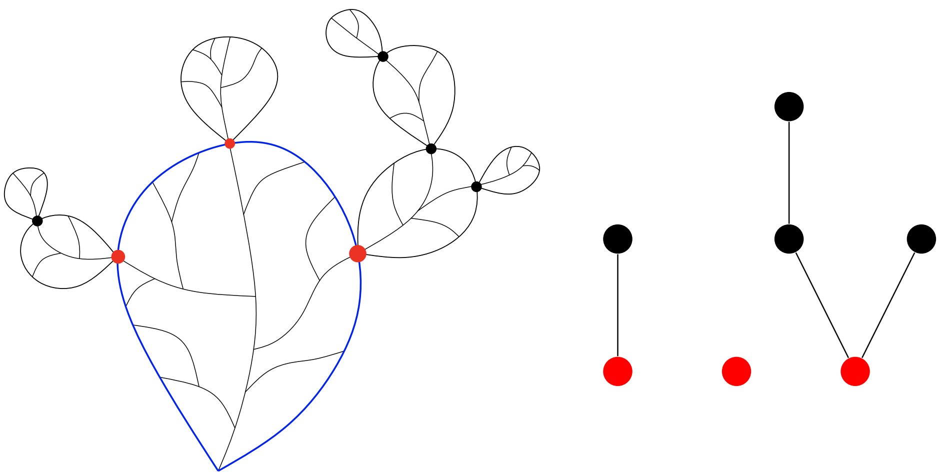

Figure 5 – Decomposition of the Lévy tree $\mathscr{T}$ along the nodes with degree larger than δ (left) and the associated discrete forest (right). In blue: the pruned subtree $\mathscr{T}^\delta$, in red: the first-generation nodes with degree larger than δ.

Theorem 0.3.16 is a special case of the main result of Abraham, Delmas and Voisin [8]. In that paper, the authors study a pruning procedure where they cut the tree according to some marks on the skeleton and on the infinite branching points. However, our proof is much simpler than theirs in this special case as the number of marks is finite.

As an application of Theorem 0.3.16, we determine the genealogical structure of the large nodes. More precisely, under $\mathbf{N}^\psi$, let $\mathbf{t}_\delta$ be the (random) discrete forest consisting of nodes with degree larger than δ and respecting the tree structure of $\mathscr{T}$; see Figure 5. Unsurprisingly, due to the branching property of the Lévy tree, the random forest $\mathbf{t}_\delta$ is a BGW forest. Given two $\mathbb{N}$-

valued random variables Z_0 and ξ, we call a (Z_0,ξ)-BGW forest a collection of Z_0 independent BGW(ξ) trees. We call Z_0 the initial distribution and ξ the offspring distribution.

Proposition 0.3.17. *Under $\mathbf{N}^\psi$, the random forest $\mathbf{t}_\delta$ is a BGW forest, with initial distribution Z_0^δ given by:*

$$\mathbf{N}^\psi\left[1-\mathrm{e}^{-\lambda Z_0^\delta}\right] = \psi_\delta^{-1}\Big((1-\mathrm{e}^{-\lambda})\bar{\pi}(\delta)\Big), \quad \forall \lambda > 0, \tag{0.3.32}$$

and offspring distribution ξ^δ given by:

$$\mathbf{N}^\psi\left[\mathrm{e}^{-\lambda \xi^\delta}\right] = \frac{1}{\bar{\pi}(\delta)}\int_{(\delta,\infty)} \mathrm{e}^{-r\psi_\delta^{-1}\left((1-\mathrm{e}^{-\lambda})\bar{\pi}(\delta)\right)}\,\pi(\mathrm{d}r), \quad \forall \lambda > 0. \tag{0.3.33}$$

One of our main results is the following theorem in which we make sense of the distribution of the Lévy tree conditioned to have a fixed maximal degree.

Theorem 0.3.18. *There exists a regular conditional probability $\mathbf{N}^\psi[\cdot|\Delta=\delta]$ for $\delta>0$ such that $\pi[\delta,\infty)>0$, which is given by, for every measurable and bounded $F\colon \mathbb{T}\to\mathbb{R}$:*

$$\mathbf{N}^\psi[F(\mathscr{T})|\Delta=\delta] = \frac{1}{\mathbf{N}^\psi[\sigma\mathbf{1}_{\{\Delta<\delta\}}]}\mathbf{N}^\psi\left[\int \mu(\mathrm{d}x)\,\mathbb{P}_\delta(\mathrm{d}\widetilde{\mathscr{T}}|\Delta\le\delta)F(\mathscr{T}\circledast(x,\widetilde{\mathscr{T}}))\mathbf{1}_{\{\Delta<\delta\}}\right]. \tag{0.3.34}$$

Furthermore, if $\langle \pi,1\rangle=\infty$, then $\mathbf{N}^\psi$-a.e. $\Delta>0$. If $\langle\pi,1\rangle<\infty$ then we have:

$$\mathbf{N}^\psi[F(\mathscr{T})\mathbf{1}_{\{\Delta=0\}}] = \mathbf{N}^{\psi_0}[F(\mathscr{T})\mathrm{e}^{-\langle\pi,1\rangle\sigma}]. \tag{0.3.35}$$

In particular, the Lévy tree conditioned to have maximal degree δ can be constructed as follows: take $\widetilde{\mathscr{T}}$ with size-biased (and degree-restricted) distribution $\mathbf{N}^\psi[\sigma\mathbf{1}_{\{\Delta\le\delta\}}]^{-1}\mathbf{N}^\psi[\cdot;\sigma\mathbf{1}_{\{\Delta\le\delta\}}]$, choose a leaf uniformly at random in $\widetilde{\mathscr{T}}$ and on this leaf graft an independent Lévy forest with initial degree δ conditioned to have maximal degree $\Delta\le\delta$. In fact, this conditioned random forest will have no nodes with degree δ other than the root, which entails that there is a unique node X_Δ with degree Δ. We show that the height of X_Δ is exponentially distributed and give a decomposition of the Lévy tree conditioned on Δ and on the height of its largest node X_Δ.

Finally, we turn to the behavior of the Lévy tree conditioned to have a large maximal degree. Similarly to the case of conditioned BGW trees, two drastically different types of limiting behavior appear. In the subcritical case, there is a condensation phenomenon where a node with infinite degree emerges at a finite exponentially distributed height. Denote by $\mathscr{T}_\Delta^-$ the pruned Lévy tree, that is the subtree obtained from $\mathscr{T}$ by removing the largest node X_Δ (and the subtree above it). We let $\mathscr{F}_\Delta^+$ be the forest above X_Δ, seen as a point measure on $\mathbb{R}_+\times\mathbb{T}$ (we refer to Chapter 3 for the precise definition); this artefact is introduced to deal with the fact that the limiting object, which we now introduce, is not a locally compact real tree. Denote by $\mathbb{P}_\infty^\psi(\mathrm{d}\mathscr{F})$ the distribution of a Poisson point measure on $\mathbb{R}_+\times\mathbb{T}$ with intensity $\mathrm{d}\ell\,\mathbf{N}^\psi[\mathrm{d}\mathscr{T}]$: this should be seen as the distribution of a Lévy forest with infinite initial degree.

Let $\mathbb{T}^*$ be the set of (isometry classes of) compact real trees that are *marked*, i.e. equipped with a distinguished vertex in addition to the root $\emptyset$. The set $\mathbb{T}^*$ can be made into a Polish space when equipped with a variant of the GHP topology which takes into account the additional mark.

Theorem 0.3.19. *Assume that ψ is subcritical and that the Lévy measure π is unbounded. Let $F\colon \mathbb{T}^* \to \mathbb{R}$ and $G\colon \mathscr{M}_p(\mathbb{R}_+ \times \mathbb{T}) \to \mathbb{R}$ be continuous and bounded and let A_δ be equal to any one of the following events: $\{\Delta = \delta\}$, $\{\Delta > \delta\}$, $\{\mathscr{T}$ has exactly one node with degree larger than $\delta\}$ or $\{\mathscr{T}$ has exactly one first-generation node with degree larger than $\delta\}$. We have:*

$$\lim_{\delta\to\infty} \mathbf{N}^\psi \left[F(\mathscr{T}_\Delta^-, X_\Delta)G(\mathscr{F}_\Delta^+)\middle| A_\delta\right] = \alpha\, \mathbf{N}^\psi\left[\int_{\mathscr{T}} F(\mathscr{T}, x)\,\mu(\mathrm{d}x)\right]\mathbb{P}_\infty^\psi(G(\mathscr{F})). \tag{0.3.36}$$

In particular, conditionally on A_δ, the height $H(X_\Delta)$ of X_Δ converges to an exponential distribution with mean $1/\alpha$.

In the critical case, He [86] showed that conditionally on $\Delta > \delta$, Lévy tree converges locally to the immortal tree. It should be no surprise then that the density version $\Delta = \delta$ gives rise to the same limiting behavior. We need to define the notion of local convergence for *locally compact* real trees. For every $h > 0$, define the restriction mapping on the set of (isometry classes of) real trees by:

$$r_h(T, d, \emptyset, \mu) = (T^h, d_{|T^h\times T^h}, \emptyset, \mu_{|T^h}) \quad \text{where } T^h = \{x \in T\colon\ H(x) \le h\}.$$

In other words, $r_h(T)$ is the real tree obtained from T by removing all nodes whose height is larger than h, equipped with the same metric and measure restricted to T^h. Recall that the Hopf-Rinow theorem implies that if T is a locally compact real tree, the closed ball $r_h(T)$ is compact. In particular, $r_h(T) \in \mathbb{T}$ and we can use the GHP topology. We say that a sequence T_n of locally compact trees converges locally to a locally compact tree T if for every $h > 0$, the sequence $r_h(T_n)$ converges for the GHP topology to $r_h(T)$ in the space $\mathbb{T}$.

Before stating our result, we need to define the immortal tree. Let $\sum_{i\in I}\delta_{(s_i,\mathscr{T}_i)}$ be a Poisson point measure on $\mathbb{R}_+\times\mathbb{T}$ with intensity $\mathrm{d}s\,\mathbf{N}_{\mathrm{B}}^\psi(\mathrm{d}\mathscr{T})$, where the measure $\mathbf{N}_{\mathrm{B}}^\psi$ is defined in (0.2.15). The immortal Lévy tree $\mathscr{T}_\infty^\psi$ with branching mechanism ψ is the real tree obtained by grafting the point measure $\sum_{i\in I}\delta_{(s_i,\mathscr{T}_i)}$ on an infinite branch. More formally, we set:

$$\mathscr{T}_\infty^\psi = \mathbb{R}_+ \circledast_{i\in I} (s_i, \mathscr{T}_i), \tag{0.3.37}$$

where $\mathbb{R}_+$ is considered as a real tree rooted at 0 and equipped with the Euclidean distance and the zero measure.

Theorem 0.3.20. *Assume that ψ is critical and that π is unbounded. Let A_δ be equal to any of the following events: $\{\Delta = \delta\}$, $\{\Delta > \delta\}$, $\{\mathscr{T}$ has exactly one node with degree larger than $\delta\}$ or $\{\mathscr{T}$ has exactly one first-generation node with degree larger than $\delta\}$. Then, conditionally on A_δ, the*

Lévy tree $\mathscr{T}$ converges in distribution locally to the immortal Lévy tree $\mathscr{T}_\infty^\psi$, i.e. we have:

$$\lim_{\delta\to\infty} \mathbf{N}^\psi \left[F(r_h(\mathscr{T}))|A_\delta\right] = \mathbb{E}\left[F(r_h(\mathscr{T}_\infty^\psi))\right]. \tag{0.3.38}$$

Remark 0.3.21. Let us briefly highlight the changes when the Lévy measure π has atoms; we refer to Chapter 3 for further details. The joint distribution of Δ and σ (Proposition 0.3.15), the decomposition of the Lévy tree along its large nodes (Theorem 0.3.16) and the branching structure of large nodes (Proposition 0.3.17) all remain unchanged. We can still make sense of the conditional distribution $\mathbf{N}^\psi[\cdot|\Delta = \delta]$ (Theorem 0.3.18) when $\pi(\{\delta\}) > 0$: instead of the conditioned Lévy tree having exactly one node with degree δ, the number of first-generation nodes with degree δ has a Poisson distribution conditioned on being at least 1. In the subcritical case, when letting $\delta \to \infty$, there is exactly one node with degree δ with high probability. As a result, we observe the same condensation as in Theorem 0.3.19, that is only one condensation node emerges. In the critical case, we show that there is local convergence to the immortal tree (as in Theorem 0.3.20) provided some condition on the size of the atoms is satisfied. However, we do not know whether this additional assumption is necessary.

We conclude with some open problems. In [60], Duquesne and Wang investigate a similar question on Lévy trees: conditioning by the diameter. They apply their results to study the asymptotic behavior of the cdf of the diameter of the normalized stable tree at zero and infinity.

Problem 0.3.22. Find the asymptotic behavior of the cdf of the maximal degree of the normalized stable tree at zero and infinity.

Finally, conditioning Lévy trees by their width could be another interesting line of research. Some results are already obtained in this area: He [86] showed that a critical Lévy tree conditioned to have large width converges locally to the immortal tree; Abraham, Delmas and He [9] studied the asymptotic behavior of the Brownian tree conditioned to have local time at level t equal to some determinstic function a_t and identified four regimes (one of them being conjectured) depending on the growth rate of a_t.

Problem 0.3.23. Determine the local limit of a (sub)critical Lévy tree conditioned by its width or its local time at a given level.

1 Global regime for general additive functionals of conditioned BGW trees

This chapter is based on the paper [7], published in *Probability Theory and Related Fields*.

We give an invariance principle for very general additive functionals of conditioned Bienaymé-Galton-Watson trees in the global regime when the offspring distribution lies in the domain of attraction of a stable distribution, the limit being an additive functional of a stable Lévy tree. This includes the case when the offspring distribution has finite variance (the Lévy tree being then the Brownian tree). We also describe, using an integral test, a phase transition for toll functions depending on the size and height.

1.1 Introduction

In view of the many applications of trees (in computer science, biology, physics, ...), the study of additive functionals on large random trees has seen a lot of development in recent years, see references below. In this paper, we consider asymptotics for general additive functionals on conditioned Bienaymé-Galton-Watson (BGW for short) trees in the so-called global regime.

Recall that a functional F defined on finite rooted ordered discrete trees is said to be additive if it satisfies the recursion

$$F(\mathbf{t}) = \sum_{i=1}^{d} F(\mathbf{t}_i) + f(\mathbf{t}), \tag{1.1.1}$$

where $\mathbf{t}_1, \ldots, \mathbf{t}_d$ are the subtrees rooted at the d children of the root of the tree $\mathbf{t}$ and f is a given toll function. Notice that this can also be written as

$$F(\mathbf{t}) = \sum_{w \in \mathbf{t}} f(\mathbf{t}_w), \tag{1.1.2}$$

where $\mathbf{t}_w$ is the subtree of $\mathbf{t}$ above the vertex w and rooted at w. Such functionals are encountered in computer science where they represent the cost of divide-and-conquer algorithms, in phylogenetics where they are used as a rough measure of tree shape to detect imbalance or in chemical graph theory where they appear as a predictive tool for some chemical properties.

Among these, we mention the total path length defined as the sum of the distances to the root of all vertices, the Wiener index [147] defined as the sum of the distances between all pairs of vertices, the shape functional, the Sackin index, the Colless index and the cophenetic index, see [144] for their definitions and also [52] for their representation using additive functionals, and the references therein. See also [139] for other functionals such that the number of matchings, dominating sets, independent sets for trees. We also mention the Shao and Sokal B_1 index [15, 144] defined by

$$B_1(\mathbf{t}) = \sum_{\substack{w\in\mathbf{t}^\circ\\ w\neq\emptyset}} \frac{1}{\mathfrak{h}(\mathbf{t}_w)}, \tag{1.1.3}$$

where for every finite rooted ordered tree $\mathbf{t}$, $\mathfrak{h}(\mathbf{t})$ is its height and $\mathbf{t}^\circ$ is the set of internal vertices. It is used for assessing the balance of phylogenetic trees, see e.g. [71, 91, 107, 138, 143].

We shall consider in this paper random discrete trees τ^n which are BGW trees conditioned to have n vertices, and then study the limit of rescaled additive functionals as n goes to infinity. One can distinguish between local and global regime. In the local regime, the toll function is small or even vanishes when the subtree is large; so the main contribution to the additive functional comes from the small subtrees. These being almost independent, we understand intuitively why the limit distribution is Gaussian. See [95, 139, 153] for asymptotic results in the local regime. In the global regime, the toll function is large when the subtree is large; so the main contribution comes from large subtrees which are strongly dependent. This intuitively explains why we expect the limit to be non-Gaussian. As far as we know, asymptotic results in the global regime deal with toll functions depending only on the size. In this paper, we shall focus on the global regime for general toll functions. In particular, our results apply to toll functions depending on the size and height. When the toll function is monomial in the size of the tree $f(\mathbf{t}) = |\mathbf{t}|^{\alpha'}$, with $|\mathbf{t}|$ the cardinal of $\mathbf{t}$, Fill and Kapur [74] observed a phase transition at $\alpha' = 1/2$ for binary trees under the Catalan model (which is a special case of conditioned BGW trees): the global regime corresponds to $\alpha' > 1/2$. This was later generalized by Fill and Janson [73] to BGW trees with critical offspring distribution with finite variance using techniques from complex analysis; they identified a local regime for $\alpha' < 0$ and an intermediate regime for $0 < \alpha' < 1/2$. When the offspring distribution has infinite variance but lies in the domain of attraction of a stable distribution with index $\gamma \in (1,2]$, Delmas, Dhersin and Sciauveau [52] proved convergence in distribution for $\alpha' \geq 1$ using stable Lévy trees and conjectured a phase transition at $\alpha' = 1/\gamma$. We shall prove this conjecture, as a particular case of our main result, see Theorem 1.1.1.

Let ξ be an $\mathbb{N}$-valued random variable. We write BGW(ξ) tree for a BGW tree with offspring distribution (the law of) ξ. We denote by τ^n a BGW(ξ) tree conditioned to have n vertices and we assume that ξ is critical, i.e. $\mathbb{E}[\xi] = 1$, and that it belongs to the domain of attraction of a stable distribution with index $\gamma \in (1,2]$, i.e. there exists a positive sequence $(b_n, n \geq 1)$ such that if $(\xi_n, n \geq 1)$ is a sequence of independent random variables with the same distribution as ξ then $b_n^{-1}\left(\sum_{k=1}^n \xi_k - n\right)$ converges in distribution towards a stable random variable whose Laplace transform is given by $\exp(\kappa\lambda^\gamma)$ for $\lambda \geq 0$, with index $\gamma \in (1,2]$ and normalizing constant

$\kappa > 0$ (the constant κ depends on the choice of the sequence $(b_n, n \geq 1)$). Notice that ξ is necessarily nondegenerate, i.e. ξ is not constant, since $\kappa > 0$. Under these assumptions, it is also well known that, as n goes to infinity, τ^n properly rescaled converges in distribution with respect to the Gromov-Hausdorff-Prokhorov topology to the stable Lévy tree $\mathscr{T}$ with index γ (and branching mechanism $\psi(\lambda) = \kappa\lambda^\gamma$, see Section 1.4.2 for a precise definition), see Aldous [16] for the finite variance case and Duquesne [54] for the general case. The stable Lévy tree is a generalization of Aldous' Brownian continuum random tree which corresponds to $\gamma = 2$. We refer the reader to Le Gall and Le Jan [119], Duquesne and Le Gall [57, 58] for the definition of Lévy trees, see also Section 1.4.2 for a brief summary in the stable case. We recall that any real tree T is endowed with the length measure $\ell(\mathrm{d}y)$ (which roughly speaking is the Lebesgue measure on the branches of the tree) and that the Lévy tree is naturally endowed with a mass measure (which roughly speaking is the uniform probability measure on the infinite set of leaves). One of our main results can be stated as follows. We refer the reader to Proposition 1.7.1 and Theorem 1.7.3 for more general statements. Recall that $\mathbf{t}^\circ$ denotes the set of internal vertices of the discrete tree $\mathbf{t}$.

Theorem 1.1.1. *Let τ^n be a BGW(ξ) tree conditioned to have n vertices, with ξ being critical and in the domain of attraction of a stable distribution with index $\gamma \in (1,2]$. We suppose moreover that the sequence $(b_n, n \geq 1)$ defined as above is such that $(b_n/n^{1/\gamma}, n \geq 1)$ is bounded away from zero and infinity. Let $\mathscr{T}$ be the stable Lévy tree with branching mechanism $\psi(\lambda) = \kappa\lambda^\gamma$. Let $\alpha', \beta \in \mathbb{R}$.*

(i) If $\gamma\alpha' + (\gamma-1)\beta > 1$, we have the convergence in distribution and of the first moment

$$\frac{b_n^{1+\beta}}{n^{1+\alpha'+\beta}} \sum_{w\in\tau^{n,\circ}} |\tau_w^n|^{\alpha'}\, \mathfrak{h}(\tau_w^n)^\beta \xrightarrow[n\to\infty]{(d)+\text{mean}} \int_{\mathscr{T}} \mathfrak{m}(\mathscr{T}_y)^{\alpha'}\, \mathfrak{h}(\mathscr{T}_y)^\beta\, \ell(\mathrm{d}y), \tag{1.1.4}$$

where the right-hand side of (1.1.4) has finite mean and, for $y \in \mathscr{T}$, $\mathscr{T}_y$ is the subtree of $\mathscr{T}$ above y, $\mathfrak{m}(\mathscr{T}_y)$ is its mass, and $\mathfrak{h}(\mathscr{T}_y)$ its height.

(ii) If $\gamma\alpha' + (\gamma-1)\beta \leq 1$, we have the convergence in distribution and of the first moment

$$\frac{b_n^{1+\beta}}{n^{1+\alpha'+\beta}} \sum_{w\in\tau^{n,\circ}} |\tau_w^n|^{\alpha'}\, \mathfrak{h}(\tau_w^n)^\beta \xrightarrow[n\to\infty]{(d)+\text{mean}} \infty. \tag{1.1.5}$$

We complete the previous result with some comments.

Remark 1.1.2. (i) From Theorem 1.1.1, we obtain a phase change for functionals of the mass and height at $\gamma\alpha' + (\gamma-1)\beta = 1$. Heuristically, the condition on α' and β is due to the fact that the height of a (unnormalized) stable Lévy tree scales as its mass to the power $(\gamma-1)/\gamma$. Let us mention that this phase change is specific to BGW trees, see Remark 1.4.14 in this direction.

(ii) See conditions (ξ1) and (ξ2) in Section 1.4 for a more detailed discussion of the assumptions on the offspring distribution. The additional boundedness assumption on

$(b_n/n^{1/\gamma}, n \geq 1)$ is also equivalent to $(\xi 2)'$. The latter can be dropped in (i) of Theorem 1.1.1 when $\alpha' \geq 1$ and $\beta \geq 0$ according to Proposition 1.4.11.

(iii) We also have the convergence (and finiteness) of the moments of all order $p > 1$ in (1.1.4) as soon as $p(\gamma\alpha + (\gamma-1)\beta) > 1-\gamma$, with $\alpha = \alpha' - 1$, see Proposition 1.7.1. In particular for $\beta = 0$, we have the convergence of all nonnegative moments for $\alpha' \geq 1$. However, in the finite variance case, for $\alpha' \in (1/2, 1)$ (and $\beta = 0$), our result is not optimal, see (vi) below.

(iv) Theorem 1.1.1 generalizes a result by Delmas, Dhersin and Sciauveau where only functionals of the mass are considered (i.e. $\beta = 0$), see [52, Lemma 4.6]. In particular, we prove the conjecture stated therein: when $\beta = 0$, there is a phase transition at $\alpha' = 1/\gamma$ (the parameter α therein corresponds to $\alpha' - 1$ here). If we fix $\alpha' = 0$ and let β vary, the phase transition occurs at $\beta = 1/(\gamma-1) \geq 1$. In particular, Shao and Sokal's B_1 index, which corresponds to $\alpha' = 0$ and $\beta = -1$, lies in the local regime, regardless the value of the index γ and is therefore not covered by our results. See also (vi) below.

(v) If the offspring distribution has finite variance $\sigma_\xi^2 \in (0,\infty)$, one can take $b_n = b\sqrt{n}$ in which case $\mathscr{T}$ is distributed as the Brownian continuum random tree with branching mechanism $\psi(\lambda) = \sigma_\xi^2 \lambda^2/(2b^2)$. For $b = \sigma_\xi$, the contour process of $\mathscr{T}$ is a standard Brownian motion under its normalized excursion measure.

(vi) Assume that the offspring distribution has finite variance $\sigma_\xi^2 \in (0,\infty)$, which implies that $\gamma = 2$. We consider the asymptotics in the local regime of $\sum_{w \in \tau^{n,\circ}} |\tau_w^n|^{\alpha'} \mathfrak{h}(\tau_w^n)^\beta$, that is when $\alpha', \beta \in \mathbb{R}$ such that $2\alpha' + \beta < 0$. Denote by $F_{\alpha',\beta}$ the additive functional (1.1.2) associated with the toll function $f_{\alpha',\beta}(\mathbf{t}) = |\mathbf{t}|^{\alpha'} \mathfrak{h}(\mathbf{t})^\beta \mathbf{1}_{\{|\mathbf{t}|>1\}}$. By [95, Theorem 1.5] and Lemma 1.4.5, we have

$$\frac{F_{\alpha',\beta}(\tau^n) - n\mu}{\sqrt{n}} \xrightarrow[n\to\infty]{(d)} \mathcal{N}(0, \varsigma^2),$$

where μ, ς^2 are finite and given by $\mu = \mathbb{E}\left[f_{\alpha',\beta}(\tau)\right]$ and by $\varsigma^2 = 2\mathbb{E}\left[f_{\alpha',\beta}(\tau)\left(F_{\alpha',\beta}(\tau) - |\tau|\mu\right)\right] - \mathrm{Var}(f_{\alpha',\beta}(\tau)) - \mu^2/\sigma_\xi^2$, and τ is the corresponding unconditioned BGW tree. In particular, this covers Shao and Sokal's B_1 index (where $\alpha' = 0$ and $\beta = -1$). Notice that this leaves a gap for $0 \leq 2\alpha' + \beta \leq 1$. At least when $\beta = 0$, the situation is well understood. Fill and Janson [73] identify three different regimes: the global regime for $\alpha' > 1/2$, the local regime for $\alpha' < 0$ and an intermediate regime for $0 < \alpha' < 1/2$. The nontrivial asymptotic behavior of $F_{\alpha',\beta}(\tau^n)$ for $\gamma \in (1,2)$ and $\gamma\alpha' + (\gamma-1)\beta \leq 1$ (that is the non global regime in the non quadratic case) is an open question.

(vii) When τ^n is uniformly distributed among the set of full binary ordered trees with n vertices (which corresponds to a conditioned BGW(ξ) tree with $\mathbb{P}(\xi = 0) = \mathbb{P}(\xi = 2) = 1/2$), Fill and Kapur [74] studied the local and global regime when the toll function is a power of the size of the tree. Concerning the global regime, they showed the convergence in distribution, using the convergence of all positive moments in (1.1.4) for $\alpha' > 1/2$ and $\beta = 0$, see Eq. (3.14) and Proposition 3.5 therein. In that case, one can take $b_n = \sqrt{n}$ and $\mathscr{T}$ is the Brownian tree with branching mechanism $\psi(\lambda) = \lambda^2/2$. See also Fill and Janson

[73] for general critical offspring distribution with finite variance. The explicit formula for the first moment of the right-hand side of (1.1.4) are given by the right-hand side of (1.1.12) with $\kappa = 1/2$ and $\alpha = \alpha' - 1$.

(viii) As an application, using (1.1.4), we obtain, when $\alpha' > 1/\gamma$, in Example 1.7.5 (with $\alpha' = \alpha + 1$) an asymptotic expansion in distribution for $b_n\, n^{-(1+\alpha')} \sum_{w\in\tau^{n,\circ}} \left|\tau^n_w\right|^{\alpha'} \log\left|\tau^n_w\right|$.

(ix) Panagiotou and Stufler [131] showed that the shape of a uniform Pólya tree is essentially given by a large BGW tree (with finite variance) that it contains. Using this decomposition, they proved the convergence of uniform Pólya trees after rescaling to the Brownian tree. One may then conjecture that additive functionals on uniform Pólya trees exhibit a phase change at $2\alpha' + \beta = 1$, similarly to the finite variance BGW trees. In this paper, we only prove that the convergence (1.1.4) holds for uniform Pólya trees when $\alpha' \geq 1$ and $\beta \geq 0$ as an easy consequence of our approach, see Remark 1.4.13. We do not investigate this further and we leave the phase transition as an open question. However, let us mention that, as shown in Remark 1.4.14, slightly modifying BGW trees may result in a different phase change.

More generally, if one views a discrete tree as a real tree, then the left-hand side in (1.1.4) is related to the discrete length measure $\ell_n(\mathrm{d}y) = \sum_{w\in\tau^n} \delta_w(\mathrm{d}y)$ of τ^n (after rescaling by b_n/n). One way to interpret the result would be to say that the sequence of measures $\int_{\tau^n} \delta_{\tau^n_y}\, \ell_n(\mathrm{d}y)$ converges in distribution to $\int_{\mathcal{T}} \delta_{\mathcal{T}_y}\, \ell(\mathrm{d}y)$ in some sense. One might then hope to prove that the mapping $T \mapsto \int_T \delta_{T_y}\, \ell(\mathrm{d}y)$ is continuous on the space of compact real trees. This is not true however, see Remark 1.4.14, one problem being that the length measure is not finite in general. To overcome this difficulty, our approach, inspired by [52], consists in considering the length measure biased by the size of the subtree above y, thus penalizing small subtrees.

More precisely let $\mathbb{T}$ be the space of (equivalent classes of) weighted rooted compact real trees (i.e. the set of quadruplets $(T, \varnothing, d, \mu)$ where (T, d) is a real tree, $\varnothing$ is a distinguished vertex of T called the root, and the mass measure μ is a finite measure on T). We recall that the length measure ℓ on a real tree (T, d) has an intrinsic definition. For every $(T, \varnothing, d, \mu) \in \mathbb{T}$, we define a measure Ψ_T on $\mathbb{T} \times \mathbb{R}_+$ as follows: for every nonnegative measurable function f defined on $\mathbb{T} \times \mathbb{R}_+$,

$$\Psi_T(f) = \int_T f\left(T_y, H(y)\right) \mu(T_y)\, \ell(\mathrm{d}y), \tag{1.1.6}$$

where $H(y) = d(\varnothing, y)$ denotes the height of y (i.e. the distance to the root) in T. We also consider the measure $\Psi^{\mathfrak{mh}}_T$ on $\mathbb{R}^2_+$ defined similarly to Ψ_T for functions depending only on the mass and height of the tree, see (1.3.2).

If $\mathbf{t}$ is a finite rooted ordered tree and $a > 0$, we denote by $a\mathbf{t}$ the real tree associated with $\mathbf{t}$, rescaled so that all edges have length a and equipped with the uniform probability measure on the set of vertices whose heights are integer multiples of a, see Section 1.2.4 for a precise definition. Furthermore, for $w \in \mathbf{t}$, we write aw for the corresponding vertex in $a\mathbf{t}$ and $a\mathbf{t}_w$ for the subtree of $a\mathbf{t}$ above aw. The height of w in $\mathbf{t}$ is denoted by $H(w)$; thus the height of aw in

$a\mathbf{t}$ is $aH(w)$. In the spirit of [52], we consider the measure $\mathscr{A}^\circ_{\mathbf{t},a}$ on $\mathbb{T}\times\mathbb{R}_+$ defined as follows: for every nonnegative measurable function f defined on $\mathbb{T}\times\mathbb{R}_+$,

$$\mathscr{A}^\circ_{\mathbf{t},a}(f) = \frac{a}{|\mathbf{t}|}\sum_{w\in\mathbf{t}^\circ}|\mathbf{t}_w|\, f\left(a\mathbf{t}_w, aH(w)\right). \tag{1.1.7}$$

In (1.1.7), instead of summing over all the internal vertices ($w\in\mathbf{t}^\circ$) one could also sum over all vertices including the leaves ($w\in\mathbf{t}$); in this case the measure is denoted by $\mathscr{A}_{\mathbf{t},a}$. The two measures are close in total variation as $d_{\mathrm{TV}}(\mathscr{A}_{\mathbf{t},a},\mathscr{A}^\circ_{\mathbf{t},a})\leq a$; see (1.4.18). We mention that the measure $\mathscr{A}_{\mathbf{t},a}$ was already considered in [52] for functions f depending only on the size.

For every finite rooted ordered tree $\mathbf{t}$ and $a>0$, we show (see Lemma 1.4.9) that the measures $\mathscr{A}^\circ_{\mathbf{t},a}$ and $\mathscr{A}_{\mathbf{t},a}$ can be approximated by $\Psi_{a\mathbf{t}}$. In Proposition 1.3.4, we give another expression for Ψ_T:

$$\Psi_T(f) = \int_T \mu(\mathrm{d}x)\int_0^{H(x)} f\left(T_{r,x}, r\right)\mathrm{d}r, \tag{1.1.8}$$

for every nonnegative measurable function f defined on $\mathbb{T}\times\mathbb{R}_+$. Here $T_{r,x}$ is the subtree of T above level r containing x. This latter expression of Ψ_T is used to prove it is continuous as a function of T, see Proposition 1.3.3.

Theorem 1.1.3. *The mapping $T\mapsto\Psi_T$, from $\mathbb{T}$ endowed with the Gromov-Hausdorff-Prokhorov topology to $\mathscr{M}(\mathbb{T}\times\mathbb{R}_+)$, the space of nonnegative finite measures on $\mathbb{T}\times\mathbb{R}_+$, endowed with the topology of weak convergence, is well defined and continuous.*

This allows us to derive a general invariance principle: for any sequence of random discrete trees $(\tau^n,\, n\in\mathbb{N})$ such that $a_n\tau^n$ converges in distribution to some random real tree $\mathscr{T}$ in the Gromov-Hausdorff-Prokhorov topology where $(a_n,\, n\in\mathbb{N})$ is a sequence of positive numbers converging to 0 and such that $(a_n\mathbb{E}[\mathfrak{h}(\tau^n)],\, n\in\mathbb{N})$ is bounded, one has the convergence in distribution of the measures $\mathscr{A}^\circ_{\tau^n,a_n}$ and $\mathscr{A}_{\tau^n,a_n}$ to $\Psi_{\mathscr{T}}$ (this is a consequence of Lemma 1.4.9 and Theorem 1.1.3). For example, this applies to Pólya trees, see Remark 1.4.13, which were shown to converge to the Brownian tree, see [82] and [131]. For BGW trees, we have the following result which is a direct consequence of the convergence of conditioned BGW trees to stable Lévy tree, see [54], and Theorem 1.1.3 and Lemma 1.4.9.

Corollary 1.1.4. *Let τ^n be a BGW(ξ) tree conditioned to have n vertices, with ξ satisfying (ξ1) and (ξ2), and $(b_n,\, n\geq 1)$ be defined as in Theorem 1.1.1. Let $\mathscr{T}$ be the stable Lévy tree with branching mechanism $\psi(\lambda)=\kappa\lambda^\gamma$. We have the following convergence in distribution and of all positive moments*

$$\frac{b_n}{n^2}\sum_{w\in\tau^{n,\circ}}|\tau^n_w|\, f\left(\frac{b_n}{n}\tau^n_w, \frac{b_n}{n}H(w)\right)\xrightarrow[n\to\infty]{(d)+\text{moments}}\Psi_{\mathscr{T}}(f),$$

where f is a bounded continuous real-valued function defined on $\mathbb{T}\times\mathbb{R}_+$.

We improve this result by allowing the function f to blow up as either the mass or the height goes to zero under the stronger assumption $(\xi 2)'$: see Proposition 1.7.1, and more precisely

Theorem 1.7.3 when f is a product of a function of the mass and a function of the height, one of them being a power function. As a particular case, property (i) of Theorem 1.1.1 gives a precise result when f is a power function of the mass and the height. Related to the latter result, we give a complete description of the finiteness of $\Psi_{\mathscr{T}}^{\mathfrak{mh}}(f)$ for power functions f where $\mathscr{T}$ is the stable Lévy tree and we also compute its first moment. We refer to Corollaries 1.6.4 and 1.6.7, and Proposition 1.6.9 for a more general statement. By convention, we write $\Psi_{\mathscr{T}}^{\mathfrak{mh}}(g(x)h(u))$ for $\Psi_{\mathscr{T}}^{\mathfrak{mh}}(f)$ where $f(x,u) = g(x)h(u)$ and we see g as a function of the mass and h as a function of the height. In particular, thanks to (1.1.6), we have for $\alpha, \beta \in \mathbb{R}$ that $\Psi_{\mathscr{T}}^{\mathfrak{mh}}(x^\alpha u^\beta) = \int_{\mathscr{T}} \mathfrak{m}(\mathscr{T}_y)^{\alpha'} \mathfrak{h}(\mathscr{T}_y)^\beta \, \ell(\mathrm{d}y)$ with $\alpha' = \alpha + 1$.

Proposition 1.1.5. *Let $\mathscr{T}$ be the stable Lévy tree with branching mechanism $\psi(\lambda) = \kappa\lambda^\gamma$ and let $\alpha, \beta \in \mathbb{R}$. We have*

$$\gamma\alpha + (\gamma-1)(\beta+1) > 0 \iff \Psi_{\mathscr{T}}^{\mathfrak{mh}}(x^\alpha u^\beta) < \infty \text{ a.s.} \iff \mathbb{E}\left[\Psi_{\mathscr{T}}^{\mathfrak{mh}}(x^\alpha u^\beta)\right] < \infty, \tag{1.1.9}$$

$$\gamma\alpha + (\gamma-1)(\beta+1) \le 0 \iff \Psi_{\mathscr{T}}^{\mathfrak{mh}}(x^\alpha u^\beta) = \infty \text{ a.s.} \iff \mathbb{E}\left[\Psi_{\mathscr{T}}^{\mathfrak{mh}}(x^\alpha u^\beta)\right] = \infty. \tag{1.1.10}$$

For every $\alpha, \beta \in \mathbb{R}$ such that $\gamma\alpha + (\gamma-1)(\beta+1) > 0$, we have

$$\mathbb{E}\left[\Psi_{\mathscr{T}}^{\mathfrak{mh}}(x^\alpha u^\beta)\right] = \frac{1}{\kappa^{1/\gamma}|\Gamma(-1/\gamma)|} \mathrm{B}\big(\alpha + (\beta+1)(1-1/\gamma), 1-1/\gamma\big) \mathbb{E}\left[\mathfrak{h}(\mathscr{T})^\beta\right], \tag{1.1.11}$$

where Γ is the gamma function and B is the beta function. Furthermore, we have $\mathbb{E}\left[\Psi_{\mathscr{T}}^{\mathfrak{mh}}(x^\alpha u^\beta)^p\right] < \infty$ for every $p \ge 1$ such that $p(\gamma\alpha + (\gamma-1)\beta) > 1-\gamma$. In the Brownian case ($\gamma = 2$), for every $\alpha, \beta \in \mathbb{R}$ such that $2\alpha + \beta + 1 > 0$, we have

$$\mathbb{E}\left[\Psi_{\mathscr{T}}^{\mathfrak{mh}}(x^\alpha u^\beta)\right] = \frac{1}{\sqrt{\pi\kappa}} \left(\frac{\pi}{\kappa}\right)^{\beta/2} \xi(\beta) \mathrm{B}\left(\alpha + \frac{\beta+1}{2}, \frac{1}{2}\right), \tag{1.1.12}$$

where ξ is the Riemann xi function defined by $\xi(s) = \frac{1}{2}s(s-1)\pi^{-s/2}\Gamma(s/2)\zeta(s)$ for every $s \in \mathbb{C}$ and ζ is the Riemann zeta function.

Thanks to Duquesne and Wang [60], $\mathbb{E}\left[\mathfrak{h}(\mathscr{T})^\beta\right]$ is finite for all $\beta \in \mathbb{R}$, so that the right-hand side of (1.1.11) is finite.

We conclude the introduction by giving a formula for the distribution of $\mathscr{T}_y$, the subtree above y, when y is chosen according to the length measure $\ell(\mathrm{d}y)$ on the stable Lévy tree $\mathscr{T}$, see Proposition 1.6.3. This is a key result for the proof of Proposition 1.1.5 and it is also interesting by itself (it is in particular related to the additive coalescent and the uniform pruning on the skeleton of the Lévy tree, see Remark 1.6.2 in this direction). Let $\mathbf{N}$ denote the excursion measure of height process H which codes the (unnormalized) stable Lévy tree $\mathscr{T}_H$. (Notice that $\mathscr{T}$ under $\mathbb{P}$ is distributed as $\mathscr{T}_H$ conditionally on $\{\mathfrak{m}(\mathscr{T}_H) = 1\}$ under $\mathbf{N}$.)

Proposition 1.1.6. *Let $\mathscr{T}$ be the stable Lévy tree with branching mechanism $\psi(\lambda) = \kappa\lambda^\gamma$ where $\kappa > 0$ and $\gamma \in (1,2]$. Let f be a nonnegative measurable function defined on $\mathbb{T}$. We have:*

$$\mathbb{E}\left[\int_{\mathscr{T}} f(\mathscr{T}_y)\,\ell(\mathrm{d}y)\right] = \mathbf{N}\left[(1-\mathfrak{m}(\mathscr{T}_H))^{-1/\gamma}\mathbf{1}_{\{\mathfrak{m}(\mathscr{T}_H)<1\}} f(\mathscr{T}_H)\right].$$

The paper is organized as follows. Section 1.2 establishes notation and defines the main objects used in this paper (discrete trees using Neveu's formalism, real trees, Gromov-Hausdorff-Prokhorov topology). In Section 1.3, we give properties of the measure Ψ_T and prove its continuity with respect to T. Section 1.4 introduces the setting of BGW trees and stable Lévy trees and contains a first convergence result for continuous functions. We gather some technical results in Section 1.5. Section 1.6 is devoted to the study of functionals of the mass and height on the stable Lévy tree and Section 1.7 presents the general convergence result for functions that may blow up and describes the phase change. Appendix 1.A introduces a space of measures and studies random elements thereof; its results are used in the proofs of Proposition 1.7.1 and Theorem 1.7.3.

1.2 Definitions and notations

1.2.1 Basic notation

Throughout the sequel, $\mathbb{N} = \{0,1,\dots\}$ will denote the set of integers, $\mathbb{N}^* = \{1,2,\dots\}$ the set of positive integers, $\mathbb{R}$ the set of real numbers, $\mathbb{R}_+ = [0,+\infty)$ the set of nonnegative real numbers and $\mathbb{R}_+^* = (0,+\infty)$ the set of positive real numbers.

1.2.2 Weak convergence in a Polish space

Let (S,ρ) be a Polish metric space. We denote by $\mathscr{B}(S)$ (resp. $\mathscr{B}_+(S)$, resp. $\mathscr{B}_b(S)$) the set of measurable functions defined on S and taking values in $[-\infty,+\infty]$ (resp. in $[0,+\infty]$, resp. in $\mathbb{R}$ and bounded) and by $\mathscr{C}(S)$ (resp. $\mathscr{C}_+(S)$, resp. $\mathscr{C}_b(S)$) the set of continuous real-valued functions defined on S (resp. nonnegative, resp. bounded). For $f \in \mathscr{B}(S)$, we set $\|f\|_\infty = \sup_{x\in S}|f(x)|$. For $f \in \mathscr{C}_b(S)$, we define its Lipschitz and bounded Lipschitz norm:

$$\|f\|_{\mathrm{L}} = \sup_{x\neq y}\frac{|f(x)-f(y)|}{\rho(x,y)} \quad\text{and}\quad \|f\|_{\mathrm{BL}} = \|f\|_\infty + \|f\|_{\mathrm{L}}.$$

We denote by $\mathscr{M}(S)$ the set of nonnegative finite measures on S. For every $\mu \in \mathscr{M}(S)$ and $f \in \mathscr{B}_+(S)$, we write $\mu(f) = \int f(x)\,\mu(\mathrm{d}x)$. We also write $f\mu$ for the measure $f(x)\mu(\mathrm{d}x)$. The set $\mathscr{M}(S)$ is endowed with the topology of weak convergence which can be metrized (see [40, Section 8.3 and Theorem 8.3.2]) by the bounded Lipschitz distance (also known as the Kantorovich-Rubinstein distance): if $\mu,\nu \in \mathscr{M}(S)$, set

$$d_{\mathrm{BL}}(\mu,\nu) = \sup\left\{|\mu(f)-\nu(f)|,\ f\in\mathscr{C}_b(S) \text{ such that } \|f\|_{\mathrm{BL}} \le 1\right\}.$$

Moreover, the space $(\mathscr{M}(S), d_{\mathrm{BL}})$ is Polish by [40, Theorem 8.9.4]. We also recall the total variation distance given by

$$\begin{aligned} d_{\mathrm{TV}}(\mu,\nu) &= \sup\{\left|\mu(A)-\nu(A)\right|,\ A\subset S \text{ measurable}\} \\ &= \frac{1}{2}\sup\left\{|\mu(f)-\nu(f)|,\ f\in\mathscr{B}(S) \text{ such that } \|f\|_\infty \le 1\right\} + \frac{1}{2}\left|\mu(1)-\nu(1)\right|. \end{aligned}$$

1.2.3 Discrete trees

We recall Neveu's formalism for rooted ordered discrete trees. Let $\mathscr{U} = \cup_{n\ge 0}(\mathbb{N}^*)^n$ be the set of labels with the convention $(\mathbb{N}^*)^0 = \{\emptyset\}$. If $v = (v^1,\dots,v^n)\in\mathscr{U}$, we denote by $H(v) = n$. By convention, we set $H(\emptyset) = 0$. If $v = (v^1,\dots,v^n), w = (w^1,\dots,w^m)\in\mathscr{U}$, we write $vw = (v^1,\dots,v^n,w^1,\dots,w^m)$ for the concatenation of v and w. In particular, $v\emptyset = \emptyset v = v$. We say that v is an ancestor of w and write $v\preccurlyeq w$ if there exists $u\in\mathscr{U}$ such that $w = vu$. If $v\preccurlyeq w$ and $v\neq w$ then we shall write $v\prec w$. The mapping $\mathrm{pr}\colon \mathscr{U}\setminus\{\emptyset\}\to\mathscr{U}$ is defined by $\mathrm{pr}(v^1,\dots,v^n) = (v^1,\dots,v^{n-1})$ (i.e. $\mathrm{pr}(v)$ is the parent of v). A finite rooted ordered tree $\mathbf{t}$ is a finite subset of $\mathscr{U}$ such that

(i) $\emptyset\in\mathbf{t}$,

(ii) $v\in\mathbf{t}\setminus\{\emptyset\}\Rightarrow \mathrm{pr}(v)\in\mathbf{t}$,

(iii) for every $v\in\mathbf{t}$, there exists a finite integer $k_v(\mathbf{t})\ge 0$ such that, for every $j\in\mathbb{N}^*$, $vj\in\mathbf{t}$ if and only if $1\le j\le k_u(\mathbf{t})$.

The number $k_v(\mathbf{t})$ is interpreted as the number of children of the vertex v in $\mathbf{t}$, $H(v)$ is its generation, $\mathrm{pr}(v)$ is its parent and more generally, the vertices $v, \mathrm{pr}(v), \mathrm{pr}^2(v), \dots \mathrm{pr}^{H(v)}(v) = \emptyset$ are its ancestors. The vertex v is called a leaf (resp. internal vertex) if $k_v(\mathbf{t}) = 0$ (resp. $k_v(\mathbf{t})>0$). The vertex $\emptyset$ is called the root of $\mathbf{t}$. We denote the set of leaves by $\mathrm{Lf}(\mathbf{t})$ and the set of internal vertices by $\mathbf{t}^\circ$. If $v\in\mathbf{t}$, we define the subtree $\mathbf{t}_v$ of $\mathbf{t}$ above v as

$$\mathbf{t}_v = \{w\in\mathscr{U}:\ vw\in\mathbf{t}\}.$$

Moreover, for every $0\le k\le H(v)$, we define the subtree $\mathbf{t}_{k,v}$ of $\mathbf{t}$ above level k containing v as

$$\mathbf{t}_{k,v} = \mathbf{t}_{\mathrm{pr}^{H(v)-k}(v)}$$

where $\mathrm{pr}^{H(v)-k}(v)$ is the unique ancestor of v with height k, with the convention that $\mathrm{pr}^0(v) = v$. We denote by $|\mathbf{t}| = \mathrm{Card}(\mathbf{t})$ the number of vertices of $\mathbf{t}$ and by $\mathfrak{h}(\mathbf{t}) = \sup_{v\in\mathbf{t}} H(v)$ the height of $\mathbf{t}$.

1.2.4 Real trees

We recall the formalism of real trees, see [69]. A metric space (T,d) is a real tree if the following two properties hold for every $x,y\in T$.

(i) (Unique geodesics). There exists a unique isometric map $f_{x,y}\colon [0,d(x,y)] \to T$ such that $f_{x,y}(0) = x$ and $f_{x,y}(d(x,y)) = y$.

(ii) (Loop-free). If φ is a continuous injective map from $[0,1]$ into T such that $\varphi(0) = x$ and $\varphi(1) = y$, then we have $\varphi([0,1]) = f_{x,y}\left([0,d(x,y)]\right)$.

For a rooted real tree $(T,\emptyset,d)$, that is a real tree with a distinguished vertex $\emptyset \in T$ called the root, we define the set of leaves by

$$\mathrm{Lf}(T) = \{x \in T \setminus \{\emptyset\}\colon\ T \setminus \{x\} \text{ is connected}\},$$

with the convention that $\mathrm{Lf}(T) = \{\emptyset\}$ if $T = \{\emptyset\}$. A weighted rooted real tree $(T,\emptyset,d,\mu)$ is a rooted real tree $(T,\emptyset,d)$ equipped with a nonnegative finite measure μ. In what follows, real trees will always be weighted and rooted and we will simply call them real trees.

Let us consider a real tree $(T,\emptyset,d,\mu)$. The total mass of the tree T is defined by $\mathfrak{m}(T) = \mu(T)$ and its height by $\mathfrak{h}(T) = \sup_{x\in T} H(x) \in [0,\infty]$, with $H(x) = d(\emptyset,x)$ denoting the height of x. Note that if (T,d) is compact, then $\mathfrak{h}(T) < \infty$. The range of the mapping $f_{x,y}$ described in (i) above is denoted by $[\![x,y]\!]$ (this is the line segment between x and y in the tree). We also write $[\![x,y[\![= [\![x,y]\!] \setminus \{y\}$. In particular, $[\![\emptyset,x]\!]$ is the path going from the root to x, which we will interpret as the ancestral line of vertex x. We define a partial order on the tree by setting $x \preccurlyeq y$ (x is an ancestor of y) if and only if $x \in [\![\emptyset,y]\!]$. If $x,y \in T$, there is a unique $z \in T$ such that $[\![\emptyset,x]\!] \cap [\![\emptyset,y]\!] = [\![\emptyset,z]\!]$. We write $z = x \wedge y$ and call it the most recent common ancestor of x and y. Let $x \in T$ be a vertex. Let $r \in [0,H(x)]$. We denote by $x_r \in T$ the unique ancestor of x with height $H(x_r) = r$. As in the discrete case, we also define the subtree T_x of T above x as

$$T_x = \left\{y \in T\colon\ x \preccurlyeq y\right\},$$

and the subtree $T_{r,x} = T_{x_r}$ of T above level r containing x as

$$T_{r,x} = \left\{y \in T\colon\ H(x\wedge y) \geq r\right\} = T_{x_r}.$$

Then T_x (resp. $T_{r,x}$) can be naturally viewed as a real tree, rooted at x (resp. at x_r) and endowed with the distance d and the measure $\mu_{|T_x} = \mu(\cdot \cap T_x)$ (resp. the measure $\mu_{|T_{r,x}}$). Note that $T_{0,x} = T$ and $T_{H(x),x} = T_x$.

Remark 1.2.1. We recall the construction of a real tree from an excursion path, see e.g. [69, Example 3.14] or [58, Section 2.1]. Let e be a positive excursion path, that is $e \in \mathscr{C}_+(\mathbb{R}_+)$ such that $e(0) = 0$, $e(s) > 0$ for $0 < s < \sigma$ and $e(s) = 0$ for $s \geq \sigma$ where $\sigma := \inf\{s > 0\colon\ e(s) = 0\} \in (0,\infty)$ is the duration of the excursion. Set $d_e(t,s) = e(t) + e(s) - 2\inf_{[t\wedge s, t\vee s]} e$ for every $t,s \in [0,\sigma]$ and define an equivalence relation on $[0,\sigma]$ by letting $t \sim_e s$ if and only if $d_e(t,s) = 0$. The real tree T_e coded by e is defined as the quotient space $[0,\sigma]/\sim_e$ rooted at $p(0)$ where $p\colon [0,\sigma] \to T_e$ is the quotient map and equipped with the distance d_e and the pushforward measure $\lambda \circ p^{-1}$ where λ is the Lebesgue measure on $[0,\sigma]$. This defines a compact weighted rooted real tree. Notice that the mass and height of T_e are given by $\mathfrak{m}(T_e) = \sigma$ and $\mathfrak{h}(T_e) = \|e\|_\infty$.

We will need to view discrete trees as real trees. Let $\mathbf{t}$ be a finite rooted ordered tree and let $a > 0$. Suppose that $\mathbf{t}$ is embedded into the plane such that the edges are straight lines with length a that only intersect at their incident vertices. Denote by $\pi_{\mathbf{t},a}: \mathbf{t} \to \mathbb{R}^2$ the embedding and by $a\mathbf{t} = \pi_{\mathbf{t},a}(\mathbf{t}) \subset \mathbb{R}^2$ the embedded set. Moreover, for a vertex $v \in \mathbf{t}$, we denote by $av = \pi_{\mathbf{t},a}(v)$ the corresponding vertex in $a\mathbf{t}$. Then $a\mathbf{t}$ can be considered as a compact real tree $(a\mathbf{t}, d_{\mathbf{t}}, \mu_{\mathbf{t}})$: the distance $d_{\mathbf{t}}(x, y)$ between two points $x, y \in a\mathbf{t}$ is defined as the shortest length of a curve that connects x and y, and the measure $\mu_{\mathbf{t}}$ is the pushforward of the uniform probability measure on the vertex set of $\mathbf{t}$ by the embedding $\pi_{\mathbf{t},a}$. In other words, $a\mathbf{t}$ is obtained from $\mathbf{t}$ by connecting every vertex to its children in such a way that all edges have length a and is equipped with the measure $\mu_{\mathbf{t}}$ supported on the set $\{av\colon\ v \in \mathbf{t}\}$ and satisfying $\mu_{\mathbf{t}}(\{av\}) = 1/|\mathbf{t}|$ for every $v \in \mathbf{t}$. The tree $a\mathbf{t}$ is naturally rooted at $a\varnothing$ (also denoted $\varnothing$). Notice that vertices of the form av with $v \in \mathbf{t}$ are precisely those vertices in $a\mathbf{t}$ whose heights are integer multiples of a. Finally, to simplify notation, for every $v \in \mathbf{t}$, we will write $a\mathbf{t}_v$ instead of $(a\mathbf{t})_{av}$ for the subtree of $a\mathbf{t}$ above av. We stress that, unless $v = \varnothing$, the measure of the compact real tree $a\mathbf{t}_v$ has mass less than one, whereas the measure of the compact real tree $a(\mathbf{t}_v)$ is by definition a probability measure.

1.2.5 Gromov-Hausdorff-Prokhorov topology

Denote by $\mathbb{T}$ the set of measure-preserving and root-preserving isometry classes of compact real trees. We will often identify a class with an element of this class. So we shall write that $(T, \varnothing, d, \mu) \in \mathbb{T}$ if $(T, \varnothing, d)$ is a rooted compact real tree and μ is a nonnegative finite measure on T. When there is no ambiguity, we may write T for $(T, \varnothing, d, \mu)$.

We start by giving the standard definition of the Gromov-Hausdorff-Prokhorov distance. Let (E, δ) be a metric space. Given a non-empty subset $A \subset E$ and $\varepsilon > 0$, the ε-neighborhood of A is $A^\varepsilon = \{x \in E\colon\ d(x, A) < \varepsilon\}$. The Hausdorff distance δ_{H} between two non-empty subsets $A, B \subset E$ is defined by

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

Next, denoting by $\mathscr{B}(E)$ the Borel σ-field on (E, δ), the Lévy-Prokhorov distance between two finite nonnegative measures μ, ν on $(E, \mathscr{B}(E))$ is

$$\delta_{\mathrm{P}}(\mu, \nu) = \inf\{\varepsilon > 0\colon\ \mu(A) \leq \nu(A^\varepsilon) + \varepsilon \text{ and } \nu(A) \leq \mu(A^\varepsilon) + \varepsilon,\ \forall A \in \mathscr{B}(E)\}.$$

We can now give the standard distance used to define the Gromov-Hausdorff-Prokhorov topology. For two compact real trees $(T, \varnothing, d, \mu), (T', \varnothing', d', \mu') \in \mathbb{T}$, set

$$d^\circ_{\mathrm{GHP}}(T, T') = \inf\left\{\delta(\varphi(\varnothing), \varphi'(\varnothing')) \vee \delta_{\mathrm{H}}(\varphi(T), \varphi'(T')) \vee \delta_{\mathrm{P}}(\mu \circ \varphi^{-1}, \mu' \circ \varphi'^{-1})\right\}, \tag{1.2.1}$$

where the infimum is taken over all isometries $\varphi\colon T \to E$ and $\varphi'\colon T' \to E$ into a common metric space (E, δ). This defines a metric which induces the Gromov-Hausdorff-Prokhorov topology on $\mathbb{T}$.

It will be convenient for our purposes to define another metric which induces the same topology on $\mathbb{T}$. Let $(T,\varnothing,d,\mu),(T',\varnothing',d',\mu')\in\mathbb{T}$. Recall that a correspondence between T and T' is a subset $\mathscr{R}\subset T\times T'$ such that for every $x\in T$, there exists $x'\in T'$ such that $(x,x')\in\mathscr{R}$, and conversely, for every $x'\in T'$, there exists $x\in T$ such that $(x,x')\in\mathscr{R}$. In other words, if we denote by $p\colon T\times T'\to T$ (resp. $p'\colon T\times T'\to T'$) the canonical projection on T (resp. on T'), a correspondence is a subset $\mathscr{R}\subset T\times T'$ such that $p(\mathscr{R})=T$ and $p'(\mathscr{R})=T'$. If $\mathscr{R}$ is a correspondence between T and T', its distortion is defined by

$$\mathrm{dis}(\mathscr{R})=\sup\left\{\left|d(x,y)-d'(x',y')\right|:\ (x,x'),(y,y')\in\mathscr{R}\right\}.$$

Next, for any nonnegative finite measure m on $T\times T'$, we define its discrepancy with respect to μ and μ' by

$$\mathrm{D}(m;\mu,\mu')=d_{\mathrm{TV}}(m\circ p^{-1},\mu)+d_{\mathrm{TV}}(m\circ p'^{-1},\mu').$$

Then the Gromov-Hausdorff-Prokhorov distance between T and T' is defined as

$$d_{\mathrm{GHP}}(T,T')=\inf\left\{\frac{1}{2}\,\mathrm{dis}(\mathscr{R})\vee\mathrm{D}(m;\mu,\mu')\vee m(\mathscr{R}^c)\right\},\tag{1.2.2}$$

where the infimum is taken over all correspondences $\mathscr{R}$ between T and T' such that $(\varnothing,\varnothing')\in\mathscr{R}$ and all nonnegative finite measures m on $T\times T'$. It can be verified that d_{GHP} is indeed a distance on $\mathbb{T}$ which is equivalent to d°_{GHP} and that the space $(\mathbb{T},d_{\mathrm{GHP}})$ is a Polish metric space, see [13].

We gather some facts about the Gromov-Hausdorff-Prokhorov distance that will be useful later. We refer the reader to [13] or [141]. We have that

$$\frac{1}{2}\left|\mathfrak{h}(T)-\mathfrak{h}(T')\right|\vee\left|\mathfrak{m}(T)-\mathfrak{m}(T')\right|\le d_{\mathrm{GHP}}(T,T')\le\left(\mathfrak{h}(T)+\mathfrak{h}(T')\right)\vee\left(\mathfrak{m}(T)+\mathfrak{m}(T')\right).\tag{1.2.3}$$

When $T'=\{\varnothing\}$ is the trivial tree consisting only of the root with mass 0, we have

$$\frac{1}{2}\mathfrak{h}(T)\vee\mathfrak{m}(T)\le d_{\mathrm{GHP}}(T,\{\varnothing\})\le\mathfrak{h}(T)\vee\mathfrak{m}(T).\tag{1.2.4}$$

We consider the subset of $\mathbb{T}$ of trees with either height or mass equal to 0:

$$\mathbb{T}_0=\left\{T\in\mathbb{T}:\ \mathfrak{m}(T)=0\text{ or }\mathfrak{h}(T)=0\right\}.\tag{1.2.5}$$

Note that $\mathbb{T}_0\subset\mathbb{T}$ is a closed subset since the mappings $\mathfrak{m}\colon\mathbb{T}\to\mathbb{R}$ and $\mathfrak{h}\colon\mathbb{T}\to\mathbb{R}$ are continuous with respect to the Gromov-Hausdorff-Prokhorov topology, thanks to (1.2.3). We now give bounds for the distance of a tree T to $\mathbb{T}_0$ which are similar to (1.2.4).

Lemma 1.2.2. *Let $T\in\mathbb{T}$. Then we have*

$$\frac{1}{2}\mathfrak{h}(T)\wedge\mathfrak{m}(T)\le d_{\mathrm{GHP}}(T,\mathbb{T}_0)\le\mathfrak{h}(T)\wedge\mathfrak{m}(T).\tag{1.2.6}$$

Proof. Let $(T, d, \emptyset, \mu) \in \mathbb{T}$ and $\delta > d_{\mathrm{GHP}}(T, \mathbb{T}_0)$. Then there exists $T' \in \mathbb{T}_0$ such that $d_{\mathrm{GHP}}(T, T') \leq \delta$. By (1.2.3), we get

$$\frac{1}{2}\left|\mathfrak{h}(T) - \mathfrak{h}(T')\right| \vee \left|\mathfrak{m}(T) - \mathfrak{m}(T')\right| \leq \delta.$$

But since $T' \in \mathbb{T}_0$, either $\mathfrak{h}(T') = 0$ or $\mathfrak{m}(T') = 0$. Therefore, either $\mathfrak{h}(T) \leq 2\delta$ or $\mathfrak{m}(T) \leq \delta$. Since $\delta > d_{\mathrm{GHP}}(T, \mathbb{T}_0)$ is arbitrary, this yields the lower bound.

To prove the upper bound, let T' be the same real tree as T but endowed with the zero measure $\mu' = 0$, and take $\mathscr{R} = \{(x, x) \colon\ x \in T\}$ and m the zero measure on $T \times T'$. Then $\mathrm{dis}(\mathscr{R}) = 0$, $m(\mathscr{R}^c) = 0$ and $\mathrm{D}(m; \mu, \mu') = \mu(T) = \mathfrak{m}(T)$. It follows that $d_{\mathrm{GHP}}(T, T') \leq \mathfrak{m}(T)$. Note that $T' \in \mathbb{T}_0$, therefore

$$d_{\mathrm{GHP}}(T, \mathbb{T}_0) \leq d_{\mathrm{GHP}}(T, T') \leq \mathfrak{m}(T).$$

Next, let $T'' = \{\emptyset\}$ be the trivial tree consisting only of the root with mass $\mathfrak{m}(T)$, i.e. endowed with the measure $\mu'' = \mathfrak{m}(T)\delta_\emptyset$. Take $\mathscr{R} = T \times \{\emptyset\}$ and $m(A \times B) = \mu(A)\delta_\emptyset(B)$. Then, we have $\mathscr{R}^c = \emptyset$, so $m(\mathscr{R}^c) = 0$. Moreover, we have

$$\mathrm{dis}(\mathscr{R}) = \sup\left\{|d(x, y)| \colon\ x, y \in T\right\} \leq 2\mathfrak{h}(T).$$

Since $m \circ p^{-1} = \mu$ and $m \circ p''^{-1} = \mathfrak{m}(T)\delta_\emptyset = \mu''$, we get $\mathrm{D}(m, \mu, \mu'') = 0$. It follows that $d_{\mathrm{GHP}}(T, T'') \leq \mathfrak{h}(T)$. Since $T'' \in \mathbb{T}_0$, we deduce that

$$d_{\mathrm{GHP}}(T, \mathbb{T}_0) \leq d_{\mathrm{GHP}}(T, T'') \leq \mathfrak{h}(T).$$

This finishes the proof of the upper bound. □

1.3 A finite measure indexed by a tree

Let $(T, \emptyset, d, \mu)$ be a compact real tree. Let $x \in T$ and $r \in [0, H(x)]$, where $H(x) = d(\emptyset, x)$. Recall that $T_{r,x} = \{y \in T : H(x \wedge y) \geq r\}$ is the subtree containing x and starting at height r, endowed with the distance d and the measure $\mu_{|T_{r,x}}$. It is straightforward to check that $T_{r,x}$ is a compact real tree and thus belongs to $\mathbb{T}$. Define a nonnegative measure Ψ_T on $\mathbb{T} \times \mathbb{R}_+$ as follows: for every $f \in \mathscr{B}_+(\mathbb{T} \times \mathbb{R}_+)$,

$$\Psi_T(f) = \int_T \mu(\mathrm{d}x) \int_0^{H(x)} f\left(T_{r,x}, r\right) \mathrm{d}r. \tag{1.3.1}$$

As we will consider functions depending only on the mass and height of the subtrees, we introduce the measure $\Psi_T^{\mathfrak{m}\mathfrak{h}}$ on $\mathbb{R}_+^2$ defined as follows: for every $f \in \mathscr{B}_+(\mathbb{R}_+^2)$,

$$\Psi_T^{\mathfrak{m}\mathfrak{h}}(f) = \int_T \mu(\mathrm{d}x) \int_0^{H(x)} f\left(\mathfrak{m}(T_{r,x}), \mathfrak{h}(T_{r,x})\right) \mathrm{d}r. \tag{1.3.2}$$

Lemma 1.3.1. *Let T be a compact real tree. The mapping $(r, x) \mapsto T_{r,x}$ from $\{(r, x) \in \mathbb{R}_+ \times T \colon\ r \leq H(x)\}$ to $\mathbb{T}$ is measurable with respect to the Borel σ-fields. Furthermore, the measure Ψ_T is finite and does not depend on the choice of representative in the equivalence class in $\mathbb{T}$ of T.*

Proof. Let $(T, \emptyset, d, \mu)$ be a compact real tree and set $A := \{(r, x) \in \mathbb{R}_+ \times T\colon\ r \leq H(x)\}$. For every $(r, x) \in A$, recall that $x_r \in T$ is the unique ancestor of x with height $H(x_r) = r$. We start by showing that the mapping $(r, x) \mapsto x_r$ is continuous from A to T. Let $(r, x), (s, y) \in A$. Without loss of generality, we can assume that $r \geq s$. If $H(x \wedge y) \geq s$, then we have $y_s \preccurlyeq x$ and thus $y_s \preccurlyeq x_r$. This implies that $d(x_r, y_s) = r - s$. If $H(x \wedge y) < s$, then we have $x_r \in [\![x \wedge y, x]\!]$ and $y_s \in [\![x \wedge y, y]\!]$. This implies that x_r and y_s belong to $[\![x, y]\!]$, and thus $d(x_r, y_s) \leq d(x, y)$. In all cases, we have

$$d(x_r, y_s) \leq d(x, y) + |r - s|.$$

This proves that $(r, x) \mapsto x_r$ is continuous.

The mapping $y \mapsto T_y$ from T to $\mathbb{T}$ is continuous from below, in the sense that for $y \in T$

$$\lim_{\substack{z \to y \\ z \preccurlyeq y}} d_{\mathrm{GHP}}(T_z, T_y) = 0. \tag{1.3.3}$$

To see this, let $\delta > 0$, $y \in T$ and $(y_n,\ n \in \mathbb{N})$ be a sequence in T converging to y such that $y_n \preccurlyeq y$ for every $n \in \mathbb{N}$. Notice that since T is compact, it holds that there is a finite number of subtrees with height larger than δ attached to the branch $[\![\emptyset, y]\!]$. Thus, there are no subtrees with height larger than δ attached to $[\![y_n, y[\![$ for n larger than some n_0. Moreover, since $T_y = \bigcap_{n \in \mathbb{N}} T_{y_n}$, we get that $\lim_{n\to\infty} \mu(T_{y_n}) = \mu(T_y)$ implying that the mass of the subtrees attached to $[\![y_n, y[\![$ goes to 0 as n goes to infinity.

Define a correspondence between T_{y_n} and T_y by

$$\mathscr{R} := \left\{(z, z)\colon\ z \in T_y\right\} \bigcup \left\{(z, y)\colon\ z \in T_{y_n} \setminus T_y\right\}.$$

It is straightforward to check that $\mathrm{dis}(\mathscr{R}) \leq 2(\delta + d(y_n, y))$ for $n \geq n_0$. Consider the measure on $T_{y_n} \times T_y$ defined by $m(\mathrm{d}x, \mathrm{d}z) = \mu_{|T_y}(\mathrm{d}z)\delta_z(\mathrm{d}x) = \mu_{|T_y}(\mathrm{d}x)\delta_x(\mathrm{d}z)$. Then we have $\mathrm{D}(m; \mu_{|T_{y_n}}, \mu_{|T_y}) \leq \mu(T_{y_n}) - \mu(T_y)$ and $m(\mathscr{R}^c) = 0$. It follows from (1.2.2) that

$$\limsup_{n\to\infty} d_{\mathrm{GHP}}(T_{y_n}, T_y) \leq \limsup_{n\to\infty} \left(\delta + d(y_n, y) + \mu(T_{y_n}) - \mu(T_y)\right) = \delta.$$

Since $\delta > 0$ is arbitrary, (1.3.3) readily follows.

Now it is not difficult to see that the continuity from below (1.3.3) of the mapping $y \mapsto T_y$ implies its measurability. By composition, it follows that the mapping $(r, x) \mapsto T_{r,x} = T_{x_r}$ from A to $\mathbb{T}$ is measurable.

Next, notice that Ψ_T is finite since

$$\Psi_T(1) = \int_T H(x)\mu(\mathrm{d}x) \leq \mathfrak{h}(T)\mathfrak{m}(T) < \infty.$$

Finally, let $f \in \mathscr{B}_+(\mathbb{T} \times \mathbb{R}_+)$ and $(T, \emptyset, d, \mu), (T', \emptyset', d', \mu')$ be two compact real trees such that there is a measure-preserving and root-preserving isometry $\varphi\colon T \to T'$. This means that

φ is an isometry satisfying $\mu' = \mu \circ \varphi^{-1}$ and $\varphi(\emptyset) = \emptyset'$. Moreover, for every $x, y \in T$, since $H(x \wedge y) = 2^{-1}\left(d(\emptyset, x) + d(\emptyset, y) - d(x, y)\right)$, we deduce that

$$H(x \wedge y) = H(\varphi(x) \wedge \varphi(y)).$$

Using this and the definitions of $T_{r,x}$ and $T'_{r,\varphi(x)}$, it is easy to see that, for every $x \in T$ and $r \in [0, H(x)]$, φ induces a measure-preserving and root-preserving isometry from $T_{r,x}$ to $T'_{r,\varphi(x)}$ and therefore $f(T_{r,x}, r) = f(T'_{r,\varphi(x)}, r)$. Since $H(x) = H(\varphi(x))$, it follows that

$$\begin{aligned}
\Psi_T(f) &= \int_T \mu(\mathrm{d}x) \int_0^{H(x)} f(T_{r,x}, r)\,\mathrm{d}r \\
&= \int_T \mu(\mathrm{d}x) \int_0^{H(\varphi(x))} f(T'_{r,\varphi(x)}, r)\,\mathrm{d}r \\
&= \int_{T'} \mu \circ \varphi^{-1}(\mathrm{d}y) \int_0^{H(y)} f(T'_{r,y}, r)\,\mathrm{d}r \\
&= \Psi_{T'}(f).
\end{aligned}$$

This proves that Ψ_T does not depend on the choice of representative in the equivalence class of T which completes the proof. □

Recall that $\mathrm{Lf}(T)$ is the set of leaves of T. It is well known that there exists a unique σ-finite measure ℓ on $(T, \mathscr{B}(T))$, called the length measure, such that $\ell(\mathrm{Lf}(T)) = 0$ and $\ell([\![x, y]\!]) = d(x, y)$, see e.g. [69, Chapter 4, §4.3.5]. The next result gives an alternative expression for Ψ_T in terms of the length measure.

Proposition 1.3.2. *Let $(T, \emptyset, d, \mu)$ be a compact real tree. For every $f \in \mathscr{B}_+(\mathbb{T} \times \mathbb{R}_+)$, we have*

$$\Psi_T(f) = \int_T \mu(T_y) f(T_y, H(y))\,\ell(\mathrm{d}y). \tag{1.3.4}$$

Proof. Let $(T, \emptyset, d, \mu)$ be a compact real tree and $f \in \mathscr{B}_+(\mathbb{T} \times \mathbb{R}_+)$. Notice that $\{(x, y) \in T^2 :\ y \preccurlyeq x\} = \{(x, y) \in T^2 :\ d(\emptyset, x) = d(\emptyset, y) + d(x, y)\}$ is closed in T^2 and thus measurable. Moreover, the mapping $y \longmapsto T_y$ is measurable from T to $\mathbb{T}$ by the proof of Lemma 1.3.1. Thus the mapping $(x, y) \longmapsto \mathbf{1}_{\{y \preccurlyeq x\}} f(T_y, H(y))$ is measurable. By Fubini's theorem, it follows that

$$\begin{aligned}
\int_T \mu(T_y) f\left(T_y, H(y)\right) \ell(\mathrm{d}y) &= \int_T \mu(\mathrm{d}x) \int_T \mathbf{1}_{\{y \preccurlyeq x\}} f\left(T_y, H(y)\right) \ell(\mathrm{d}y) \\
&= \int_T \mu(\mathrm{d}x) \int_{[\![\emptyset, x]\!]} f\left(T_y, H(y)\right) \ell(\mathrm{d}y).
\end{aligned}$$

Let $x \in T$ and let $f_{\emptyset,x} : [0, H(x)] \to [\![\emptyset, x]\!]$ be the unique isometry such that $f_{\emptyset,x}(0) = \emptyset$ and $f_{\emptyset,x}(H(x)) = x$. Using that $\ell_{|[\![\emptyset,x]\!]} = \lambda \circ f_{\emptyset,x}^{-1}$ where λ is the Lebesgue measure on $[0, H(x)]$, we get that

$$\int_{[\![\emptyset, x]\!]} f\left(T_y, H(y)\right) \ell(\mathrm{d}y) = \int_0^{H(x)} f\left(T_{f_{\emptyset,x}(r)}, H(f_{\emptyset,x}(r))\right) \mathrm{d}r.$$

Since $f_{\emptyset,x}$ is an isometry, for every $r \in [0, H(x)]$, $f_{\emptyset,x}(r)$ is the unique ancestor of x at height r, that is x_r, and $H(f_{\emptyset,x}(r)) = r$. As $T_{f_{\emptyset,x}(r)} = T_{x_r} = T_{r,x}$ for every $r \in [0, H(x)]$, it follows that

$$\int_T \mu(T_y) f\left(T_y, H(y)\right) \ell(\mathrm{d}y) = \int_T \mu(\mathrm{d}x) \int_0^{H(x)} f\left(T_{r,x}, r\right) \mathrm{d}r.$$

This concludes the proof. □

The main result of this section concerns the continuity of the mapping $\Psi\colon T \mapsto \Psi_T$.

Proposition 1.3.3. *The mapping $\Psi\colon T \mapsto \Psi_T$, from $\mathbb{T}$ endowed with the Gromov-Hausdorff-Prokhorov topology to $\mathscr{M}(\mathbb{T} \times \mathbb{R}_+)$ endowed with the topology of weak convergence, is well defined and continuous.*

The end of this section is devoted to the proof of Proposition 1.3.3. For T a compact real tree, $x \in T$, $s \in [0, +\infty]$, $r \in [0, s \wedge H(x)]$, we define the following set of elements of T such that their common ancestor with x has height in $[r, s]$:

$$T_{[r,s],x} = \{y \in T\colon\ H(y \wedge x) \in [r, s]\}.$$

Recall that x_r is the ancestor of x at height r in T, and is also seen as the root of the tree $T_{r,x}$. We shall see $T_{[r,s],x}$ as a compact real tree rooted at x_r with measure $\mu_{|T_{[r,s],x}} = \mu(\cdot \cap T_{[r,s],x})$ and thus $T_{[r,s],x} \in \mathbb{T}$. Recall that $\mathfrak{m}(T_{[r,s],x}) = \mu(T_{[r,s],x})$ denotes its mass and $\mathfrak{h}(T_{[r,s],x}) = \sup\{H(y)\colon\ y \in T_{[r,s],x} \subset T\} - r$ its height. Notice in particular that $T_{[r,+\infty],x} = T_{r,x}$ for $r \in [0, H(x)]$.

We first establish an estimate for the Gromov-Hausdorff-Prokhorov distance between subtrees of two real trees in terms of the distance between the trees themselves.

Lemma 1.3.4. *Let T, T' be compact real trees and let $\delta > d_{\mathrm{GHP}}(T, T')$. Let $\mathscr{R}$ be a correspondence between T and T' such that $(\emptyset, \emptyset') \in \mathscr{R}$ and let m be a measure on $T \times T'$ such that*

$$\frac{1}{2}\,\mathrm{dis}(\mathscr{R}) \vee \mathrm{D}(m; \mu, \mu') \vee m(\mathscr{R}^c) \le \delta.$$

Then for every (x, x') in $\mathscr{R}$ and every $r \ge 0$ such that $6\delta \le r \le H(x) \wedge H(x')$, we have

$$d_{\mathrm{GHP}}(T_{r,x}, T'_{r,x'}) \le 8\delta + 2\mathfrak{m}\left(T_{[r-6\delta, r+6\delta],x}\right) + 2\mathfrak{h}(T_{[r-3\delta, r+6\delta],x}). \tag{1.3.5}$$

Proof. We shall bound $d_{\mathrm{GHP}}(T_{r,x}, T'_{r,x'})$ from above by

$$\frac{1}{2}\,\mathrm{dis}(\widetilde{\mathscr{R}}) \vee \mathrm{D}(\widetilde{m}; \widetilde{\mu}, \widetilde{\mu}') \vee \widetilde{m}(\widetilde{\mathscr{R}}^c)$$

where $\widetilde{\mathscr{R}}$ is a well chosen correspondence between $T_{r,x}$ and $T'_{r,x'}$ and $\widetilde{m}$ (resp. $\widetilde{\mu}$, $\widetilde{\mu}'$) is the restriction of the measure m (resp. μ, μ') to $T_{r,x} \times T'_{r,x'}$ (resp. $T_{r,x}$, $T'_{r,x'}$). Notice that, for every $(t, t'), (s, s') \in \mathscr{R}$, we have

$$\left|d(t, s) - d'(t', s')\right| \le \mathrm{dis}(\mathscr{R}) \le 2\delta. \tag{1.3.6}$$

In particular, taking $(s,s') = (\emptyset,\emptyset') \in \mathscr{R}$ yields $\left|H(t) - H(t')\right| \le 2\delta$. Using this, we get that for $(t,t') \in \mathscr{R}$

$$H(t\wedge x) - 3\delta \le H(t'\wedge x') \le H(t\wedge x) + 3\delta. \tag{1.3.7}$$

Let $(t,t') \in \mathscr{R}$. Assume that $H(t\wedge x) \ge r + 3\delta$. Then, we get that $t \in T_{r,x}$ and that $H(t'\wedge x') \ge r$ by (1.3.7), that is $t' \in T'_{r,x'}$. This gives that $(t,t') \in T_{r,x}\times T'_{r,x'}$. Similarly, if $H(t'\wedge x') \ge r+3\delta$, we get $(t,t') \in T_{r,x}\times T'_{r,x'}$. Therefore, the following set

$$\widetilde{\mathscr{R}} = \{(t,t') \in \mathscr{R} \colon \max(H(t\wedge x), H(t'\wedge x')) \ge r+3\delta\} \bigcup \left(T_{[r,r+3\delta],x}\times\{x'_r\}\right) \bigcup \left(\{x_r\}\times T'_{[r,r+3\delta],x'}\right)$$

is a correspondence between $T_{r,x}$ and $T'_{r,x'}$. Using (1.3.6) and (1.3.7) and distinguishing according to whether an element $(t,t') \in \widetilde{\mathscr{R}}$ lies or not in $\mathscr{R}$, it is not difficult to establish the following bound for its distortion:

$$\operatorname{dis}(\widetilde{\mathscr{R}}) \le 10\delta + 2\mathfrak{h}(T_{[r,r+3\delta],x}) + 2\mathfrak{h}(T'_{[r,r+3\delta],x'}). \tag{1.3.8}$$

Denote by $\widetilde{m}$ the restriction of the measure m to $T_{r,x}\times T'_{r,x'}$. Routine arguments yield the following bound for its distortion:

$$\mathrm{D}(\widetilde{m};\widetilde{\mu},\widetilde{\mu}') \le \mathfrak{m}\left(T_{[r-3\delta,r+3\delta],x}\right) + \mathfrak{m}\left(T'_{[r-3\delta,r+3\delta],x'}\right) + 6\delta. \tag{1.3.9}$$

Furthermore, we have

$$\widetilde{m}(\widetilde{\mathscr{R}}^c) \le \mathfrak{m}\left(T_{[r,r+3\delta],x}\right) + 2\delta. \tag{1.3.10}$$

Combining (1.3.8), (1.3.9) and (1.3.10) and using the definition of the Gromov-Hausdorff-Prokhorov distance, we get

$$d_{\mathrm{GHP}}(T_{r,x}, T'_{r,x'}) \le 6\delta + \mathfrak{m}\left(T_{[r-3\delta,r+3\delta],x}\right) + \mathfrak{m}\left(T'_{[r-3\delta,r+3\delta],x'}\right) + \mathfrak{h}\left(T_{[r,r+3\delta],x}\right) + \mathfrak{h}\left(T'_{[r,r+3\delta],x'}\right). \tag{1.3.11}$$

Thanks to (1.3.7), it is straightforward to prove that

$$\mathfrak{m}\left(T'_{[r-3\delta,r+3\delta],x'}\right) \le \mathfrak{m}\left(T_{[r-6\delta,r+6\delta],x}\right) + 3\delta, \tag{1.3.12}$$

$$\mathfrak{h}\left(T'_{[r,r+3\delta],x'}\right) \le \mathfrak{h}\left(T_{[r-3\delta,r+6\delta],x}\right) - \delta. \tag{1.3.13}$$

Using (1.3.12) and (1.3.13) in conjunction with (1.3.11) yields the result. □

Proof of Proposition 1.3.3. Fix a compact real tree $T = (T,d,\emptyset,\mu)$. We will show that $\Psi_{T'} \to \Psi_T$ weakly as $T' \to T$ for d_{GHP}. Let $\varepsilon > 0$ and let $T' = (T',d',\emptyset',\mu')$ be a compact real tree such that $d_{\mathrm{GHP}}(T,T') \le \varepsilon$. Then there exist a correspondence $\mathscr{R}$ between T and T' and a measure m on $T\times T'$ such that $(\emptyset,\emptyset') \in \mathscr{R}$, $m(\mathscr{R}^c) \le \varepsilon$, $\operatorname{dis}(\mathscr{R}) \le 2\varepsilon$ and $\mathrm{D}(m;\mu,\mu') \le \varepsilon$. In particular, we will make constant use of the inequalities $|m(T\times T') - \mathfrak{m}(T)| \le \varepsilon$ and $|H(x) - H(x')| \le 2\varepsilon$ for

$(x,x') \in \mathscr{R}$. Let $f \in \mathscr{C}_b(\mathbb{T}\times\mathbb{R}_+)$ be Lipschitz. Write

$$\Psi_T(f) - \Psi_{T'}(f) = A_1 + A_2 + A_3 + A_4,$$

where

$$\begin{aligned}
A_1 &= \int_T \mu(\mathrm{d}x)\int_0^{H(x)} f(T_{r,x},r)\mathrm{d}r - \int_T m\circ p^{-1}(\mathrm{d}x)\int_0^{H(x)} f(T_{r,x},r)\mathrm{d}r\\
A_2 &= \int_{\mathscr{R}} m(\mathrm{d}x,\mathrm{d}x')\left(\int_0^{H(x)} f(T_{r,x},r)\mathrm{d}r - \int_0^{H(x')} f(T'_{r,x'},r)\mathrm{d}r\right)\\
A_3 &= \int_{\mathscr{R}^c} m(\mathrm{d}x,\mathrm{d}x')\left(\int_0^{H(x)} f(T_{r,x},r)\mathrm{d}r - \int_0^{H(x')} f(T'_{r,x'},r)\mathrm{d}r\right)\\
A_4 &= \int_{T'} m\circ p'^{-1}(\mathrm{d}x')\int_0^{H(x')} f(T'_{r,x'},r)\mathrm{d}r - \int_{T'} \mu(\mathrm{d}x')\int_0^{H(x')} f(T'_{r,x'},r)\mathrm{d}r.
\end{aligned}$$

Notice that

$$|A_1| \le 2d_{\mathrm{TV}}(m\circ p^{-1},\mu)\sup_{x\in T}\int_0^{H(x)} f(T_{r,x},r)\mathrm{d}r \le 2\mathfrak{h}(T)\,\|f\|_\infty\,\varepsilon. \tag{1.3.14}$$

Similarly, we have

$$|A_4| \le 2\mathfrak{h}(T')\,\|f\|_\infty\,\varepsilon \le 2(\mathfrak{h}(T)+2\varepsilon)\,\|f\|_\infty\,\varepsilon, \tag{1.3.15}$$

where in the second inequality we used that $\mathfrak{h}(T') \le \mathfrak{h}(T) + 2d_{\mathrm{GHP}}(T,T') \le \mathfrak{h}(T)+2\varepsilon$ by (1.2.3). Next, we have

$$|A_3| \le m(\mathscr{R}^c)(\mathfrak{h}(T)+\mathfrak{h}(T'))\,\|f\|_\infty \le 2(\mathfrak{h}(T)+\varepsilon)\,\|f\|_\infty\,\varepsilon. \tag{1.3.16}$$

We now provide a bound for A_2. We have

$$\begin{aligned}
A_2 &= \int_{\mathscr{R}} \mathbf{1}_{\{H(x)\ge H(x')\}} m(\mathrm{d}x,\mathrm{d}x')\left(\int_0^{H(x)} f(T_{r,x},r)\mathrm{d}r - \int_0^{H(x')} f(T'_{r,x'},r)\mathrm{d}r\right)\\
&\quad + \int_{\mathscr{R}} \mathbf{1}_{\{H(x)< H(x')\}} m(\mathrm{d}x,\mathrm{d}x')\left(\int_0^{H(x)} f(T_{r,x},r)\mathrm{d}r - \int_0^{H(x')} f(T'_{r,x'},r)\mathrm{d}r\right).
\end{aligned} \tag{1.3.17}$$

We only treat the first term, the second one being similar. We have

$$\begin{aligned}
&\int_{\mathscr{R}} \mathbf{1}_{\{H(x)\ge H(x')\}} m(\mathrm{d}x,\mathrm{d}x')\left(\int_0^{H(x)} f(T_{r,x},r)\mathrm{d}r - \int_0^{H(x')} f(T'_{r,x'},r)\mathrm{d}r\right)\\
&\quad = \int_{\mathscr{R}} \mathbf{1}_{\{H(x)\ge H(x')\}} m(\mathrm{d}x,\mathrm{d}x')\left(\int_0^{H(x')} \left(f(T_{r,x},r) - f(T'_{r,x'},r)\right)\mathrm{d}r + \int_{H(x')}^{H(x)} f(T_{r,x},r)\mathrm{d}r\right).
\end{aligned}$$

On the one hand, we get

$$\begin{aligned}
\left|\int_{\mathscr{R}} \mathbf{1}_{\{H(x)\ge H(x')\}} m(\mathrm{d}x,\mathrm{d}x')\int_{H(x')}^{H(x)} f(T_{r,x},r)\mathrm{d}r\right| &\le \int_{\mathscr{R}} \|f\|_\infty\,|H(x)-H(x')|\,m(\mathrm{d}x,\mathrm{d}x')\\
&\le \|f\|_\infty\, m(T\times T')\,\mathrm{dis}(\mathscr{R})
\end{aligned}$$

$$\leq 2\left\|f\right\|_{\infty}(\mathfrak{m}(T)+\varepsilon)\varepsilon. \tag{1.3.18}$$

On the other hand, we have

$$\begin{aligned}
&\left|\int_{\mathscr{R}} \mathbf{1}_{\{H(x)\geq H(x')\}} m(\mathrm{d}x,\mathrm{d}x') \int_0^{H(x')} \left(f(T_{r,x},r)-f(T'_{r,x},r)\right)\mathrm{d}r\right| \\
&\quad\leq \left\|f\right\|_{\mathrm{L}} \int_{\mathscr{R}} \mathbf{1}_{\{H(x)\geq H(x')\}} m(\mathrm{d}x,\mathrm{d}x') \int_0^{H(x')} d_{\mathrm{GHP}}\left(T_{r,x},T'_{r,x}\right)\mathbf{1}_{\{r\geq 6\varepsilon\}}\mathrm{d}r \\
&\qquad\qquad\qquad\qquad + \int_{\mathscr{R}} m(\mathrm{d}x,\mathrm{d}x') \int_0^{6\varepsilon} 2\left\|f\right\|_{\infty}\mathrm{d}r \\
&\quad\leq 2\left\|f\right\|_{\mathrm{L}} \int m(\mathrm{d}x,\mathrm{d}x') \int_0^{H(x)} \left(\mathfrak{m}(T_{[r-6\varepsilon,r+6\varepsilon],x})+\mathfrak{h}(T_{[r-3\varepsilon,r+6\varepsilon],x})\right)\mathbf{1}_{\{r\geq 6\varepsilon\}}\,\mathrm{d}r \\
&\qquad\qquad + 8\left\|f\right\|_{\mathrm{L}}\mathfrak{h}(T)(\mathfrak{m}(T)+\varepsilon)\varepsilon + 12\left\|f\right\|_{\infty}(\mathfrak{m}(T)+\varepsilon)\varepsilon.
\end{aligned} \tag{1.3.19}$$

where we used (1.3.5) for the last inequality. Using Fubini's theorem, we get

$$\begin{aligned}
&\int m(\mathrm{d}x,\mathrm{d}x') \int_0^{H(x)} \mathfrak{m}(T_{[r-6\varepsilon,r+6\varepsilon],x})\mathbf{1}_{\{r\geq 6\varepsilon\}}\,\mathrm{d}r \\
&\qquad = \int m(\mathrm{d}x,\mathrm{d}x') \int_0^{H(x)} \mu(t\colon\ H(t\wedge x)\in[r-6\varepsilon,r+6\varepsilon])\mathbf{1}_{\{r\geq 6\varepsilon\}}\,\mathrm{d}r \\
&\qquad = \int m(\mathrm{d}x,\mathrm{d}x') \int_T \mu(\mathrm{d}t) \int_0^{H(x)} \mathbf{1}_{\{H(t\wedge x)\in[r-6\varepsilon,r+6\varepsilon]\}}\mathbf{1}_{\{r\geq 6\varepsilon\}}\,\mathrm{d}r \\
&\qquad \leq 12\,\mathfrak{m}(T)(\mathfrak{m}(T)+\varepsilon)\varepsilon.
\end{aligned} \tag{1.3.20}$$

Moreover, since T is compact, it holds that for every $x\in T$ and every $\delta>0$, there is a finite number of subtrees with height larger than δ attached to the branch $[\![\emptyset,x]\!]$. Let $r\in(0,H(x))$. Recall that x_r is the unique ancestor of x with height $H(x_r)=r$. Assume that x_r is not a branching point. Then, for every $\delta>0$ and for $\varepsilon>0$ small enough (depending on δ), there are no subtrees with height larger than δ attached to $[\![x_{r-3\varepsilon},x_{r+6\varepsilon}]\!]$. (To be precise, if $y\in[\![x_{r-3\varepsilon},x_{r+6\varepsilon}]\!]$ is a branching point, the tree attached at y is $T_{[s,s],x}$ with $s=H(y)$). Therefore, we have $\mathfrak{h}(T_{[r-3\varepsilon,r+6\varepsilon],x})\leq\delta+9\varepsilon$. This proves that, for every $r\in(0,H(x))$ such that x_r is not a branching point,

$$\lim_{\varepsilon\to 0}\mathfrak{h}(T_{[r-3\varepsilon,r+6\varepsilon],x})=0. \tag{1.3.21}$$

But since T is compact, there are (at most) countably many $r\in(0,H(x))$ such that x_r is a branching point. It follows that (1.3.21) holds for every $x\in T$ and $\mathrm{d}r$-a.e. $r\in[0,H(x)]$. Notice that $\mathfrak{h}(T_{r-3\varepsilon,r+6\varepsilon,x})\leq\mathfrak{h}(T)$ and the measure $\mathbf{1}_{\{0\leq r\leq H(x)\}}\,\mu(\mathrm{d}x)\mathrm{d}r$ is finite as its total mass is less than $\mathfrak{h}(T)\mathfrak{m}(T)$ which is finite. We get by the dominated convergence theorem that

$$\lim_{\varepsilon\to 0}\int_T \mu(\mathrm{d}x)\int_0^{H(x)} \mathfrak{h}(T_{[r-3\varepsilon,r+6\varepsilon],x})\mathbf{1}_{\{r\geq 6\varepsilon\}}\,\mathrm{d}r=0.$$

Since

$$\left|\int_T \left(m\circ p^{-1}(\mathrm{d}x)-\mu(\mathrm{d}x)\right)\int_0^{H(x)} \mathfrak{h}(T_{[r-3\varepsilon,r+6\varepsilon],x})\mathbf{1}_{\{r\geq 6\varepsilon\}}\,\mathrm{d}r\right| \leq 2\mathfrak{h}(T)^2 d_{\mathrm{TV}}(m\circ p^{-1},\mu)\leq 2\mathfrak{h}(T)^2\varepsilon,$$

it follows that

$$\lim_{\varepsilon\to 0}\int m(\mathrm{d}x,\mathrm{d}x')\int_0^{H(x)} \mathfrak{h}(T_{[r-3\varepsilon,r+6\varepsilon],x})\mathbf{1}_{\{r\geq 6\varepsilon\}}\,\mathrm{d}r = 0. \tag{1.3.22}$$

Thus, by (1.3.14)–(1.3.16), (1.3.18)–(1.3.20) and (1.3.22), we deduce that

$$\lim_{\varepsilon\to 0}\sup_{d_{\mathrm{GHP}}(T,T')<\varepsilon}\Psi_{T'}(f)=\Psi_T(f)$$

for every Lipschitz function $f\in\mathscr{C}_b(\mathbb{T}\times\mathbb{R}_+)$. This proves that $\Psi\colon\mathbb{T}\to\mathscr{M}(\mathbb{T}\times\mathbb{R}_+)$ is continuous which concludes the proof. $\square$

1.4 Bienaymé-Galton-Watson trees and stable Lévy trees

Throughout this work, we fix a random variable ξ whose distribution is critical and belongs to the domain of attraction of a stable distribution with index $\gamma\in(1,2]$. More precisely, we assume that ξ takes values in $\mathbb{N}=\{0,1,2,\dots\}$ and that it satisfies the following conditions:

(ξ1) ξ is critical, i.e. $\mathbb{E}[\xi]=1$,

(ξ2) ξ belongs to the domain of attraction of a stable distribution with index $\gamma\in(1,2]$, i.e. $\mathbb{E}\left[\xi^2\mathbf{1}_{\{\xi\leq n\}}\right]=n^{2-\gamma}L(n)$, where $L\colon\mathbb{R}_+\to\mathbb{R}_+$ is a slowly varying function.

By [72, Theorem XVII.5.2] or [93, Theorem 5.2], assumption (ξ2) is equivalent to the existence of a positive sequence $(b_n,\, n\geq 1)$ such that, if $(\xi_n,\, n\geq 1)$ is a sequence of independent random variables with the same distribution as ξ, then

$$\frac{1}{b_n}\left(\sum_{k=1}^n \xi_k - n\right)\xrightarrow[n\to\infty]{(d)} X_1, \tag{1.4.1}$$

where $(X_t,\, t\geq 0)$ is a strictly stable spectrally positive Lévy process with Laplace transform $\mathbb{E}\left[\exp(-\lambda X_t)\right]=\exp(t\kappa\lambda^\gamma)$ where $\gamma\in(1,2]$ and $\kappa>0$. Note that we have automatically $b_n/n\to 0$ as $n\to\infty$. In most of our results, we make the following stronger assumption on ξ:

(ξ2)$'$ $\mathbb{E}\left[\xi^2\mathbf{1}_{\{\xi\leq n\}}\right]=n^{2-\gamma}L(n)$ where $L\colon\mathbb{R}_+\to\mathbb{R}_+$ is a slowly varying function which is bounded away from zero and infinity.

Assumption (ξ2)$'$ is equivalent to the normalizing sequence $(b_n,\, n\geq 1)$ which appears in (1.4.1) satisfying

$$\underline{b}n^{1/\gamma}\leq b_n\leq \overline{b}n^{1/\gamma},\quad \forall n\geq 1, \tag{1.4.2}$$

for some constants $0 < \underline{b} < \overline{b} < \infty$. Indeed, if $\gamma = 2$, we have the convergence of $nb_n^{-2}L(b_n)$ to some positive constant by [93, Theorem 5.2 and Eq. (5.44)]. Similarly, if $\gamma \in (1,2)$, using [93, Theorem 5.3 and Eq. (5.7)], we have as $n \to \infty$ that

$$n\mathbb{P}(\xi > b_n) \sim \frac{2-\gamma}{\gamma} nb_n^{-\gamma} L(b_n).$$

On the other hand, [93, Eq. (5.10)] entails the convergence of $n\mathbb{P}(\xi > b_n)$ to some positive constant. Therefore, for $\gamma \in (1,2]$, the sequence $n^{1/\gamma} b_n^{-1} L(b_n)^{1/\gamma}$ converges to some positive constant. Thus, if L is bounded away from 0 and infinity, then (1.4.2) follows. The proof of the converse (which we shall not use) is left for the reader.

1.4.1 Results on conditioned Bienaymé-Galton-Watson trees

Recall that the span of the integer-valued random variable ξ is the largest integer λ_0 such that a.s. $\xi \in a + \lambda_0\mathbb{Z}$ for some $a \in \mathbb{Z}$. As we only consider ξ with $\mathbb{P}(\xi = 0) > 0$, the span is the largest integer λ_0 such that a.s. $\xi \in \lambda_0\mathbb{Z}$, i.e. the greatest common divisor of $\{k \geq 1 : \mathbb{P}(\xi = k) > 0\}$.

Assume that ξ satisfies (ξ1) and (ξ2) and denote by $\mathfrak{g}$ the density of the random variable X_1 appearing in (1.4.1). Then the function $\mathfrak{g}$ is continuous on $\mathbb{R}$ (in fact infinitely differentiable) and satisfies

$$\mathfrak{g}(0) = \frac{1}{\kappa^{1/\gamma} \left|\Gamma(-1/\gamma)\right|}, \tag{1.4.3}$$

where Γ is Euler's gamma function, see [72, Lemma XVII.6.1] or [93, Example 3.15 and Eq. (4.6)]. In particular, when $\gamma = 2$, $\mathfrak{g}$ is the density of a centered Gaussian distribution with variance 2κ and we have

$$\mathfrak{g}(0) = \frac{1}{2\sqrt{\kappa\pi}}. \tag{1.4.4}$$

Recall that $(\xi_n, n \geq 1)$ is a sequence of independent random variables with the same distribution as ξ and define $S_n = \sum_{k=1}^n \xi_k$. The following result is a direct consequence of the local limit theorem, see e.g. [90, Chapter 4, Theorem 4.2.1].

Lemma 1.4.1 (Local limit theorem)**.** *Assume that ξ satisfies (ξ1) and (ξ2) and denote its span by λ_0. We have*

$$\lim_{n\to\infty} \sup_{k\geq 0} \left| \frac{b_n}{\lambda_0} \mathbb{P}(S_n = \lambda_0 k) - \mathfrak{g}\left(\frac{\lambda_0 k - n}{b_n}\right)\right| = 0,$$

where $\mathfrak{g}$ is the density of the random variable X_1 defined in (1.4.1)*. In particular, for any fixed $k \geq 0$, we have as $n \to \infty$ with $n \equiv k \pmod{\lambda_0}$,*

$$\mathbb{P}(S_n = n - k) \sim \frac{\lambda_0 \mathfrak{g}(0)}{b_n}. \tag{1.4.5}$$

Let τ be a BGW(ξ) tree, see e.g. Athreya and Ney [23]. By the well-known Otter-Dwass formula, we have, for every $n \geq 1$,

$$\mathbb{P}(|\tau| = n) = \frac{1}{n}\mathbb{P}(S_n = n-1). \tag{1.4.6}$$

In particular, we get $\mathbb{P}(|\tau| = n) = 0$ if $n \not\equiv 1 \pmod{\lambda_0}$ while $\mathbb{P}(|\tau| = n) > 0$ for all large n with $n \equiv 1 \pmod{\lambda_0}$ by Lemma 1.4.1. We denote by Δ the support of the random variable $|\tau|$ when τ is not reduced to the root, that is

$$\Delta = \{n \geq 2 : \ \mathbb{P}(|\tau| = n) > 0\}. \tag{1.4.7}$$

In particular, the previous discussion implies that $\Delta \subset 1 + \lambda_0\mathbb{N}$ and conversely, $1 + \lambda_0 n \in \Delta$ for all large n. In what follows, we only consider $n \in \Delta$ and convergences should be understood along the set Δ.

We will also need the following sub-exponential tail bounds for the height of conditioned BGW trees, see [112, Theorem 2] and the discussion thereafter. For every $n \in \Delta$, τ^n will denote a BGW(ξ) tree conditioned to have n vertices, that is τ^n is distributed as τ conditionally on $\{|\tau| = n\}$.

Lemma 1.4.2. *Assume that ξ satisfies (ξ1) and (ξ2). For every $\alpha \in (0, \gamma/(\gamma-1))$ and every $\beta \in (0,\gamma)$, there exist two finite constants $C_0, c_0 > 0$ such that for every $y \geq 0$ and $n \in \Delta$, we have*

$$\mathbb{P}\left(\frac{b_n}{n}\mathfrak{h}(\tau^n) \leq y\right) \leq C_0 \exp\left(-c_0 y^{-\alpha}\right), \tag{1.4.8}$$

$$\mathbb{P}\left(\frac{b_n}{n}\mathfrak{h}(\tau^n) \geq y\right) \leq C_0 \exp\left(-c_0 y^{\beta}\right). \tag{1.4.9}$$

Remark 1.4.3.

(i) If moreover ξ satisfies $(\xi 2)'$, then we can take $\alpha = \gamma/(\gamma-1)$ in (1.4.8), see Appendix 1.B.

(ii) If ξ has finite variance $\sigma_\xi^2 \in (0,\infty)$ (in which case $(\xi 2)'$ is satisfied), we have $\gamma = 2$ and we can take $b_n = \sigma_\xi\sqrt{n}$ in (1.4.1) with $\kappa = 1/2$ (this is just the central limit theorem). Then both (1.4.8) and (1.4.9) hold with $\alpha = \beta = 2$, see [14, Theorem 1.1 and Theorem 1.2].

An immediate consequence of Lemma 1.4.2 is the following estimate for the moments of $\mathfrak{h}(\tau^n)$ which extends [14, Corollary 1.3].

Lemma 1.4.4. *Assume that ξ satisfies (ξ1) and (ξ2). For every $p \in \mathbb{R}$, we have*

$$\sup_{n\in\Delta} \mathbb{E}\left[\left(\frac{b_n}{n}\mathfrak{h}(\tau^n)\right)^p\right] < \infty.$$

Proof. Let $p > 0$. Fix $\beta \in (0,\gamma)$. By Lemma 1.4.2, we have for every $n \in \Delta$

$$\mathbb{E}\left[\left(\frac{b_n}{n}\mathfrak{h}(\tau^n)\right)^p\right] = p\int_0^\infty y^{p-1}\,\mathbb{P}\left(\frac{b_n}{n}\mathfrak{h}(\tau^n) > y\right)\mathrm{d}y \leq C_0 p \int_0^\infty y^{p-1}\mathrm{e}^{-c_0 y^\beta}\,\mathrm{d}y < \infty.$$

Similarly, fix $\alpha \in (0, \gamma/(\gamma-1))$ and apply Lemma 1.4.2 to get

$$\mathbb{E}\left[\left(\frac{b_n}{n}\mathfrak{h}(\tau^n)\right)^{-p}\right] = p\int_0^\infty y^{p-1}\,\mathbb{P}\left(\frac{b_n}{n}\mathfrak{h}(\tau^n) < \frac{1}{y}\right)\mathrm{d}y \le C_0 p\int_0^\infty y^{p-1}e^{-c_0 y^\alpha}\,\mathrm{d}y < \infty.$$

This proves the result. □

We end this section with the following lemma used in the proof of Remark 1.1.2-(vi).

Lemma 1.4.5. *Assume that ξ has finite variance $\sigma_\xi^2 \in (0,\infty)$. Let $\alpha', \beta \in \mathbb{R}$ such that $2\alpha' + \beta < 0$ and set $f_{\alpha',\beta}(\mathbf{t}) = |\mathbf{t}|^{\alpha'}\,\mathfrak{h}(\mathbf{t})^\beta \mathbf{1}_{\{|\mathbf{t}|>1\}}$. Then we have*

$$\mathbb{E}\left[f_{\alpha',\beta}(\tau)\right] < \infty, \quad \lim_{n\to\infty}\mathbb{E}\left[f_{\alpha',\beta}(\tau^n)^2\right] = 0 \quad \textit{and} \quad \sum_{n\in\Delta}\frac{\sqrt{\mathbb{E}\left[f_{\alpha',\beta}(\tau^n)^2\right]}}{n} < \infty.$$

Proof. Recall from (1.4.7) the definition of Δ. We have

$$\mathbb{E}\left[f_{\alpha',\beta}(\tau)\right] = \sum_{n\in\Delta} n^{\alpha'}\,\mathbb{E}\left[\mathfrak{h}(\tau^n)^\beta\right]\mathbb{P}\left(|\tau| = n\right).$$

Using (1.4.6) and (1.4.5), (1.4.4) with $b_n = \sigma_\xi\sqrt{n}$ and $\kappa = 1/2$, we have as $n \to \infty$ that

$$\mathbb{P}\left(|\tau| = n\right) \sim \frac{\lambda_0}{\sqrt{2\pi\sigma_\xi^2}}n^{-3/2}.$$

Since $\mathbb{E}\left[\mathfrak{h}(\tau^n)^\beta\right] = O(n^{\beta/2})$ as $n \to \infty$ by Lemma 1.4.4, we get that

$$\mathbb{E}\left[f_{\alpha',\beta}(\tau)\right] \le C\sum_{n\in\Delta} n^{-3/2+\alpha'+\beta/2} < \infty.$$

Applying Lemma 1.4.4 again gives $\mathbb{E}\left[f_{\alpha',\beta}(\tau^n)^2\right] = n^{2\alpha'}\,\mathbb{E}\left[\mathfrak{h}(\tau^n)^{2\beta}\right]\mathbf{1}_{\{n>1\}} \le Mn^{2\alpha'+\beta}$ for some finite constant $M > 0$, and the last term converges to 0 as $n \to \infty$. Finally, we have

$$\sum_{n\in\Delta}\frac{\sqrt{\mathbb{E}\left[f_{\alpha',\beta}(\tau^n)^2\right]}}{n} \le \sqrt{M}\sum_{n\in\Delta} n^{-1+\alpha'+\beta/2} < \infty.$$

□

1.4.2 Stable Lévy trees

Let us briefly recall the definition of the height process and the associated Lévy tree, see e.g. [119, 57, 54, 110]. Recall that $(X_t,\, t \ge 0)$ is a strictly stable Lévy process with Laplace exponent $\psi(\lambda) = \kappa\lambda^\gamma$ where $\gamma \in (1,2]$ and $\kappa > 0$. For $\gamma \in (1,2)$, denote by π the associated Lévy measure

$$\pi(\mathrm{d}x) = \frac{\kappa\gamma(\gamma-1)}{\Gamma(2-\gamma)}\frac{\mathrm{d}x}{x^{1+\gamma}}. \tag{1.4.10}$$

Le Gall and Le Jan [119] proved that there exists a continuous process $(H(t),\, t \geq 0)$ called the ψ-height process such that for every $t \geq 0$, we have the following convergence in probability

$$H(t) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_0^t \mathbf{1}_{\{X_s < I_t^s + \varepsilon\}} \, \mathrm{d}s,$$

where $I_t^s = \inf_{[s,t]} X$. In the Brownian case, H is a (scaled) reflected Brownian motion. Let $\mathbf{N}$ be the excursion measure of H above 0 and set

$$\sigma = \inf\{s > 0 \colon\ H(s) = 0\} \quad \text{and} \quad \mathfrak{h} = \sup_{s \geq 0} H(s) \tag{1.4.11}$$

for the duration of the excursion and its maximum. We choose to normalize the excursion measure $\mathbf{N}$ such that the distribution of σ under $\mathbf{N}$ is π_* given by

$$\pi_*(\mathrm{d}x) = \mathbf{N}[\sigma \in \mathrm{d}x] = \mathfrak{g}(0) \frac{\mathrm{d}x}{x^{1+1/\gamma}}, \tag{1.4.12}$$

with $\mathfrak{g}(0)$ given in (1.4.3). Furthermore, by [58, Eq. (14)], the distribution of $\mathfrak{h}$ under $\mathbf{N}$ is given by

$$\mathbf{N}[\mathfrak{h} > x] = \big(\kappa(\gamma - 1)x\big)^{-1/(\gamma-1)}. \tag{1.4.13}$$

We have the following equality in "distribution" for the height process, see e.g. [60, Eq. (40)],

$$(H(xt),\, t \geq 0) \quad \text{under } x^{1/\gamma}\mathbf{N} \quad \overset{(d)}{=} \quad x^{1-1/\gamma} H \quad \text{under } \mathbf{N}.$$

Using this, one can make sense of the conditional probability measure $\mathbf{N}^{(x)}[\bullet] = \mathbf{N}[\bullet | \sigma = x]$ such that $\mathbf{N}^{(x)}$-a.s., $\sigma = x$ and

$$\mathbf{N}[\bullet] = \int_0^\infty \mathbf{N}^{(x)}[\bullet]\, \pi_*(\mathrm{d}x).$$

Informally, $\mathbf{N}^{(x)}$ can be seen as the distribution of the excursion of H with duration x. Moreover, the height process H has the following scaling property

$$(H(s),\, s \in [0,x]) \text{ under } \mathbf{N}^{(x)} \quad \overset{(d)}{=} \quad \big(x^{1-1/\gamma} H(s/x),\, s \in [0,x]\big) \text{ under } \mathbf{N}^{(1)}. \tag{1.4.14}$$

See also Lemma 1.6.11 for the scaling property of H and related processes.

We call the stable Lévy tree with branching mechanism $\psi(\lambda) = \kappa\lambda^\gamma$, the compact real tree $\mathscr{T}$ coded by the ψ-height process H under $\mathbf{N}^{(1)}$. See Remark 1.2.1 for the coding of real trees by excursion paths. Thanks to the Ray-Knight theorem, see [57, Theorem 1.4.1], the stable Lévy tree codes the genealogy of the stable continuous-state branching process.

Remark 1.4.6. Notice that $\sigma = \mathfrak{m}(\mathscr{T}_H)$ and $\mathfrak{h} = \mathfrak{h}(\mathscr{T}_H)$ are the mass and the height of the tree $\mathscr{T}_H$ coded by the height process H under $\mathbf{N}$. Furthermore, for $s \in [0,\sigma]$, the notation $H(s)$ is consistent with the one introduced in Section 1.2.4 since $H(s)$ is the height of s in the tree coded by H under $\mathbf{N}$.

Remark 1.4.7. In the Brownian case $\psi(\lambda) = \kappa\lambda^2$, the ψ-height process H is distributed under $\mathbf{N}^{(1)}$ as $\sqrt{2/\kappa}\,\mathbf{e}$ where $\mathbf{e}$ is the normalized Brownian excursion. In particular, $\kappa = 1/2$ corresponds to Aldous' normalization of the Brownian tree [18, Corollary 22], while $\kappa = 2$ corresponds to Le Gall's [116, Definition 2.2].

1.4.3 Convergence of continuous functionals

For every $n \in \Delta$, we let τ^n be a BGW(ξ) tree conditioned to have n vertices, and let $\mathscr{T}^n = (b_n/n)\tau^n$ be the associated real tree rescaled so that all edges have length b_n/n and equipped with the uniform probability measure on the set of vertices whose heights are integer multiples of b_n/n. Duquesne [54] (see also [110]) showed that the convergence in distribution

$$\mathscr{T}^n \xrightarrow[n\to\infty]{(d)} \mathscr{T} \tag{1.4.15}$$

holds in the space $\mathbb{T}$ where $\mathscr{T}$ is the stable Lévy tree with branching mechanism $\psi(\lambda) = \kappa\lambda^\gamma$.

The following result is an immediate consequence of Proposition 1.3.3. Recall from (1.3.1) and (1.3.2) the definitions of the measures Ψ_T and $\Psi_T^{\mathfrak{mh}}$.

Corollary 1.4.8. *Assume that ξ satisfies (ξ1) and (ξ2). Let τ^n be a BGW(ξ) tree conditioned to have n vertices and let $\mathscr{T}^n = (b_n/n)\tau^n$ be the associated real tree rescaled so that all edges have length b_n/n (where b_n is the normalizing sequence in* (1.4.1)*). Then we have the convergence in distribution $\Psi_{\mathscr{T}^n} \xrightarrow{(d)} \Psi_{\mathscr{T}}$ in $\mathscr{M}(\mathbb{T}\times\mathbb{R}_+)$, where $\mathscr{T}$ is the stable Lévy tree with branching mechanism $\psi(\lambda) = \kappa\lambda^\gamma$. In particular, we have $\Psi^{\mathfrak{mh}}_{\mathscr{T}^n} \xrightarrow{(d)} \Psi^{\mathfrak{mh}}_{\mathscr{T}}$ in $\mathscr{M}(\mathbb{R}_+^2)$.*

The convergence in distribution obtained in Corollary 1.4.8 is unsatisfactory to study the asymptotics of additive functionals of large BGW trees as it involves the real tree $\mathscr{T}^n$ instead of the (discrete) BGW tree τ^n. To remedy this, we shall introduce a discrete version of the measure Ψ_T when T is associated with a discrete tree. Let $\mathbf{t}$ be a discrete tree and $a > 0$. Recall that $a\mathbf{t}$ denotes the real tree associated to $\mathbf{t}$ where the branches have length a, and that for $v \in \mathbf{t}$, av denotes the corresponding vertex in $a\mathbf{t}$, see Section 1.2.4 for the definitions. We define two nonnegative measures $\mathscr{A}^\circ_{\mathbf{t},a}$ and $\mathscr{A}_{\mathbf{t},a}$ on $\mathbb{T}\times\mathbb{R}_+$ as follows: for every $f \in \mathscr{B}_+(\mathbb{T}\times\mathbb{R}_+)$,

$$\boxed{\mathscr{A}^\circ_{\mathbf{t},a}(f) = \frac{a}{|\mathbf{t}|}\sum_{w\in\mathbf{t}^\circ} |\mathbf{t}_w|\, f(a\mathbf{t}_w, aH(w))} \quad \text{and} \quad \boxed{\mathscr{A}_{\mathbf{t},a}(f) = \frac{a}{|\mathbf{t}|}\sum_{w\in\mathbf{t}} |\mathbf{t}_w|\, f(a\mathbf{t}_w, aH(w))}, \tag{1.4.16}$$

where $a\mathbf{t}_w$ is the subtree of $a\mathbf{t}$ above aw. Note that the sum is over all internal vertices of $\mathbf{t}$ for $\mathscr{A}^\circ_{\mathbf{t},a}$, while for $\mathscr{A}_{\mathbf{t},a}$ the sum extends over all vertices including the leaves. In other words, the measure $\mathscr{A}^\circ_{\mathbf{t},a}$ ignores the subtrees rooted at a leaf of $\mathbf{t}$ (which are trivial trees consisting only of a root equipped with a scaled Dirac measure). Let us take a moment to explain why we introduce the measure $\mathscr{A}^\circ_{\mathbf{t},a}$. While $\mathscr{A}_{\mathbf{t},a}$ seems more natural, the measure $\mathscr{A}^\circ_{\mathbf{t},a}$ has the advantage of putting no mass on the set

$$\mathbb{T}_0 \times \mathbb{R}_+ = \left\{T \in \mathbb{T}\colon\ \mathfrak{m}(T) = 0 \text{ or } \mathfrak{h}(T) = 0\right\} \times \mathbb{R}_+.$$

This will be useful as we are interested in sums of the form (1.4.16) where the function f may blow up on $\mathbb{T}_0 \times \mathbb{R}_+$. We now give estimates for the distances between the three measures $\mathscr{A}^\circ_{\mathbf{t},a}$, $\mathscr{A}_{\mathbf{t},a}$ and $\Psi_{a\mathbf{t}}$, on $\mathbb{T} \times \mathbb{R}_+$, which are associated with the discrete tree $\mathbf{t}$ and $a > 0$.

Lemma 1.4.9. *Let* $\mathbf{t}$ *be a discrete tree and let* $a > 0$. *We have*

$$d_{\mathrm{BL}}\left(\Psi_{a\mathbf{t}}, \mathscr{A}_{\mathbf{t},a}\right) \le a\left(\frac{3}{4}\mathscr{A}_{\mathbf{t},a}(1) + 1\right), \tag{1.4.17}$$

$$d_{\mathrm{TV}}(\mathscr{A}_{\mathbf{t},a}, \mathscr{A}^\circ_{\mathbf{t},a}) \le a. \tag{1.4.18}$$

Proof. Let $f \in \mathscr{C}_b(\mathbb{T} \times \mathbb{R}_+)$ be Lipschitz. Recall that $T = a\mathbf{t}$ is the real tree associated with $\mathbf{t}$, rescaled so that all edges have length a and equipped with the uniform probability measure on the set of vertices whose height is an integer multiple of a. Recall also that for $v \in \mathbf{t}$, av denotes the corresponding vertex in $T = a\mathbf{t}$. In particular, $H(av) = aH(v)$, where $H(av)$ is the height of av in the real tree $a\mathbf{t}$ and $H(v)$ is the height of v in the discrete tree $\mathbf{t}$. Thus, we have

$$\begin{aligned}\Psi_T(f) = \frac{1}{|\mathbf{t}|}\sum_{v\in\mathbf{t}}\int_0^{H(av)} f(T_{r,av}, r)\,\mathrm{d}r &= \frac{1}{|\mathbf{t}|}\sum_{v\in\mathbf{t}}\int_0^{aH(v)} f(T_{r,av}, r)\,\mathrm{d}r \\ &= \frac{a}{|\mathbf{t}|}\sum_{v\in\mathbf{t}}\sum_{k=1}^{H(v)}\int_{k-1}^{k} f\left(T_{ar,av}, ar\right)\mathrm{d}r.\end{aligned}$$

On the other hand, note that for every $1 \le k \le H(v)$, we have $T_{ak,av} = T_{aw}$ where $w \in \mathbf{t}$ is the unique ancestor of v with height k. Thus, we have

$$\sum_{v\in\mathbf{t}}\sum_{k=1}^{H(v)} f\left(T_{ak,av}, ak\right) = \sum_{v\in\mathbf{t}}\sum_{\substack{w\preccurlyeq v\\ w\neq\emptyset}} f\left(T_{aw}, aH(w)\right) = \sum_{w\neq\emptyset}|\mathbf{t}_w| f\left(T_{aw}, aH(w)\right) = \frac{|\mathbf{t}|}{a}\mathscr{A}_{\mathbf{t},a}(f) - |\mathbf{t}| f(T,0).$$

Therefore, we deduce that

$$\begin{aligned}\left|\Psi_T(f) - \mathscr{A}_{\mathbf{t},a}(f)\right| &\le \frac{a}{|\mathbf{t}|}\sum_{v\in\mathbf{t}}\sum_{k=1}^{H(v)}\int_{k-1}^{k}\left|f\left(T_{ar,av}, ar\right) - f\left(T_{ak,av}, ak\right)\right|\mathrm{d}r + a\|f\|_\infty \\ &\le \frac{a}{|\mathbf{t}|}\sum_{v\in\mathbf{t}}\sum_{k=1}^{H(v)}\int_{k-1}^{k}\|f\|_{\mathrm{L}}\left(d_{\mathrm{GHP}}\left(T_{ar,av}, T_{ak,av}\right) + a(k-r)\right)\mathrm{d}r + a\|f\|_\infty.\end{aligned} \tag{1.4.19}$$

Since for $k-1 < r \le k$, the tree $T_{ar,av}$ is obtained by grafting $T_{ak,av}$ on top of a branch of height $a(k-r)$ and no mass, it is straightforward to check that $d_{\mathrm{GHP}}\left(T_{ar,av}, T_{ak,av}\right) \le a(k-r)/2$. It follows that

$$\left|\Psi_T(f) - \mathscr{A}_{\mathbf{t},a}(f)\right| \le \frac{a}{|\mathbf{t}|}\sum_{v\in\mathbf{t}}\sum_{k=1}^{H(v)}\frac{3a}{4}\|f\|_{\mathrm{L}} + a\|f\|_\infty \le \frac{3a}{4}\|f\|_{\mathrm{L}}\mathscr{A}_{\mathbf{t},a}(1) + a\|f\|_\infty.$$

By definition of the distance d_{BL}, we deduce that

$$d_{\mathrm{BL}}\left(\Psi_T, \mathscr{A}_{\mathbf{t},a}\right) \le a\left(\frac{3}{4}\mathscr{A}_{\mathbf{t},a}(1) + 1\right).$$

Next, let $f \in \mathscr{B}_b(\mathbb{T} \times \mathbb{R}_+)$. We have

$$\left|\mathscr{A}_{\mathbf{t},a}(f) - \mathscr{A}^{\circ}_{\mathbf{t},a}(f)\right| = \frac{a}{|\mathbf{t}|}\left|\sum_{w\in \mathrm{Lf}(\mathbf{t})} |\mathbf{t}_w| f\left(T_{aw}, aH(w)\right)\right| \leq \frac{a}{|\mathbf{t}|}\,|\mathrm{Lf}(\mathbf{t})|\,\|f\|_\infty \leq a\,\|f\|_\infty.$$

Taking the supremum over all $f \in \mathscr{B}_b(\mathbb{T}\times\mathbb{R}_+)$ such that $\|f\|_\infty \leq 1$ yields $d_{\mathrm{TV}}\left(\mathscr{A}_{\mathbf{t},a}, \mathscr{A}^{\circ}_{\mathbf{t},a}\right) \leq a$. $\square$

We now restate the convergence of Corollary 1.4.8 in terms of the discrete trees τ^n. To avoid cumbersome notations, we write

$$\boxed{\mathscr{A}^{\circ}_n = \mathscr{A}^{\circ}_{\tau^n, b_n/n}} \quad \text{and} \quad \boxed{\mathscr{A}_n = \mathscr{A}_{\tau^n, b_n/n}}.$$

Recall that for a discrete tree $\mathbf{t}$, $w \in \mathbf{t}$ and $a > 0$, we have that $\mathfrak{h}(a\mathbf{t}_w) = a\mathfrak{h}(\mathbf{t}_w)$ and $\mathfrak{m}(a\mathbf{t}_w) = |\mathbf{t}_w|/|\mathbf{t}|$. We shall also consider the following variant of the measure $\mathscr{A}^{\circ}_n$ for functions depending only on the mass and height: for every measurable function f belonging to $\mathscr{B}_+([0,1]\times\mathbb{R}_+)$,

$$\boxed{\mathscr{A}^{\mathfrak{m}\mathfrak{h},\circ}_n(f) = \frac{b_n}{n^2}\sum_{w\in\tau^{n,\circ}} |\tau^n_w|\, f\left(\frac{|\tau^n_w|}{n}, \frac{b_n}{n}\mathfrak{h}(\tau^n_w)\right)}. \tag{1.4.20}$$

We have the following upper bound of their total mass.

Lemma 1.4.10. *We have:*

$$\mathscr{A}^{\circ}_n(1) \leq \frac{b_n}{n}\mathfrak{h}(\tau^n) \quad \textit{and} \quad \mathscr{A}_n(1) \leq \frac{b_n}{n}\left(\mathfrak{h}(\tau^n)+1\right). \tag{1.4.21}$$

Proof. The proof is elementary as

$$\mathscr{A}^{\circ}_n(1) = \frac{b_n}{n^2}\sum_{w\in\tau^{n,\circ}} |\tau^n_w| = \frac{b_n}{n^2}\sum_{w\in\tau^{n,\circ}}\sum_{w\preccurlyeq v} 1 \leq \frac{b_n}{n^2}\sum_{v\in\tau^n}\mathfrak{h}(\tau^n) \leq \frac{b_n}{n}\mathfrak{h}(\tau^n),$$

$$\mathscr{A}_n(1) = \frac{b_n}{n^2}\sum_{w\in\tau^n} |\tau^n_w| = \mathscr{A}^{\circ}_n(1) + \frac{b_n}{n^2}|\mathrm{Lf}(\mathbf{t})| \leq \frac{b_n}{n}\left(\mathfrak{h}(\tau^n)+1\right).$$

$\square$

We have the following convergence of $\mathscr{A}^{\circ}_n$ as n goes to infinity.

Corollary 1.4.11. *Assume that ξ satisfies (ξ1) and (ξ2) and let τ^n be a BGW(ξ) tree conditioned to have n vertices. Then for every $f \in \mathscr{C}_b(\mathbb{T}\times\mathbb{R}_+)$, we have the convergence in distribution and of all positive moments*

$$\mathscr{A}^{\circ}_n(f) = \frac{b_n}{n^2}\sum_{w\in\tau^{n,\circ}} |\tau^n_w|\, f\left(\frac{b_n}{n}\tau^n_w, \frac{b_n}{n}H(w)\right) \xrightarrow[n\to\infty]{(d)+\text{moments}} \Psi_{\mathscr{T}}(f), \tag{1.4.22}$$

where $\mathcal{T}$ is the stable Lévy tree with branching mechanism $\psi(\lambda) = \kappa\lambda^\gamma$. In particular, for every $f \in \mathscr{C}_b([0,1] \times \mathbb{R}_+)$, we have

$$\mathscr{A}_n^{\mathrm{m}\mathfrak{h},\circ}(f) = \frac{b_n}{n^2} \sum_{w \in \tau^{n,\circ}} |\tau_w^n| f\left(\frac{|\tau_w^n|}{n}, \frac{b_n}{n}\mathfrak{h}(\tau_w^n)\right) \xrightarrow[n\to\infty]{(d)+\text{moments}} \Psi_{\mathcal{T}}^{\mathrm{m}\mathfrak{h}}(f). \tag{1.4.23}$$

Remark 1.4.12. By (1.4.18), we have that a.s. and in L^1

$$d_{\mathrm{TV}}\left(\mathscr{A}_n, \mathscr{A}_n^\circ\right) \xrightarrow[n\to\infty]{} 0.$$

In particular, the convergences of Corollary 1.4.11 still hold if we sum over τ^n instead of $\tau^{n,\circ}$.

Remark 1.4.13. Another model of random trees is the uniform Pólya trees which are rooted, unlabelled and unordered trees. In [131], Panagiotou and Stufler show that the scaling limit of uniform Pólya trees is the Brownian tree and that the sub-exponential tail bounds of Lemma 1.4.2 hold in this case with $\alpha = \beta = 2$. Let $\Omega \subset \mathbb{N}$ be such that $\Omega \cap \{0,1\} \neq \Omega$ and let T^n denote the uniform random unordered tree with n vertices and vertex outdegree in Ω. Then there exists a finite constant $c_\Omega > 0$ such that $(c_\Omega/\sqrt{n})\mathsf{T}^n$ converges in distribution to the Brownian tree $\mathcal{T}$ with branching mechanism $\psi(\lambda) = 2\lambda^2$. Thus, the result of Corollary 1.4.11 holds for T^n and the proof is exactly the same as in the BGW case: for every $f \in \mathscr{C}_b(\mathbb{T} \times \mathbb{R}_+)$,

$$\frac{c_\Omega}{n^{3/2}} \sum_{w \in \mathsf{T}^{n,\circ}} |\mathsf{T}_w^n| f\left(\frac{c_\Omega}{\sqrt{n}}\mathsf{T}_w^n, \frac{c_\Omega}{\sqrt{n}} H(w)\right) \xrightarrow[n\to\infty]{(d)+\text{moments}} \Psi_{\mathcal{T}}(f).$$

Proof of Corollary 1.4.11. Denote by $\mathcal{T}^n = (b_n/n)\tau^n$ the real tree associated with τ^n rescaled so that all edges have length b_n/n and equipped with the uniform probability measure on the set of vertices whose heights are integer multiples of b_n/n. By Lemma 1.4.9, we have

$$d_{\mathrm{BL}}\left(\Psi_{\mathcal{T}^n}, \mathscr{A}_n^\circ\right) \le d_{\mathrm{BL}}\left(\Psi_{\mathcal{T}^n}, \mathscr{A}_n\right) + 2d_{\mathrm{TV}}(\mathscr{A}_n, \mathscr{A}_n^\circ) \le \frac{b_n}{n}\left(\frac{3}{4}\mathscr{A}_n(1) + 2\right).$$

Thanks to (1.4.21) and Lemma 1.4.4, we have that $M = \sup_{n\in\Delta} \mathbb{E}[\mathscr{A}_n(1)]$ is finite. It follows that

$$\limsup_{n\to\infty} \mathbb{E}\left[d_{\mathrm{BL}}\left(\Psi_{\mathcal{T}^n}, \mathscr{A}_n^\circ\right)\right] \le \lim_{n\to\infty} \frac{b_n}{n}\left(\frac{3M}{4} + 2\right) = 0.$$

Thus, using that $\Psi_{\mathcal{T}^n} \xrightarrow{(d)} \Psi_{\mathcal{T}}$ in $\mathscr{M}(\mathbb{T} \times \mathbb{R}_+)$ by Corollary 1.4.8, Slutsky's lemma yields the convergence in distribution $\mathscr{A}_n^\circ \xrightarrow{(d)} \Psi_{\mathcal{T}}$ in $\mathscr{M}(\mathbb{T} \times \mathbb{R}_+)$ which proves (1.4.22).

Let $f \in \mathscr{C}_b(\mathbb{T} \times \mathbb{R}_+)$. Using Skorokhod's representation theorem, we may assume that the convergence (1.4.22) holds almost surely. To prove the convergence of positive moments, it suffices to show that the family $(\mathscr{A}_n^\circ(f), n \in \Delta)$ is bounded in L^p for every $p \in [1,\infty)$. This is the case as by (1.4.21), we have $\mathscr{A}_n^\circ(f) \le \|f\|_\infty \mathscr{A}_n^\circ(1) \le \|f\|_\infty \frac{b_n}{n}\mathfrak{h}(\tau^n)$, and the family $(\frac{b_n}{n}\mathfrak{h}(\tau^n), n \in \Delta)$ is bounded in L^p for every $p \in [1,\infty)$ by Lemma 1.4.4. This completes the proof. □

The Gromov-Hausdorff-Prokhorov convergence (1.4.15) allowed us to derive an invariance principle (1.4.22) for a certain class of additive functionals on BGW trees, namely those associated with real-valued continuous bounded functions f defined on $\mathbb{T} \times \mathbb{R}_+$. In the sequel, we will be looking at a similar invariance principle when f blows up on $\mathbb{T}_0 \times \mathbb{R}_+$. It is not surprising that the Gromov-Hausdorff-Prokhorov convergence alone does not allow us to say anything about the convergence of $\Psi_{\mathcal{T}^n}(f)$ in this case as the next remark illustrates.

Remark 1.4.14. Let τ^n be a Catalan tree with n vertices, where $n \in \Delta = 2\mathbb{N}+1$. In other words, τ^n is uniformly distributed among the set of full binary ordered trees with n vertices, which corresponds to a BGW(ξ) tree with $\mathbb{P}(\xi = 0) = \mathbb{P}(\xi = 2) = 1/2$ conditioned to have size n. Notice that ξ has finite variance $\sigma_\xi^2 = 1$. Take $b_n = \sqrt{n}/2$ so that by (1.4.15), $\mathcal{T}^n = (1/2\sqrt{n})\tau^n$ converges in distribution in $\mathbb{T}$ to the Brownian continuum random tree $\mathcal{T}$ with branching mechanism $\psi(\lambda) = 2\lambda^2$. In fact, it is well known, see e.g. [137, Theorem 7.9], that there is a representation of $\mathcal{T}^n$ such that the almost sure convergence holds. Denote by $\mathcal{T}_\varepsilon^n$ the real tree obtained from $\mathcal{T}^n$ by stretching the leaves by a distance of $\varepsilon \geq 0$ and equip it with the uniform probability measure on the set of branching points and leaves. Fix $0 < \alpha < 1/2$ and set $\varepsilon_n = n^{-\alpha}$. It is clear from this construction that $\mathcal{T}_{\varepsilon_n}^n$ is a $\mathbb{T}$-valued random variable and that a.s.

$$d_{\mathrm{GHP}}\left(\mathcal{T}_{\varepsilon_n}^n, \mathcal{T}^n\right) \leq \varepsilon_n.$$

So it follows that $\mathcal{T}_{\varepsilon_n}^n$ converges to $\mathcal{T}$ a.s. in the sense of the Gromov-Hausdorff-Prokhorov distance. We consider $f(T, r) = \mathfrak{m}(T)^{-\alpha}$ and if $\nu \in \mathcal{M}(\mathbb{T} \times \mathbb{R}_+)$ we write $\nu(x^{-\alpha})$ for $\nu(f)$. According to [52, Theorem 3.1], we have the following a.s. convergence $\mathcal{A}_n(x^{-\alpha}) \xrightarrow[n\to\infty]{} \Psi_{\mathcal{T}}(x^{-\alpha})$. In conjunction with the identity $\Psi_{\mathcal{T}^n}(x^{-\alpha}) = \mathcal{A}_n(x^{-\alpha}) - 1/(2\sqrt{n})$, this proves the a.s. convergence

$$\Psi_{\mathcal{T}^n}(x^{-\alpha}) \xrightarrow[n\to\infty]{} \Psi_{\mathcal{T}}(x^{-\alpha}).$$

On the other hand, we have

$$\Psi_{\mathcal{T}_{\varepsilon_n}^n}(x^{-\alpha}) - \Psi_{\mathcal{T}^n}(x^{-\alpha}) = \frac{1}{|\tau^n|} \sum_{w \in \mathrm{Lf}(\tau^n)} \int_{(2\sqrt{n})^{-1}H(w)}^{(2\sqrt{n})^{-1}H(w)+\varepsilon_n} \left(\frac{|\tau_w^n|}{|\tau^n|}\right)^{-\alpha} \mathrm{d}r = \frac{n+1}{2} n^{\alpha-1} \varepsilon_n$$

since $|\tau^n| = n$ and $|\mathrm{Lf}(\tau^n)| = (n+1)/2$. Thus, we get

$$\Psi_{\mathcal{T}_{\varepsilon_n}^n}(x^{-\alpha}) - \Psi_{\mathcal{T}^n}(x^{-\alpha}) \xrightarrow[n\to\infty]{} \frac{1}{2}.$$

In conclusion, even though we have the a.s. convergence $\mathcal{T}_{\varepsilon_n}^n$ towards $\mathcal{T}$ in $\mathbb{T}$, $\Psi_{\mathcal{T}_{\varepsilon_n}^n}(x^{-\alpha})$ does not converge to $\Psi_{\mathcal{T}}(x^{-\alpha})$ for $\alpha \in (0, 1/2)$. This proves that the continuity of $\Psi_T(f)$ in T when f blows up on $\mathbb{T}_0$, which has been observed in [52], is indeed specific to BGW trees.

1.5 Technical lemmas

In this section, we gather some technical results that will be used later. The next lemma, which gives sufficient conditions for boundedness in L^1 of functionals of the mass and height on BGW trees, will be a key ingredient in proving our convergence results. Recall that τ is a BGW(ξ) tree and τ^n is a BGW(ξ) conditioned to have n vertices. Recall from (1.4.20) the definition of the measure $\mathscr{A}_n^{\mathfrak{mh},\circ}$ and notice that $\mathscr{A}_n^{\mathfrak{mh},\circ}([0,1]\times\mathbb{R}_+ \setminus (0,1]\times\mathbb{R}_+^*) = 0$. For this reason, we also see $\mathscr{A}_n^{\mathfrak{mh},\circ}$ as a measure on $(0,1]\times\mathbb{R}_+^*$. By convention, we write $\mathscr{A}_n^{\mathfrak{mh},\circ}(g(x)h(u))$ for $\mathscr{A}_n^{\mathfrak{mh},\circ}(f)$ where $f(x,u)=g(x)h(u)$, and we see g as a function of the mass and h as a function of the height.

Lemma 1.5.1. *Assume that ξ satisfies $(\xi 1)$ and $(\xi 2)'$. Suppose that $f \in \mathscr{B}_+((0,1]\times\mathbb{R}_+^*)$ satisfies one of the following assumptions:*

(i) f is of the form $f(x,u)=g(x)u^\beta$ or $f(x,u)=x^\alpha h(u)$ where $\alpha,\beta\in\mathbb{R}$ and g,h are nonincreasing and

$$\int_0 f(x^{\gamma/(\gamma-1)}, x)\,\mathrm{d}x < \infty. \tag{1.5.1}$$

(ii) $f(x,u)=g(x)\mathrm{e}^{u^\eta}\mathbf{1}_{[1,\infty)}(u)$ where $\eta\in(0,\gamma)$ and $g\in\mathscr{B}_+((0,1])$ is nonincreasing and satisfies $\int_0 g(x)\mathrm{e}^{-x^{-r_0}}\,\mathrm{d}x<\infty$ for some $r_0\in(0,\gamma-1)$.

Then, we have

$$\sup_{n\in\Delta}\mathbb{E}\left[\mathscr{A}_n^{\mathfrak{mh},\circ}(f)\right]<\infty.$$

Proof of Lemma 1.5.1. Here c, C and M denote positive finite constants that may vary from expression to expression (but are independent of n and x). Let $n\in\Delta$ so that $\mathbb{P}(S_n=n-1)>0$. Observe that $w\in\tau^{n,\circ}$ if and only if $|\tau_w^n|>1$ and that the root $\varnothing$ is the only vertex in τ^n such that $|\tau_w^n|=n$. Thus, for every $f\in\mathscr{B}_+([0,1]\times\mathbb{R}_+)$, we have the decomposition

$$\begin{aligned}\mathbb{E}\left[\mathscr{A}_n^{\mathfrak{mh},\circ}(f)\right] &= \frac{b_n}{n^2}\mathbb{E}\left[\sum_{w\in\tau^{n,\circ}}|\tau_w^n| f\left(\frac{|\tau_w^n|}{n},\frac{b_n}{n}\mathfrak{h}(\tau_w^n)\right)\right]\\ &= \frac{b_n}{n^2}\mathbb{E}\left[\sum_{w\in\tau^n}\mathbf{1}_{\{1<|\tau_w^n|<n\}}|\tau_w^n| f\left(\frac{|\tau_w^n|}{n},\frac{b_n}{n}\mathfrak{h}(\tau_w^n)\right)\right] + \frac{b_n}{n}\mathbb{E}\left[f\left(1,\frac{b_n}{n}\mathfrak{h}(\tau^n)\right)\right].\end{aligned}$$

By [95, Lemma 5.1], we have

$$\begin{aligned}&\frac{b_n}{n^2}\mathbb{E}\left[\sum_{w\in\tau^n}\mathbf{1}_{\{1<|\tau_w^n|<n\}}|\tau_w^n| f\left(\frac{|\tau_w^n|}{n},\frac{b_n}{n}\mathfrak{h}(\tau_w^n)\right)\right]\\ &\qquad = \frac{b_n}{n}\sum_{k=1}^{n}\frac{\mathbb{P}(S_k=k-1)\,\mathbb{P}(S_{n-k}=n-k)}{\mathbb{P}(S_n=n-1)}\mathbb{E}\left[f\left(\frac{k}{n},\frac{b_n}{n}\mathfrak{h}(\tau^k)\right)\right]\mathbf{1}_{\{1<k<n\}},\end{aligned} \tag{1.5.2}$$

where by convention the summand is zero for $k \notin \Delta$. Using Lemma 1.4.1 and (1.4.2), we get for every $n \in \Delta$ and every $1 < k < n$,

$$b_n \frac{\mathbb{P}(S_k = k-1)\,\mathbb{P}(S_{n-k} = n-k)}{\mathbb{P}(S_n = n-1)} \le C \frac{b_n^2}{b_k b_{n-k}} \le C \left(\frac{n^2}{k(n-k)} \right)^{1/\gamma}.$$

We deduce that

$$\begin{aligned} \frac{b_n}{n^2} \mathbb{E}\left[\sum_{w \in \tau^{n,\circ}} |\tau_w^n| f\left(\frac{|\tau_w^n|}{n}, \frac{b_n}{n} \mathfrak{h}(\tau_w^n) \right) \right] &\le \frac{C}{n} \sum_{k=1}^{n} g_n(k) + \frac{b_n}{n} \mathbb{E}\left[f\left(1, \frac{b_n}{n} \mathfrak{h}(\tau^n)\right) \right] \\ &= C \int_0^1 g_n(\lceil nx \rceil)\, \mathrm{d}x + \frac{b_n}{n} \mathbb{E}\left[f\left(1, \frac{b_n}{n} \mathfrak{h}(\tau^n)\right) \right], \end{aligned} \tag{1.5.3}$$

where we set

$$g_n(k) = \left(\frac{n^2}{k(n-k)} \right)^{1/\gamma} \mathbb{E}\left[f\left(\frac{k}{n}, \frac{b_n}{n} \mathfrak{h}(\tau^k) \right) \right] \mathbf{1}_{\{1<k<n\}} \quad \text{for all } k \in \Delta, \tag{1.5.4}$$

and $g_n(k) = 0$ for $k \notin \Delta$. We will constantly make use of the following inequality

$$c \left(\frac{k}{n} \right)^{1-1/\gamma} \le \frac{b_n}{n} \frac{k}{b_k} \le C \left(\frac{k}{n} \right)^{1-1/\gamma} \quad \text{for all } 1 \le k \le n, \tag{1.5.5}$$

which follows easily from (1.4.2).

First case. Assume (i). First, we consider the case $f(x,u) = g(x) u^\beta$. Since $b_n/n \to 0$, we deduce from Lemma 1.4.4 that

$$\lim_{n\to\infty} \frac{b_n}{n} \mathbb{E}\left[f\left(1, \frac{b_n}{n} \mathfrak{h}(\tau^n)\right) \right] = g(1) \lim_{n\to\infty} \frac{b_n}{n} \mathbb{E}\left[\left(\frac{b_n}{n} \mathfrak{h}(\tau^n) \right)^\beta \right] = 0. \tag{1.5.6}$$

For every $1/n < x \le (n-1)/n$, it holds that $x \le \lceil nx \rceil / n \le 2x$ and $n - \lceil nx \rceil \ge n(1-x)/2$. Thus, for every $x \in (0,1)$, using Lemma 1.4.4 for the last inequality, we have

$$\begin{aligned} g_n(\lceil nx \rceil) &\le M x^{-1/\gamma} (1-x)^{-1/\gamma} g\left(\frac{\lceil nx \rceil}{n} \right) \mathbb{E}\left[\left(\frac{b_n}{n} \mathfrak{h}(\tau^{\lceil nx \rceil}) \right)^\beta \right] \mathbf{1}_{\{1<nx\le n-1\}} \\ &\le M x^{-1/\gamma} (1-x)^{-1/\gamma} g(x) \left(\frac{b_n}{n} \frac{\lceil nx \rceil}{b_{\lceil nx \rceil}} \right)^\beta \sup_{k\in\Delta} \mathbb{E}\left[\left(\frac{b_k}{k} \mathfrak{h}(\tau^k) \right)^\beta \right] \mathbf{1}_{\{1<nx\le n-1\}} \\ &\le M x^{(\beta+1)(1-1/\gamma)-1} (1-x)^{-1/\gamma} g(x). \end{aligned}$$

It follows that

$$\int_0^1 g_n(\lceil nx \rceil)\, \mathrm{d}x \le M \int_0^1 g(x) x^{(\beta+1)(1-1/\gamma)-1} (1-x)^{-1/\gamma}\, \mathrm{d}x, \tag{1.5.7}$$

where the right-hand side is finite by (1.5.1) as $\gamma > 1$. Combining (1.5.6) and (1.5.7), it follows from (1.5.3) that

$$\sup_{n\in\Delta} \mathbb{E}\left[\mathscr{A}_n^{\mathfrak{mh},\circ}(f)\right] = \sup_{n\in\Delta} \frac{b_n}{n^2} \mathbb{E}\left[\sum_{w\in\tau^{n,\circ}} |\tau_w^n| f\left(\frac{|\tau_w^n|}{n}, \frac{b_n}{n}\mathfrak{h}(\tau_w^n)\right)\right] < \infty.$$

Next, we consider the case $f(x,u) = x^\alpha h(u)$. By Lemma 1.4.2 and (i) from Remark 1.4.3, we have, for every $k \in \Delta$,

$$\mathbb{P}\left(\frac{b_k}{k}\mathfrak{h}(\tau^k) \le y\right) \le 1 \wedge \left(C_0 \exp\left(-c_0 y^{-\gamma/(\gamma-1)}\right)\right). \tag{1.5.8}$$

Denoting by Y a random variable whose cdf is given by the right-hand side of (1.5.8) and using (1.5.5), we get, for every $2 \le k \le n$,

$$\frac{b_n}{n}\mathfrak{h}(\tau^k) \ge_{\mathrm{st}} \frac{b_n}{n}\frac{k}{b_k} Y \ge c\left(\frac{k}{n}\right)^{1-1/\gamma} Y, \tag{1.5.9}$$

where $\ge_{\mathrm{st}}$ denotes the usual stochastic order. In particular, since Y has density

$$y \mapsto C y^{-(2\gamma-1)/(\gamma-1)} \exp\left(-c_0 y^{-\gamma/(\gamma-1)}\right) \mathbf{1}_{[0,a]}(y)$$

for some $a > 0$, the first inequality in (1.5.9) applied with $k = n$ gives, for every $n \in \Delta$,

$$\mathbb{E}\left[h\left(\frac{b_n}{n}\mathfrak{h}(\tau^n)\right)\right] \le \mathbb{E}[h(Y)] \le C\int_0^\infty h(y)\mathrm{e}^{-c_0 y^{-\gamma/(\gamma-1)}} \frac{\mathrm{d}y}{y^{(2\gamma-1)/(\gamma-1)}}. \tag{1.5.10}$$

Note that the last integral is finite: indeed, since h is nonincreasing, we have

$$\int_1^\infty h(y)\mathrm{e}^{-c_0 y^{-\gamma/(\gamma-1)}} \frac{\mathrm{d}y}{y^{(2\gamma-1)/(\gamma-1)}} \le h(1)\int_1^\infty \frac{\mathrm{d}y}{y^{(2\gamma-1)/(\gamma-1)}} < \infty,$$

and by (1.5.1)

$$\int_0^1 h(y)\mathrm{e}^{-c_0 y^{-\gamma/(\gamma-1)}} \frac{\mathrm{d}y}{y^{(2\gamma-1)/(\gamma-1)}} \le \sup_{0<y\le 1} \frac{\mathrm{e}^{-c_0 y^{-\gamma/(\gamma-1)}}}{y^{1+(\alpha+1)\gamma/(\gamma-1)}} \int_0^1 h(y) y^{\alpha\gamma/(\gamma-1)}\,\mathrm{d}y < \infty. \tag{1.5.11}$$

Then, applying (1.5.9) with $k = \lceil nx \rceil$ and using the fact that h is nonincreasing, we get for every $x \in (0,1)$:

$$\begin{aligned}
g_n(\lceil nx\rceil) &\le M x^{-1/\gamma}(1-x)^{-1/\gamma}\left(\frac{\lceil nx\rceil}{n}\right)^\alpha \mathbb{E}\left[h\left(\frac{b_n}{n}\mathfrak{h}(\tau^{\lceil nx\rceil})\right)\right]\mathbf{1}_{\{1<nx\le n-1\}} \\
&\le M x^{\alpha-1/\gamma}(1-x)^{-1/\gamma}\,\mathbb{E}\left[h\left(c x^{1-1/\gamma} Y\right)\right] \\
&\le M x^{\alpha-1/\gamma}(1-x)^{-1/\gamma}\int_0^a h\left(c x^{1-1/\gamma} y\right)\mathrm{e}^{-c_0 y^{-\gamma/(\gamma-1)}} \frac{\mathrm{d}y}{y^{(2\gamma-1)/(\gamma-1)}} \\
&\le M x^{1+\alpha-1/\gamma}(1-x)^{-1/\gamma}\int_0^{acx^{1-1/\gamma}} h(u)\mathrm{e}^{-rxu^{-\gamma/(\gamma-1)}} \frac{\mathrm{d}u}{u^{(2\gamma-1)/(\gamma-1)}},
\end{aligned}$$

for some positive constant $r > 0$, where in the last inequality we made the change of variable $u = cx^{1-1/\gamma}y$. Therefore we have

$$\int_0^1 g_n(\lceil nx \rceil)\,\mathrm{d}x \le M \int_0^1 x^{1+\alpha-1/\gamma}(1-x)^{-1/\gamma}\,\mathrm{d}x \int_0^{acx^{1-1/\gamma}} h(u)\mathrm{e}^{-rxu^{-\gamma/(\gamma-1)}} \frac{\mathrm{d}u}{u^{(2\gamma-1)/(\gamma-1)}}. \tag{1.5.12}$$

It remains to check that the last integral is finite. But, arguing as in (1.5.11) with r instead of c_0, we have

$$\begin{aligned}\int_{1/2}^1 x^{1+\alpha-1/\gamma}(1-x)^{-1/\gamma}\,\mathrm{d}x \int_0^{acx^{1-1/\gamma}} h(u)\mathrm{e}^{-rxu^{-\gamma/(\gamma-1)}} \frac{\mathrm{d}u}{u^{(2\gamma-1)/(\gamma-1)}}&\\ \le M \int_{1/2}^1 (1-x)^{-1/\gamma}\,\mathrm{d}x \int_0^{ac} h(u)\mathrm{e}^{-ru^{-\gamma/(\gamma-1)}/2} \frac{\mathrm{d}u}{u^{(2\gamma-1)/(\gamma-1)}} < \infty.&\end{aligned}$$

Let $\delta = \gamma/(\gamma-1)$. Making the change of variable $y = xu^{-\delta}$ with u fixed, we have, thanks to (1.5.1),

$$\begin{aligned}&\int_0^{1/2} x^{1+\alpha-1/\gamma}(1-x)^{-1/\gamma}\,\mathrm{d}x \int_0^{acx^{1-1/\gamma}} h(u)\mathrm{e}^{-rxu^{-\delta}} \frac{\mathrm{d}u}{u^{1+\delta}}\\ &\qquad \le \max_{0\le x\le 1/2}(1-x)^{-1/\gamma} \int_{(ac)^{-\delta}}^{\infty} y^{1+\alpha-1/\gamma}\mathrm{e}^{-ry}\,\mathrm{d}y \int_0^\infty h(u)u^{\alpha\delta}\mathbf{1}_{\{yu^\delta\le 1/2\}}\,\mathrm{d}u\\ &\qquad \le 2^{1/\gamma}\int_{(ac)^{-\delta}}^{\infty} y^{1+\alpha-1/\gamma}\mathrm{e}^{-ry}\,\mathrm{d}y \int_0^{ac} h(u)u^{\alpha\delta}\,\mathrm{d}u < \infty.\end{aligned}$$

The right-hand side of (1.5.10) and (1.5.12) being finite and $(b_n/n,\, n \ge 1)$ being bounded, we deduce from (1.5.3) that

$$\sup_{n\in\Delta}\mathbb{E}\left[\mathscr{A}_n^{\mathfrak{m}\mathfrak{h},\circ}(f)\right] = \sup_{n\in\Delta}\frac{b_n}{n^2}\mathbb{E}\left[\sum_{w\in\tau^{n,\circ}} |\tau_w^n| f\left(\frac{|\tau_w^n|}{n}, \frac{b_n}{n}\mathfrak{h}(\tau_w^n)\right)\right] < \infty.$$

Second case. Assume (ii). Fix $\eta \in (0,\gamma)$ and set $h(u) = \mathrm{e}^{u^\eta}\mathbf{1}_{\{u\ge 1\}}$. Choose $\beta \in (\eta,\gamma)$ such that $\beta(1-1/\gamma) > r_0$. By (1.4.9) and (1.5.5), we have, for every $k \in \Delta$ such that $2 \le k \le n$,

$$\frac{b_n}{n}\mathfrak{h}(\tau^k) \le_{\mathrm{st}} \frac{b_n}{n}\frac{k}{b_k} Z \le C\left(\frac{k}{n}\right)^{1-1/\gamma} Z, \tag{1.5.13}$$

where Z has density $z \mapsto Mz^{\beta-1}\mathrm{e}^{-c_0 z^\beta}\mathbf{1}_{[a,\infty)}(z)$ for some $a > 0$. So, we get for $x \in (0,1)$

$$\begin{aligned} g_n(\lceil nx \rceil) &\le Mx^{-1/\gamma}(1-x)^{-1/\gamma} g\left(\frac{\lceil nx \rceil}{n}\right)\mathbb{E}\left[h\left(\frac{b_n}{n}\mathfrak{h}(\tau^{\lceil nx \rceil})\right)\right]\\ &\le Mx^{-1/\gamma}(1-x)^{-1/\gamma} g(x)\mathbb{E}\left[h\left(Cx^{1-1/\gamma}Z\right)\right]\\ &\le Mx^{-1/\gamma}(1-x)^{-1/\gamma} g(x)\int_a^\infty h\left(Cx^{1-1/\gamma}z\right) z^{\beta-1}\mathrm{e}^{-c_0 z^\beta}\,\mathrm{d}z\\ &\le Mx^{-1/\gamma}(1-x)^{-1/\gamma} g(x)\int_a^\infty z^{\beta-1}\mathrm{e}^{c_1 z^\eta - c_0 z^\beta}\mathbf{1}_{\{Cx^{1-1/\gamma}z\ge 1\}}\,\mathrm{d}z,\end{aligned}$$

where we used (1.5.5) for the first and second inequalities, the monotonicity of g and h for the second and the fact that $\left(Cx^{1-1/\gamma}z\right)^{\eta} \le c_1 z^{\eta}$ for some finite constant $c_1 > 0$ for the last. Notice that if $r < c_0$, then the function $z \mapsto \mathrm{e}^{c_1 z^{\eta} - (c_0 - r) z^{\beta}}$ is bounded on $\mathbb{R}_+$ as $\beta > \eta$. It follows that

$$\begin{aligned}\int_0^1 g_n(\lceil nx \rceil)\, \mathrm{d}x &\le M \int_0^1 x^{-1/\gamma}(1-x)^{-1/\gamma} g(x)\, \mathrm{d}x \int_0^{\infty} z^{\beta-1} \mathrm{e}^{-rz^{\beta}} \mathbf{1}_{\{Cx^{1-1/\gamma}z \ge 1\}}\, \mathrm{d}z \\ &\le M \int_0^1 x^{-1/\gamma}(1-x)^{-1/\gamma} \mathrm{e}^{-rC^{-\beta}x^{-\beta(1-1/\gamma)}} g(x)\, \mathrm{d}x \\ &\le M \int_0^1 (1-x)^{-1/\gamma} \mathrm{e}^{-x^{-r_0}} g(x)\, \mathrm{d}x < \infty, \end{aligned} \tag{1.5.14}$$

where in the last inequality we used that the function $x \mapsto x^{-1/\gamma} \mathrm{e}^{x^{-r_0} - rC^{-\beta}x^{-\beta(1-1/\gamma)}}$ is bounded on $(0,1]$ as $\beta(1-1/\gamma) > r_0$. On the other hand, we have

$$\frac{b_n}{n} \mathbb{E}\left[f\left(1, \frac{b_n}{n} \mathfrak{h}(\tau^n)\right)\right] \le \frac{b_n}{n} g(1) \mathbb{E}[h(Z)] \le M \frac{b_n}{n} \int_1^{\infty} z^{\beta-1} \mathrm{e}^{cz^{\eta} - c_0 z^{\beta}}\, \mathrm{d}z \le M, \tag{1.5.15}$$

where we used the first inequality from (1.5.13) with $k = n$ and the fact that h is nondecreasing for the first inequality and that b_n/n converges to 0 as $n \to \infty$ for the last. Combining (1.5.14) and (1.5.15), we deduce from (1.5.3) that

$$\sup_{n \in \Delta} \mathbb{E}\left[\mathscr{A}_n^{\mathrm{m}\mathfrak{h},\circ}(f)\right] = \sup_{n \in \Delta} \frac{b_n}{n^2} \mathbb{E}\left[\sum_{w \in \tau^{n,\circ}} |\tau_w^n| f\left(\frac{|\tau_w^n|}{n}, \frac{b_n}{n} \mathfrak{h}(\tau_w^n)\right)\right] < \infty.$$

□

As a consequence of the following lemma, we get that $(\mathscr{A}_n^{\mathrm{m}\mathfrak{h},\circ}(x^{\alpha} u^{\beta}),\, n \in \Delta)$ is bounded in L^p for some $p > 1$.

Lemma 1.5.2. *Let $\alpha, \beta \in \mathbb{R}$ such that $\gamma\alpha + (\gamma-1)(\beta+1) > 0$. For every $p \ge 1$ such that $p(\gamma\alpha + (\gamma-1)\beta) > 1 - \gamma$ and $\delta \in \mathbb{R}$, we have:*

$$\sup_{n \in \Delta} \mathbb{E}\left[\left(\frac{b_n}{n} \mathfrak{h}(\tau^n)\right)^{\delta} \mathscr{A}_n^{\mathrm{m}\mathfrak{h},\circ}(x^{\alpha} u^{\beta})^p\right] < \infty. \tag{1.5.16}$$

Proof. Set $M_n = \frac{b_n}{n} \mathfrak{h}(\tau^n)$ for $n \in \Delta$. Let $p_0, q_0 \in (1, \infty)$ such that $1/p_0 + 1/q_0 = 1$. By Hölder's inequality, we have

$$\mathscr{A}_n^{\mathrm{m}\mathfrak{h},\circ}(x^{\alpha} u^{\beta})^{p_0} \le \mathscr{A}_n^{\mathrm{m}\mathfrak{h},\circ}(1)^{p_0/q_0} \mathscr{A}_n^{\mathrm{m}\mathfrak{h},\circ}(x^{p_0\alpha} u^{p_0\beta}) \le M_n^{p_0/q_0} \mathscr{A}_n^{\mathrm{m}\mathfrak{h},\circ}(x^{p_0\alpha} u^{p_0\beta}), \tag{1.5.17}$$

where for the last inequality we used the fact that $\mathscr{A}_n^{\mathrm{m}\mathfrak{h},\circ}(1) = \mathscr{A}_n^{\circ}(1)$ and $\mathscr{A}_n^{\circ}(1) \le M_n$ which holds thanks to (1.4.21). Assume that $p_0 > p$ satisfies $p_0(\gamma\alpha + (\gamma-1)\beta) > 1 - \gamma$. Set $r = p_0/p$ and s such that $1/r + 1/s = 1$. We deduce that

$$\mathbb{E}\left[M_n^{\delta} \mathscr{A}_n^{\mathrm{m}\mathfrak{h},\circ}(x^{\alpha} u^{\beta})^p\right] = \mathbb{E}\left[M_n^{\delta + p/q_0} M_n^{-p/q_0} \mathscr{A}_n^{\mathrm{m}\mathfrak{h},\circ}(x^{\alpha} u^{\beta})^p\right]$$

$$\le \mathbb{E}\left[M_n^{s(\delta+p/q_0)}\right]^{1/s} \mathbb{E}\left[M_n^{-p_0/q_0} \mathscr{A}_n^{\mathfrak{mh},\circ}(x^\alpha u^\beta)^{p_0}\right]^{1/r}$$
$$\le \mathbb{E}\left[M_n^{s(\delta+p/q_0)}\right]^{1/s} \mathbb{E}\left[\mathscr{A}_n^{\mathfrak{mh},\circ}(x^{p_0\alpha} u^{p_0\beta})\right]^{1/r},$$

where we used Hölder's inequality for the first inequality and (1.5.17) for the second. Since $p_0(\gamma\alpha+(\gamma-1)\beta) > 1-\gamma$, the function $f(x,u) = x^{p_0\alpha}u^{p_0\beta}$ satisfies assumption (i) of Lemma 1.5.1. We deduce that $\sup_{n\in\Delta}\mathbb{E}\left[\mathscr{A}_n^{\mathfrak{mh},\circ}(x^{p_0\alpha}u^{p_0\beta})\right] < \infty$. Then use Lemma 1.4.4 to get (1.5.16). □

1.6 Functionals of the mass and height on the stable Lévy tree

In this section, our goal is to study the finiteness and compute the first moment of the random variable $\Psi_{\mathscr{T}}^{\mathfrak{mh}}(f)$ where $\mathscr{T}$ is the stable Lévy tree and f is a measurable function. Recall from Section 1.4.2 that H denotes the ψ-height process under its excursion measure $\mathbf{N}$, σ is the duration of an excursion and $\mathfrak{h}$ is its height. Notice that σ and $\mathfrak{h}$ are the mass and the height of the tree $\mathscr{T}_H$ coded by H. Furthermore, the stable Lévy tree $\mathscr{T}$ (under $\mathbb{P}$) is the real tree $\mathscr{T}_H$ coded by H, see Remark 1.2.1, under $\mathbf{N}^{(1)}[\bullet] = \mathbf{N}[\bullet\,|\,\sigma=1]$.

1.6.1 On the fragmentation (on the skeleton) of Lévy trees

In this section only we consider a general continuous height process H under its excursion measure $\mathbf{N}$ associated with a branching mechanism $\psi(\lambda) = a\lambda+\beta(\lambda^2/2)+\int\pi(\mathrm{d}r)(\mathrm{e}^{-\lambda r}-1+\lambda r)$ with $a,\beta\ge 0$, π a σ-finite measure on $(0,\infty)$ such that $\int\pi(\mathrm{d}r)(r\wedge r^2) < \infty$ and such that $\int^\infty \mathrm{d}\lambda/\psi(\lambda) < \infty$. We refer to [57, Section 1] for a complete presentation of the subject.

We will present a decomposition of a general Lévy tree using Bismut's decomposition. Define the length and height of the excursion of H above level r that straddles s:

$$\sigma_{r,s} = \int_0^\sigma \mathbf{1}_{\{m(s,t)\ge r\}}\,\mathrm{d}t = \mathrm{T}^+_{r,s} - \mathrm{T}^-_{r,s} \quad\text{and}\quad \mathfrak{h}_{r,s} = \sup_{t\in[\mathrm{T}^-_{r,s},\mathrm{T}^+_{r,s}]} H(t) - r, \tag{1.6.1}$$

where $m(s,t) = \inf_{[s\wedge t, s\vee t]} H$ is the minimum of H between times s,t and $\mathrm{T}^-_{r,s} = \sup\{t < s\colon H(t) = r\}$ and $\mathrm{T}^+_{r,s} = \inf\{t > s\colon H(t) = r\}$ are the beginning and the end of the excursion of H above level r that straddles time s, see Figure 1.1. Then, we consider $H^+_{r,s} = (H^+_{r,s}(t), t\ge 0)$ the excursion of H above level r that straddles s defined by:

$$H^+_{r,s}(t) = H\big((t+\mathrm{T}^-_{r,s})\wedge \mathrm{T}^+_{r,s}\big) - r,$$

and $H^-_{r,s} = (H^-_{r,s}(t), t\ge 0)$ the excursion of H below defined as $H^-_{r,s}(t) = H(t)$ for $t\in[0,\mathrm{T}^-_{r,s}]$ and $H^-_{r,s}(t+\sigma_{r,s})$ for $t > \mathrm{T}^-_{r,s}$. Notice that the duration and height of the excursion $H^+_{r,s}$ are given by $\sigma^+_{r,s} = \sigma_{r,s}$ and $\mathfrak{h}_{r,s}$; that the duration of the excursion $H^-_{r,s}$ is given by $\sigma^-_{r,s} = \sigma - \sigma_{r,s}$; and that

$$\sigma = \sigma^+_{r,s} + \sigma^-_{r,s}. \tag{1.6.2}$$

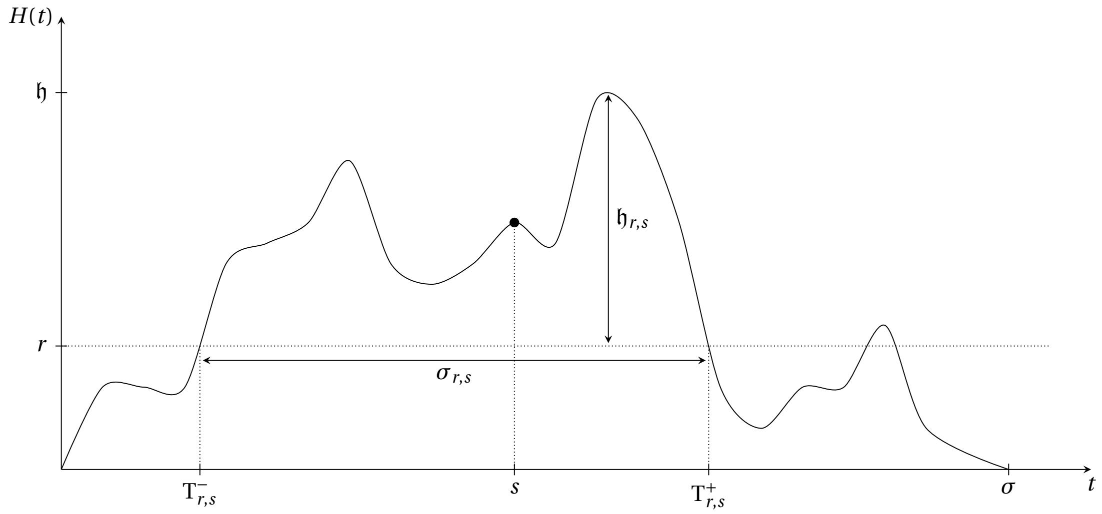

Figure 1.1 – The duration $\sigma_{r,s}$ and the height $\mathfrak{h}_{r,s}$ of the excursion of H above level r that straddles time s.

Recall notations from Remark 1.2.1. For $s \in [0,\sigma]$ and $r \in [0, H(s)]$, the function $H^+_{r,s}$ codes the subtree $\mathcal{T}_{r,s} := (\mathcal{T}_H)_{r,\, p(s)}$ and $H^-_{r,s}$ codes the subtree $\mathcal{T}^-_{r,s} := (\mathcal{T}_H \setminus \mathcal{T}_{r,s}) \cup \{x_{r,s}\}$, where $x_{r,s}$ is the ancestor of $p(s)$, the image of s in $\mathcal{T}_H$, at distance r from the root of $\mathcal{T}_H$. The next lemma says that when s and r are chosen "uniformly" under $\mathbf{N}$, then the random trees $\mathcal{T}_{r,s}$ and $\mathcal{T}^-_{r,s}$ are independent and distributed as $\mathcal{T}_H$ under $\mathbf{N}[\sigma\bullet]$. This result is a consequence of Bismut's decomposition of the excursion of the height process.

Lemma 1.6.1. *Let H be a continuous height process associated with a general branching mechanism under its excursion measure* $\mathbf{N}$. *Then for every nonnegative measurable functions f_+ and f_- defined on $\mathcal{C}_+(\mathbb{R}_+)$, we have:*

$$\mathbf{N}\left[\int_0^\sigma \mathrm{d}s \int_0^{H(s)} f_+(H^+_{r,s})\, f_-(H^-_{r,s})\, \mathrm{d}r\right] = \mathbf{N}\left[\sigma f_+(H)\right] \mathbf{N}\left[\sigma f_-(H)\right].$$

Remark 1.6.2. Lemma 1.6.1 allows to recover directly the distribution of the size of the two fragments given by the fragmentation measure $q^{ske}(\mathrm{d}s,\mathrm{d}r) = 2\beta\sigma^{-1}_{r,s}\mathbf{1}_{[0,H(s)]}(r)\,\mathrm{d}s\,\mathrm{d}r$ on the skeleton in [152, Lemma 5.1]. The Brownian case ($\pi = 0$ and $\beta > 0$) appears already in [20] and then in [10].

Proof. We follow the proof of [58, Lemma 3.4] and use notations from [57] on the càd-làg Markov process process $\left(\rho_s, \eta_s;\, s \in [0,\sigma]\right)$ under $\mathbf{N}$, which is an $\mathcal{M}(\mathbb{R}_+)^2$-valued process. The process (ρ, η) is a Markov process which allows to recover the (*a priori* non-Markovian) height process as a.s. $[0, H(t)] = \mathrm{Supp}\ (\rho_t) = \mathrm{Supp}\ (\eta_t)$. (The process ρ is called the exploration process associated with H and is strong Markov.) Thanks to [57, Proposition 3.1.3], we have that:

$$\mathbf{N}\left[\int_0^\sigma \mathrm{d}s\, F(\rho_s, \eta_s)\right] = \int \mathbb{M}(\mathrm{d}\mu, \mathrm{d}\nu)\, F(\mu, \nu), \tag{1.6.3}$$

where $\mathbb{M} = \int_0^\infty \mathrm{d}t\, \mathrm{e}^{-at}\, \mathbb{M}_{[0,t]}$ and, for any interval I, $\mathbb{M}_I$ is the law on $\mathcal{M}(\mathbb{R}_+)^2$ of the pair (μ_I, ν_I) defined by:

$$\mu_I(f) = \int \mathcal{N}(\mathrm{d}r, \mathrm{d}\ell, \mathrm{d}x)\, \mathbf{1}_I(r)\, x f(r) + \beta \int_I \mathrm{d}r\, f(r),$$
$$\nu_I(f) = \int \mathcal{N}(\mathrm{d}r, \mathrm{d}\ell, \mathrm{d}x)\, \mathbf{1}_I(r)(\ell - x) f(r) + \beta \int_I \mathrm{d}r\, f(r),$$

with $\mathcal{N}(\mathrm{d}r, \mathrm{d}\ell, \mathrm{d}x)$ a Poisson point measure on $(\mathbb{R}_+)^3$ with intensity $\mathrm{d}r\, \pi(\mathrm{d}\ell)\, \mathbf{1}_{[0,\ell]}(x)\, \mathrm{d}x$. We write $\tilde{\rho} = (\rho, \eta)$ and $\tilde{\eta} = (\eta, \rho)$. We recall that the process $\left(\rho_s;\, s \in [0,\sigma]\right)$ is strong Markov under $\mathbf{N}$, see [57, Proposition 1.2.3], and the time reversal property of (ρ, η), see [57, Corollary 3.1.6], that is $\left(\tilde{\rho}_s;\, s \in [0,\sigma]\right)$ and $\left(\tilde{\eta}_{(\sigma-s)-};\, s \in [0,\sigma]\right)$ have the same distribution under $\mathbf{N}$.

For a measure μ on $\mathbb{R}_+$ and $u > 0$ we define the measure $\mu^{[u]}$, the measure μ erased up to level u and shifted by u, by $\mu^{[u]}(f) = \int f(r-u) \mathbf{1}_{\{r>u\}}\, \mu(\mathrm{d}r)$ for $f \in \mathcal{B}_+(\mathbb{R}_+)$. We write $\tilde{\rho}^{[u]} = (\rho^{[u]}, \eta^{[u]})$ and similarly for $\tilde{\eta}$. Let F_i^ε, for $\varepsilon \in \{+,-\}$ and $i \in \{\mathrm{g}, \mathrm{d}\}$, be measurable nonnegative functionals defined on the set of càd-làg $\mathcal{M}(\mathbb{R}_+)^2$-valued functions. We shall compute:

$$A = \mathbf{N}\left[\int_0^\sigma \mathrm{d}s \int_0^{H(s)} \mathrm{d}r\, F_\mathrm{d}^+\left(\tilde{\rho}_{s+t}^{[r]};\, t \in [0, \mathrm{T}_{r,s}^+ - s]\right) F_\mathrm{g}^+\left(\tilde{\eta}_{(s-t)-}^{[r]};\, t \in [0, s - \mathrm{T}_{r,s}^-]\right)\right.$$
$$\left. F_\mathrm{d}^-\left(\tilde{\rho}_{\mathrm{T}_{r,s}^+ + t};\, t \in [0, \sigma - \mathrm{T}_{r,s}^+]\right) F_\mathrm{g}^-\left(\tilde{\eta}_{(\mathrm{T}_{r,s}^- - t)-};\, t \in [0, \mathrm{T}_{r,s}^-]\right)\right].$$

We write $\mathbf{1}_{[0,r]}\tilde{\rho} = (\mathbf{1}_{[0,r]}\rho, \mathbf{1}_{[0,r]}\eta)$. Using the Markov property of $\tilde{\rho}$ at time s, the time reversal property, again the Markov property of $\tilde{\rho}$ at time s, (1.6.3) and the transition kernel of $\tilde{\rho}$ given in [57, Proposition 3.1.2], we get that:

$$A = \mathbf{N}\left[\int_0^\sigma \mathrm{d}s \int_0^{H(s)} \mathrm{d}r\, G^+\left(\tilde{\rho}_s^{[r]}\right) G^-\left(\mathbf{1}_{[0,r]}\tilde{\rho}_s\right)\right],$$

for some measurable nonnegative functions G^- and G^+ such that for $\varepsilon \in \{+,-\}$

$$\mathbb{M}[G^\varepsilon] = \mathbf{N}\left[\int_0^\sigma \mathrm{d}s\, F_\mathrm{d}^\varepsilon(\tilde{\rho}_{s+t}, t \in [0, \sigma - s]) F_\mathrm{g}^\varepsilon(\tilde{\rho}_{(s-t)-}, t \in [0,s])\right]. \tag{1.6.4}$$

Then using (1.6.3) and the definition of $\mathbb{M}$, we get, with $\tilde{\mu} = (\mu, \nu)$:

$$\begin{aligned}
A &= \int_0^\infty \mathrm{d}t\, \mathrm{e}^{-at} \int_0^t \mathrm{d}r\, \mathbb{M}_{[0,t]}(\mathrm{d}\tilde{\mu})\, G^+\left(\tilde{\mu}^{[r]}\right) G^-\left(\mathbf{1}_{[0,r]}\tilde{\mu}\right) \\
&= \int_0^\infty \mathrm{d}t\, \mathrm{e}^{-at} \int_0^t \mathrm{d}r\, \mathbb{M}_{[0,t-r]}[G^+]\, \mathbb{M}_{[0,r]}[G^-] \\
&= \left(\int_0^\infty \mathrm{d}r\, \mathrm{e}^{-ar}\, \mathbb{M}_{[0,r]}[G^+]\right)\left(\int_0^\infty \mathrm{d}r\, \mathrm{e}^{-ar}\, \mathbb{M}_{[0,r]}[G^-]\right) \\
&= \mathbb{M}[G^+]\, \mathbb{M}[G^-],
\end{aligned}$$

where we used the independence property, that is $\mathbb{M}_I * \mathbb{M}_J = \mathbb{M}_{I \cup J}$ when I and J are disjoint, for the second equality. We deduce from (1.6.4) and the monotone class theorem that for any measurable nonnegative functionals F^+ and F^- defined on the set of càd-làg $\mathscr{M}(\mathbb{R}_+)^2$-valued functions, we have:

$$\begin{aligned}\mathbf{N}\left[\int_0^\sigma \mathrm{d}s\int_0^{H(s)} \mathrm{d}r\, F^+(\tilde\rho_{t+T^-_{r,s}}; t\in[0,\sigma_{r,s}])F^-(\tilde\rho_{t+\sigma_{r,s}\mathbf{1}_{\{t>T^-_{r,s}\}}}; t\in[0,\sigma-\sigma_{r,s}])\right]\\ = \mathbf{N}\left[\int_0^\sigma \mathrm{d}s\, F^+(\tilde\rho_t; t\in[0,\sigma])\right]\mathbf{N}\left[\int_0^\sigma \mathrm{d}s\, F^-(\tilde\rho_t; t\in[0,\sigma])\right].\\ = \mathbf{N}\left[\sigma F^+(\tilde\rho_t; t\in[0,\sigma])\right]\mathbf{N}\left[\sigma F^-(\tilde\rho_t; t\in[0,\sigma])\right].\end{aligned}$$

Then use that H is a measurable functional of the exploration process $\tilde\rho$ to conclude. □

1.6.2 First moment of $\Psi_{\mathscr{T}}$

We start with the main result of this section which gives the first moment of functionals of the stable Lévy tree. Recall that $\mathscr{T}_H$ is the real tree coded by H, see Remark 1.2.1.

Proposition 1.6.3. *Let $\mathscr{T}$ be the stable Lévy tree with branching mechanism $\psi(\lambda)=\kappa\lambda^\gamma$ where $\kappa>0$ and $\gamma\in(1,2]$. Let $f\in\mathscr{B}_+(\mathbb{T})$, and set $\tilde f(T,r)=f(T)$ for $T\in\mathbb{T}$ and $r\in\mathbb{R}_+$. We have:*

$$\mathbb{E}\left[\Psi_{\mathscr{T}}(\tilde f)\right]=\mathbf{N}\left[\sigma(1-\sigma)^{-1/\gamma}f(\mathscr{T}_H)\mathbf{1}_{\{\sigma<1\}}\right]. \tag{1.6.5}$$

Proof. Let $f\in\mathscr{B}_+(\mathbb{T})$ and set $\tilde f(T,r)=f(T)$ for $T\in\mathbb{T}$ and $r\in\mathbb{R}_+$. Using notations from Section 1.6.1, we have $\Psi_{\mathscr{T}_H}(\tilde f)=\int_0^\sigma \mathrm{d}s\int_0^{H(s)} f(\mathscr{T}_{H^+_{r,s}})\,\mathrm{d}r$. Thus, on the one hand, we get for $\lambda>0$

$$\begin{aligned}\mathbf{N}\left[\exp\{-\lambda\sigma\}\Psi_{\mathscr{T}_H}(\tilde f)\right]&=\mathbf{N}\left[\int_0^\sigma \mathrm{d}s\int_0^{H(s)}\exp\left\{-\lambda\sigma^+_{r,s}\right\}f(\mathscr{T}_{H^+_{r,s}})\exp\left\{-\lambda\sigma^-_{r,s}\right\}\mathrm{d}r\right]\\ &=\mathbf{N}\left[\sigma\exp\{-\lambda\sigma\}\right]\mathbf{N}\left[\sigma\exp\{-\lambda\sigma\}f(\mathscr{T}_H)\right]\\ &=\mathfrak{g}(0)^2\int_0^\infty \exp\{-\lambda u\}\mathbf{N}^{(u)}\left[f(\mathscr{T}_H)\right]\frac{\mathrm{d}u}{u^{1/\gamma}}\int_0^\infty\exp\{-\lambda y\}\frac{\mathrm{d}y}{y^{1/\gamma}}\\ &=\mathfrak{g}(0)^2\int_0^\infty\exp\{-\lambda r\}\,\mathrm{d}r\int_0^r\mathbf{N}^{(u)}\left[f(\mathscr{T}_H)\right]\frac{\mathrm{d}u}{(u(r-u))^{1/\gamma}},\end{aligned} \tag{1.6.6}$$

where we used (1.6.2) for the first equality, Lemma 1.6.1 for the second, (1.4.12) for the third and the change of variable $r=u+y$ for the last. On the other hand, we consider the random variable $H^r=(r^{1-1/\gamma}H(s/r), s\in[0,r])$ for $r>0$. According to (1.4.14), H^r under $\mathbf{N}^{(1)}$ is distributed as H under $\mathbf{N}^{(r)}$. Then, we have for $\lambda>0$

$$\mathbf{N}\left[\exp\{-\lambda\sigma\}\Psi_{\mathscr{T}_H}(\tilde f)\right]=\mathfrak{g}(0)\int_0^\infty\exp\{-\lambda r\}\mathbb{E}\left[\Psi_{\mathscr{T}_{H^r}}(\tilde f)\right]\frac{\mathrm{d}r}{r^{1+1/\gamma}}. \tag{1.6.7}$$

Comparing (1.6.6) and (1.6.7), we deduce that $\mathrm{d}r$-a.e., for $r>0$

$$\mathbb{E}\left[\Psi_{\mathscr{T}_{H^r}}(\tilde f)\right]=r^{1+1/\gamma}\mathfrak{g}(0)\int_0^r\frac{\mathbf{N}^{(u)}\left[f(\mathscr{T}_H)\right]}{(r-u)^{1/\gamma}}\frac{\mathrm{d}u}{u^{1/\gamma}}=r^{1+1/\gamma}\mathbf{N}\left[\sigma(r-\sigma)^{-1/\gamma}f(\mathscr{T}_H)\mathbf{1}_{\{\sigma<r\}}\right]. \tag{1.6.8}$$

From now on, we assume that $f \in \mathscr{C}_+(\mathbb{T})$ is bounded and that there exists $\varepsilon > 0$ such that $f(T) = 0$ if $\mathfrak{m}(T) > 1-\varepsilon$. As $\mathfrak{m}(\mathscr{T}_H) = \sigma$, the map $r \mapsto \mathbf{N}\left[\sigma(r-\sigma)^{-1/\gamma} f(\mathscr{T}_H)\mathbf{1}_{\{\sigma<r\}}\right]$ is continuous at $r = 1$ by dominated convergence. By definition of H^r and the continuity of the height function, we get that a.s. $\lim_{r\to 1} \left\|H^r - H^1\right\|_\infty = 0$. Following [5, Proposition 2.10], we get that the $\mathbb{T}$-valued function $r \mapsto \mathscr{T}_{H^r}$ is then a.s. continuous at $r = 1$. We deduce from Proposition 1.3.3 that $r \mapsto \Psi_{\mathscr{T}_{H^r}}(\tilde{f})$ is continuous at $r = 1$. We also have

$$\Psi_{\mathscr{T}_{H^r}}(\tilde{f}) \le \mathfrak{m}(\mathscr{T}_{H^r})\mathfrak{h}(\mathscr{T}_{H^r}) \left\|f\right\|_\infty \le r^{2-1/\gamma}\mathfrak{h}(H^1) \left\|f\right\|_\infty.$$

Since $\mathfrak{h}(H^1)$ is integrable, we deduce by dominated convergence that the map $r \mapsto \mathbb{E}\left[\Psi_{\mathscr{T}_{H^r}}(\tilde{f})\right]$ is continuous at $r = 1$. We deduce from (1.6.8) that for all $f \in \mathscr{C}_+(\mathbb{T})$ bounded and such that there exists $\varepsilon > 0$ for which $f(T) = 0$ if $\mathfrak{m}(T) > 1-\varepsilon$, we have:

$$\mathbb{E}\left[\Psi_{\mathscr{T}_{H^1}}(\tilde{f})\right] = \mathbf{N}\left[\sigma(1-\sigma)^{-1/\gamma} f(\mathscr{T}_H)\mathbf{1}_{\{\sigma<1\}}\right].$$

By monotone convergence, this equality holds if $f \in \mathscr{C}_+(\mathbb{T})$ is bounded. Then use that $\mathscr{T}_{H^1}$ is distributed as $\mathscr{T}$ to get (1.6.5). □

The next result is a direct consequence of Proposition 1.6.3, combined with the fact that π_*, defined in (1.4.12), is the distribution of σ under $\mathbf{N}$. Recall the notation $\Psi^{\mathfrak{mh}}_{\mathscr{T}}(g(x)h(u))$ which means that g is a function of the mass and h a function of the height.

Corollary 1.6.4. *Let $\mathscr{T}$ be the stable Lévy tree with branching mechanism $\psi(\lambda) = \kappa\lambda^\gamma$ where $\kappa > 0$ and $\gamma \in (1,2]$. Then we have for every $f \in \mathscr{B}_+([0,1]\times\mathbb{R}_+)$*

$$\mathbb{E}\left[\Psi^{\mathfrak{mh}}_{\mathscr{T}}(f)\right] = \mathfrak{g}(0)\int_0^1 x^{-1/\gamma}(1-x)^{-1/\gamma}\,\mathbb{E}\left[f\left(x, x^{1-1/\gamma}\mathfrak{h}(\mathscr{T})\right)\right]\mathrm{d}x, \tag{1.6.9}$$

where $\mathfrak{g}(0)$ is given in (1.4.3). In particular, we have for every $g \in \mathscr{B}_+([0,1])$

$$\mathbb{E}\left[\Psi^{\mathfrak{mh}}_{\mathscr{T}}(g(x))\right] = \mathfrak{g}(0)\int_0^1 x^{-1/\gamma}(1-x)^{-1/\gamma} g(x)\,\mathrm{d}x.$$

Remark 1.6.5. An equivalent way to state (1.6.9) is the following equality of measures

$$\mathbb{E}\left[\Psi^{\mathfrak{mh}}_{\mathscr{T}}(f)\right] = C(\gamma,\kappa)\,\mathbb{E}\left[f\left(V, V^{1-1/\gamma}\mathfrak{h}(\mathscr{T})\right)\right] \quad \text{with} \quad C(\gamma,\kappa) = \mathrm{B}(1-1/\gamma, 1-1/\gamma)\mathfrak{g}(0),$$

where V is a random variable with distribution $\mathrm{Beta}(1-1/\gamma, 1-1/\gamma)$, independent of $\mathfrak{h}(\mathscr{T})$ and B is the beta function. Using (1.3.4), this can be interpreted in the following way where we recall that ℓ denotes the length measure on a real tree: taking a stable Lévy tree $\mathscr{T}$ under $\mathbb{P}$ and simultaneously choosing a vertex $y \in \mathscr{T}$ uniformly according to the measure $C(\gamma,\kappa)^{-1}\mu(\mathscr{T}_y)\ell(\mathrm{d}y)$, then the mass and height of the subtree $\mathscr{T}_y$ are jointly distributed as V and $V^{1-1/\gamma}\mathfrak{h}(\mathscr{T})$.

While the measure $\mathbb{E}\left[\Psi^{\mathfrak{mh}}_{\mathscr{T}}(\bullet)\right]$ is not known explicitly, its moments can be expressed in terms of the moments of $\mathfrak{h}(\mathscr{T})$.

Corollary 1.6.6. *Let $\mathscr{T}$ be the stable Lévy tree with branching mechanism $\psi(\lambda) = \kappa\lambda^\gamma$. For every $\alpha, \beta \in \mathbb{C}$ such that $\Re(\gamma\alpha + (\gamma-1)(\beta+1)) > 0$, we have*

$$\mathbb{E}\left[\Psi^{\mathfrak{mh}}_{\mathscr{T}}(x^\alpha u^\beta)\right] = \mathfrak{g}(0)\mathrm{B}\big(\alpha + (\beta+1)(1-1/\gamma), 1-1/\gamma\big)\mathbb{E}\left[\mathfrak{h}(\mathscr{T})^\beta\right], \tag{1.6.10}$$

where B *is the beta function.*

Observe that $\mathfrak{h}(\mathscr{T})$ has finite moments of all order. This can be seen as a consequence of the convergence in distribution $\frac{b_n}{n}\mathfrak{h}(\tau^n) \xrightarrow{(d)} \mathfrak{h}(\mathscr{T})$ together with the fact that $\left(\frac{b_n}{n}\mathfrak{h}(\tau^n), n \in \mathbb{N}\right)$ is bounded in L^p for every $p \in \mathbb{R}$ by Lemma 1.4.4. The first moment of $\mathfrak{h}(\mathscr{T})$ is given in [60, Proposition 3.4]. We shall discuss the other moments in a future work.

Note that by taking $\beta = 0$, we recover [52, Lemma 4.6]. Heuristically, the condition $\Re(\gamma\alpha + (\gamma-1)(\beta+1)) > 0$ is due to the fact that under the excursion measure $\mathbf{N}$, the height $\mathfrak{h}$ scales as $\sigma^{1-1/\gamma}$ (see also Lemma 1.6.11 below), implying that for $\alpha, \beta \in \mathbb{R}$:

$$\mathbb{E}\left[\int_{\mathscr{T}} \mu(\mathrm{d}x) \int_0^{H(x)} \mathfrak{m}(\mathscr{T}_{r,x})^\alpha\, \mathfrak{h}(\mathscr{T}_{r,x})^\beta \,\mathrm{d}r\right] < \infty \Longleftrightarrow \mathbb{E}\left[\int_{\mathscr{T}} \mu(\mathrm{d}x) \int_0^{H(x)} \mathfrak{m}(\mathscr{T}_{r,x})^{\alpha+\beta(1-1/\gamma)} \,\mathrm{d}r\right] < \infty.$$

Thus, the condition on α, β corresponds to the phase transition observed in [52, Lemma 4.6 and Remark 4.8] for functionals depending only on the mass (that is $\beta = 0$).

In the Brownian case, $\mathfrak{h}(\mathscr{T})$ is the maximum of the (scaled) Brownian excursion whose moments are known explicitly. Therefore we get an explicit formula for the moments of the measure $\mathbb{E}\left[\Psi^{\mathfrak{mh}}_{\mathscr{T}}(\bullet)\right]$.

Corollary 1.6.7. *Let $\mathscr{T}$ be the Brownian tree with branching mechanism $\psi(\lambda) = \kappa\lambda^2$. For every $\alpha, \beta \in \mathbb{C}$ such that $\Re(2\alpha + \beta + 1) > 0$, we have*

$$\mathbb{E}\left[\Psi^{\mathfrak{mh}}_{\mathscr{T}}(x^\alpha u^\beta)\right] = \frac{1}{\sqrt{\pi\kappa}}\left(\frac{\pi}{\kappa}\right)^{\beta/2} \xi(\beta)\mathrm{B}\left(\alpha + \frac{\beta+1}{2}, \frac{1}{2}\right), \tag{1.6.11}$$

where ξ is the Riemann xi function defined by $\xi(s) = \frac{1}{2}s(s-1)\pi^{-s/2}\Gamma(s/2)\zeta(s)$ for every $s \in \mathbb{C}$ and ζ is the Riemann zeta function.

Proof. The normalized excursion of the height process H is distributed as $\sqrt{2/\kappa}\,B_{\mathrm{ex}}$ where B_{ex} is the normalized Brownian excursion, see e.g. [57]. Therefore we get the identity $\mathfrak{h}(\mathscr{T}) \overset{(d)}{=} \sqrt{2/\kappa}\max B_{\mathrm{ex}}$. By [37, Proposition 2.1 and Eq. (4.10)], we have

$$\mathbb{E}\left[(\max B_{\mathrm{ex}})^\beta\right] = 2\left(\frac{\pi}{2}\right)^{\beta/2}\xi(\beta), \quad \forall \beta \in \mathbb{C}.$$

The result follows then from Corollary 1.6.6 and the value of $\mathfrak{g}(0)$ given in (1.4.4) . □

1.6.3 Finiteness of $\Psi_{\mathscr{T}}^{\mathfrak{mh}}(f)$

This section is devoted to the study of the finiteness of functionals of the mass and height on the stable Lévy tree. Arguing as in the proof of Lemma 1.5.2 and using Corollary 1.6.6 and the fact that $\mathfrak{h}(\mathscr{T})$ has finite moments of all orders, we get the following result.

Lemma 1.6.8. *Let $\mathscr{T}$ be the stable Lévy tree with branching mechanism $\psi(\lambda) = \kappa\lambda^\gamma$ where $\kappa > 0$ and $\gamma \in (1,2]$. Let $\alpha, \beta \in \mathbb{R}$ such that $\gamma\alpha + (\gamma-1)(\beta+1) > 0$. For every $p \geq 1$ such that $p(\gamma\alpha + (\gamma-1)\beta) > 1-\gamma$ and $\delta \in \mathbb{R}$, we have:*

$$\mathbb{E}\left[\mathfrak{h}(\mathscr{T})^\delta \Psi_{\mathscr{T}}^{\mathfrak{mh}}(x^\alpha u^\beta)^p\right] < \infty. \tag{1.6.12}$$

We now state the main result of this section which gives an integral test for the finiteness of functionals of the mass and height on the stable Lévy tree.

Proposition 1.6.9. *Let $\mathscr{T}$ be the stable Lévy tree with branching mechanism $\psi(\lambda) = \kappa\lambda^\gamma$ where $\kappa > 0$ and $\gamma \in (1,2]$. Let $f \in \mathscr{B}_+([0,1] \times \mathbb{R}_+)$ be of the form $f(x,u) = g(x)u^\beta$ or $f(x,u) = x^\alpha h(u)$ where $\alpha, \beta \in \mathbb{R}$, and g, h nonincreasing. Then we have*

$$\Psi_{\mathscr{T}}^{\mathfrak{mh}}(f) \begin{cases} < \infty & a.s., \\ = \infty & a.s., \end{cases} \tag{1.6.13}$$

according as

$$\int_0 f(x^{\gamma/(\gamma-1)}, x)\,\mathrm{d}x \begin{cases} < \infty, \\ = \infty. \end{cases} \tag{1.6.14}$$

Furthermore, if $\Psi_{\mathscr{T}}^{\mathfrak{mh}}(f)$ is a.s. finite then we have $\mathbb{E}\left[\Psi_{\mathscr{T}}^{\mathfrak{mh}}(f)\right] < \infty$.

Proof. We first prove that if $\int_0 f(x^{\gamma/(\gamma-1)}, x)\,\mathrm{d}x$ is finite then $\mathbb{E}\left[\Psi_{\mathscr{T}}^{\mathfrak{mh}}(f)\right]$ is finite and thus $\Psi_{\mathscr{T}}^{\mathfrak{mh}}(f)$ is a.s. finite.

Let $\beta \in \mathbb{R}$ and $g \in \mathscr{B}_+([0,1])$ be such that $\int_0 g(x^{\gamma/(\gamma-1)}) x^\beta\,\mathrm{d}x < \infty$. Recall that $\mathfrak{h}(\mathscr{T})$ has finite moments of all orders. Thus, by (1.6.9), we have

$$\mathbb{E}\left[\Psi_{\mathscr{T}}^{\mathfrak{mh}}(g(x)u^\beta)\right] = \mathfrak{g}(0)\,\mathbb{E}\left[\mathfrak{h}(\mathscr{T})^\beta\right] \int_0^1 g(x) x^{(\beta+1)(1-1/\gamma)-1} (1-x)^{-1/\gamma}\,\mathrm{d}x < \infty.$$

Next, let $\alpha \in \mathbb{R}$ and $h \in \mathscr{B}_+(\mathbb{R}_+)$ be nonincreasing such that $\int_0 h(x) x^{\alpha\gamma/(\gamma-1)}\,\mathrm{d}x < \infty$. Again by (1.6.9), we have

$$\mathbb{E}\left[\Psi_{\mathscr{T}}^{\mathfrak{mh}}(x^\alpha h(u))\right] = \mathfrak{g}(0) \int_0^1 x^{\alpha-1/\gamma} (1-x)^{-1/\gamma}\,\mathbb{E}\left[h\left(x^{1-1/\gamma}\mathfrak{h}(\mathscr{T})\right)\right]\mathrm{d}x.$$

Now, letting k go to infinity in (1.5.8) and using the continuity of the cdf of $\mathfrak{h}(\mathscr{T})$ (see [60]), we get that

$$\mathbb{P}\left(\mathfrak{h}(\mathscr{T}) \leq y\right) \leq 1 \wedge \left(C_0 \exp\left(-c_0 y^{-\gamma/(\gamma-1)}\right)\right) \quad \text{for all} \quad y \geq 0.$$

We deduce that $\mathfrak{h}(\mathscr{T}) \geq_{\mathrm{st}} Y$ where the cdf of the random variable Y is given by the right-hand side of the inequality above. Using that h is nonincreasing and repeating the same computations as in the proof of Lemma 1.5.1 (cf. (1.5.12)), we deduce that

$$\mathbb{E}\left[\Psi_{\mathscr{T}}^{\mathfrak{m}\mathfrak{h}}(x^\alpha h(u))\right] \leq \mathfrak{g}(0) \int_0^1 x^{\alpha-1/\gamma}(1-x)^{-1/\gamma}\, \mathbb{E}\left[h\left(x^{1-1/\gamma} Y\right)\right] \mathrm{d}x < \infty.$$

This finishes the proof of the finite case. The infinite case is more delicate and its proof is postponed to Section 1.6.4. □

We end this section with a complete description of the behavior of polynomial functionals of the mass and height on the stable Lévy tree, which is a particular case of Proposition 1.6.9 (and Lemma 1.6.8 for $\alpha > 0$ and $\beta > 0$).

Corollary 1.6.10. *Let $\mathscr{T}$ be the stable Lévy tree with branching mechanism $\psi(\lambda) = \kappa\lambda^\gamma$ with $\kappa > 0$ and $\gamma \in (1,2]$, and let $\alpha, \beta \in \mathbb{R}$. Then we have*

$$\gamma\alpha + (\gamma-1)(\beta+1) > 0 \iff \Psi_{\mathscr{T}}^{\mathfrak{m}\mathfrak{h}}(x^\alpha u^\beta) < \infty \text{ a.s.} \iff \mathbb{E}\left[\Psi_{\mathscr{T}}^{\mathfrak{m}\mathfrak{h}}(x^\alpha u^\beta)\right] < \infty, \tag{1.6.15}$$

$$\gamma\alpha + (\gamma-1)(\beta+1) \leq 0 \iff \Psi_{\mathscr{T}}^{\mathfrak{m}\mathfrak{h}}(x^\alpha u^\beta) = \infty \text{ a.s.} \iff \mathbb{E}\left[\Psi_{\mathscr{T}}^{\mathfrak{m}\mathfrak{h}}(x^\alpha u^\beta)\right] = \infty. \tag{1.6.16}$$

1.6.4 Proof of the infinite case in Proposition 1.6.9

Recall that H denotes the height process under the excursion measure $\mathbf{N}$. Recall that $\sigma_{r,s}$ and $\mathfrak{h}_{r,s}$ are the length and height of the excursion of H above level r that straddles s, see Section 1.6.1. Let $f \in \mathscr{B}_+([0,1] \times \mathbb{R}_+)$. Set

$$Z_f = \int_0^\sigma \mathrm{d}s \int_0^{H(s)} f(\sigma_{r,s}, \mathfrak{h}_{r,s})\, \mathrm{d}r. \tag{1.6.17}$$

Notice that under $\mathbf{N}^{(1)}$, the random variable Z_f is distributed as $\Psi_{\mathscr{T}}^{\mathfrak{m}\mathfrak{h}}(f)$ under $\mathbb{P}$. Using the scaling property (1.4.14) of the height process, we have the following more general result which is partially given in [52] (notice that there is a misprint in the first line of p.34 therein).

Lemma 1.6.11. *Let $\psi(\lambda) = \kappa\lambda^\gamma$ with $\kappa > 0$ and $\gamma \in (1,2]$ and let H be the ψ-height process. For every $x > 0$, the random variable*

$$\left((H(s),\, s \in [0,x]), (\sigma_{r,s}, \mathfrak{h}_{r,s};\, r \in [0, H(s)], s \in [0,x])\right)$$

under $\mathbf{N}^{(x)}$ is distributed as the following random variable under $\mathbf{N}^{(1)}$

$$\left(\left(x^{1-1/\gamma} H(s/x),\, s \in [0,x]\right), \left(x\sigma_{x^{-1+1/\gamma} r, s/x}, x^{1-1/\gamma} \mathfrak{h}_{x^{-1+1/\gamma} r, s/x};\, r \in [0, x^{1-1/\gamma} H(s/x)], s \in [0,x]\right)\right).$$

In particular, the random variable $\big((H(s),\, s\in[0,x])\,, Z_f\big)$ *under* $\mathbf{N}^{(x)}$ *is distributed as the random variable* $\Big(\big(x^{1-1/\gamma}H(s/x),\, s\in[0,x]\big)\,, x^{2-1/\gamma}Z_{f_x}\Big)$ *under* $\mathbf{N}^{(1)}$*, where* f_x *is defined by* $f_x(y,u) = f(xy, x^{1-1/\gamma}u)$ *for* $x>0$.

Conditionally on H, let U be uniformly distributed on $[0,\sigma]$ under $\mathbf{N}[\sigma\bullet]$. Using Bismut's decomposition, see e.g. [58, Theorem 4.5] or [3, Theorem 2.1], we get that under $\mathbf{N}[\sigma\bullet]$, the random variable $H(U)$ has Lebesgue distribution on $(0,\infty)$ and, conditionally on $\{H(U)=t\}$, the process $\big((\sigma_{t-r,U},\mathfrak{h}_{t-r,U})\,, 0\le r\le t\big)$ is distributed as $((\mathsf{S}_r,\mathsf{H}_r)\,, 0\le r\le t)$ where

$$\mathsf{S}_r = \sum_{s\le r}\mathfrak{m}(\mathfrak{T}_s) \quad\text{and}\quad \mathsf{H}_r = \max_{s\le r}(\mathfrak{h}(\mathfrak{T}_s)+r-s)\,, \quad \forall 0\le r\le t, \tag{1.6.18}$$

where $\mathfrak{m}(\mathfrak{T}_s)$ (resp. $\mathfrak{h}(\mathfrak{T}_s)$) stands for the mass (resp. the height) of the real tree $\mathfrak{T}_s$, and $\mathfrak{T}=(\mathfrak{T}_s,\, s\ge 0)$ is a $\mathbb{T}$-valued Poisson point process on $[0,t]$ whose intensity is given below. If $\gamma=2$, the Poisson point process $\mathfrak{T}$ has intensity $2\kappa\mathbf{N}$. To describe the intensity of $\mathfrak{T}$ for $\gamma\in(1,2)$, we introduce the probability distribution $\mathbf{P}_a$ on $\mathbb{T}$ which is the law of a random tree obtained by gluing a family of trees $(T_i, i\in I)$ at their root, with $\sum_{i\in I}\delta_{T_i}(\mathrm{d}T)$ a $\mathbb{T}$-valued Poisson point measure with intensity $a\mathbf{N}[\mathrm{d}T]$, see also [3, Section 2.6] for more details on $\mathbf{P}_a$. If $\gamma\in(1,2)$, the Poisson point process $\mathfrak{T}$ has intensity $\int_0^\infty a\pi(\mathrm{d}a)\mathbf{P}_a(\mathrm{d}T)$ where π is the Lévy measure associated with ψ given by (1.4.10). In particular, we get the equality in law

$$\int_0^{H(U)} f(\sigma_{r,U},\mathfrak{h}_{r,U})\,\mathrm{d}r \text{ under } \mathbf{N}[\sigma\bullet|H(U)=t] \quad\overset{(d)}{=}\quad \int_0^t f(\mathsf{S}_r,\mathsf{H}_r)\,\mathrm{d}r. \tag{1.6.19}$$

In the proof of [52, Lemma 4.6], see Section 8.6 and more precisely (8.20) therein, it is proven that S is a stable subordinator with Laplace transform $\mathbb{E}\left[\exp(-\lambda\mathsf{S}_1)\right]=\exp(-\gamma\kappa^{1/\gamma}\lambda^{1-1/\gamma})$. We shall determine the intensity of the Poisson point process $\mathfrak{h}(\mathfrak{T})=(\mathfrak{h}(\mathfrak{T}_s), 0\le s\le t)$. If $\gamma=2$, $\mathfrak{h}(\mathfrak{T})$ has intensity $2\kappa\mathbf{N}[\mathfrak{h}\in\mathrm{d}x]$. But, by [58, Eq. (14)], we have $\mathbf{N}[\mathfrak{h}>x]=1/(\kappa x)$. Differentiating with respect to x, we get $\mathbf{N}[\mathfrak{h}\in\mathrm{d}x]=\kappa^{-1}x^{-2}\mathbf{1}_{\{x>0\}}\,\mathrm{d}x$, so that $\mathfrak{h}(\mathfrak{T})$ has intensity $2x^{-2}\mathbf{1}_{\{x>0\}}\,\mathrm{d}x$. If $1<\gamma<2$, $\mathfrak{h}(\mathfrak{T})$ has intensity

$$\int_0^\infty a\pi(\mathrm{d}a)\,\mathbf{P}_a(\mathfrak{h}\in\mathrm{d}x).$$

Using (1.4.13) and the definition of $\mathbf{P}_a$, we have $\mathbf{P}_a(\mathfrak{h}\le x)=\mathrm{e}^{-a\mathbf{N}[\mathfrak{h}>x]}=\mathrm{e}^{-Cax^{-1/(\gamma-1)}}$ where $C=(\kappa(\gamma-1))^{-1/(\gamma-1)}$. Differentiating with respect to x, we obtain

$$\mathbf{P}_a(\mathfrak{h}\in\mathrm{d}x) = \frac{Cax^{-\gamma/(\gamma-1)}}{\gamma-1}\mathrm{e}^{-Cax^{-1/(\gamma-1)}}\mathbf{1}_{\{x>0\}}\,\mathrm{d}x.$$

Since $\pi(\mathrm{d}a)=C'a^{-1-\gamma}\,\mathrm{d}a$ where $C'=\kappa\gamma(\gamma-1)/\Gamma(2-\gamma)$ (see (1.4.10)), we deduce that, for $x>0$,

$$\begin{aligned}\int_0^\infty a\pi(\mathrm{d}a)\,\mathbf{P}_a(\mathfrak{h}\in\mathrm{d}x) &= \frac{CC'}{\gamma-1}\left(\int_0^\infty a^{1-\gamma}x^{-\gamma/(\gamma-1)}\mathrm{e}^{-Cax^{-1/(\gamma-1)}}\,\mathrm{d}a\right)\mathbf{1}_{\{x>0\}}\,\mathrm{d}x\\ &= \frac{C^{\gamma-1}C'\Gamma(2-\gamma)}{\gamma-1}\mathbf{1}_{\{x>0\}}\,\frac{\mathrm{d}x}{x^2}\end{aligned}$$

$$= \frac{\gamma}{\gamma-1}\mathbf{1}_{\{x>0\}}\frac{\mathrm{d}x}{x^2}.$$

In all cases, for $\gamma \in (1,2]$, we get that $\mathfrak{h}(\mathfrak{T})$ is a Poisson point process with intensity $(\gamma/(\gamma-1))x^{-2}\mathbf{1}_{\{x>0\}}\,\mathrm{d}x$. Intuitively, this implies that S_r is of order $r^{\gamma/(\gamma-1)}$ while H_r is of order r as $r\to 0$ which, together with (1.6.19), explains the form of the integral test (1.6.14).

Our goal now is to show that

$$\int_0 f(x^{\gamma/(\gamma-1)},x)\,\mathrm{d}x=\infty \quad\Longrightarrow\quad \int_0 f(\mathsf{S}_t,\mathsf{H}_t)\,\mathrm{d}t=\infty \quad \text{a.s.}$$

under the assumptions of Proposition 1.6.9. To do this, we adapt the proof of Theorem 1 in [66] which gives a necessary and sufficient condition for the divergence of integrals of Lévy processes. We first consider the case $f(x,u)=x^\alpha h(u)$.

Lemma 1.6.12. *Let $\alpha > -1+1/\gamma$ and $h\in\mathscr{B}_+(\mathbb{R}_+)$ be nonincreasing such that $\int_0 h(x)x^{\alpha\gamma/(\gamma-1)}\,\mathrm{d}x = \infty$. We have that a.s.*

$$\int_0 \mathsf{S}_t^\alpha h(\mathsf{H}_t)\,\mathrm{d}t=\infty.$$

Proof. Define the first passage time for $a>0$

$$\mathrm{T}(a) := \inf\{t>0:\ \mathsf{H}_t\geq a\}. \tag{1.6.20}$$

Since $t\mapsto \mathsf{H}_t$ is right-continuous, we have

$$\{\mathrm{T}(a)>t\}=\{\mathsf{H}_t<a\}. \tag{1.6.21}$$

Furthermore, since $\mathsf{H}_0=0$, it holds that a.s. $\mathrm{T}(a)>0$ for every $a>0$.

Set $F(t)=\int_0^t \mathsf{S}_s^\alpha\,\mathrm{d}s$. Clearly $F(t)<\infty$ a.s. if $\alpha\geq 0$. If $-1+1/\gamma<\alpha<0$, we have

$$\mathbb{E}[F(t)]=\int_0^t \mathbb{E}\left[\mathsf{S}_s^\alpha\right]\mathrm{d}s=\mathbb{E}\left[\mathsf{S}_1^\alpha\right]\int_0^t s^{\alpha\gamma/(\gamma-1)}\,\mathrm{d}s,$$

where we used that S is stable with index $1-1/\gamma$. Now the last integral is finite because of the condition on α and

$$\mathbb{E}\left[\mathsf{S}_1^\alpha\right]=\frac{1}{\Gamma(|\alpha|)}\int_0^\infty \mathbb{E}\left[\mathrm{e}^{-\lambda \mathsf{S}_1}\right]\lambda^{-1-\alpha}\,\mathrm{d}\lambda=\frac{1}{\Gamma(|\alpha|)}\int_0^\infty \mathrm{e}^{-\gamma\kappa^{1/\gamma}\lambda^{1-1/\gamma}}\lambda^{-1-\alpha}\,\mathrm{d}\lambda<\infty.$$

Thus, we get $F(t)<\infty$ a.s. for $\alpha>-1+1/\gamma$. Furthermore, F is nondecreasing and we have

$$\int_0^1 \mathsf{S}_t^\alpha h(\mathsf{H}_t)\,\mathrm{d}t=\int_0^1 h(\mathsf{H}_t)\,\mathrm{d}F(t). \tag{1.6.22}$$

We shall need the first and second moment of $F(\mathrm{T}(a))$ for $a > 0$. Using (1.6.21), we have that

$$\mathbb{E}\left[F(\mathrm{T}(a))\right] = \int_0^\infty \mathbb{E}\left[\mathsf{S}_t^\alpha \mathbf{1}_{\{\mathrm{T}(a)>t\}}\right] \mathrm{d}t = \int_0^\infty \mathbb{E}\left[\mathsf{S}_t^\alpha \mathbf{1}_{\{\mathsf{H}_t<a\}}\right] \mathrm{d}t.$$

On the other hand, notice that for every $s \in [0,\sigma]$, it holds that $\sigma_{0,s} = \sigma$ is the total mass and $H_{0,s} = \mathfrak{h}$ is the total height. Thus, using Bismut's decomposition, we have

$$\mathbf{N}\left[\sigma^{\alpha+1}\mathbf{1}_{\{\mathfrak{h}<a\}}\right] = \int_0^\infty \mathbf{N}\left[\sigma\sigma_{0,U}^\alpha \mathbf{1}_{\{H_{0,U}<a\}}\middle| H(U)=t\right] \mathrm{d}t = \int_0^\infty \mathbb{E}\left[\mathsf{S}_t^\alpha \mathbf{1}_{\{\mathsf{H}_t<a\}}\right] \mathrm{d}t, \tag{1.6.23}$$

where we recall that conditionally on H, under $\mathbf{N}[\sigma\bullet]$, U is uniformly distributed on $[0,\sigma]$ and $(\sigma_{0,U}, H_{0,U})$ conditionally on $\{H(U)=t\}$ is then distributed as $(\mathsf{S}_t, \mathsf{H}_t)$. We deduce that

$$\begin{aligned}
\mathbb{E}\left[F(\mathrm{T}(a))\right] &= \mathbf{N}\left[\sigma^{\alpha+1}\mathbf{1}_{\{\mathfrak{h}<a\}}\right] \\
&= \mathfrak{g}(0)\int_0^\infty x^{-1-1/\gamma}\mathbf{N}^{(x)}\left[\sigma^{\alpha+1}\mathbf{1}_{\{\mathfrak{h}<a\}}\right]\mathrm{d}x \\
&= \mathfrak{g}(0)\int_0^\infty x^{\alpha-1/\gamma}\mathbf{N}^{(1)}\left[x^{1-1/\gamma}\mathfrak{h}<a\right]\mathrm{d}x \\
&= \frac{\gamma\mathfrak{g}(0)}{(\alpha+1)\gamma-1}\mathbf{N}^{(1)}\left[\mathfrak{h}^{-1-\alpha\gamma/(\gamma-1)}\right]a^{1+\alpha\gamma/(\gamma-1)},
\end{aligned} \tag{1.6.24}$$

where we disintegrated with respect to σ for the second equality and used the scaling property (1.4.14) of the height process for the third. Recall that $\mathfrak{h}$ has finite moments of all orders under $\mathbf{N}^{(1)}$, so that $\mathbb{E}[F(\mathrm{T}(a))]$ is finite for all $a > 0$. Next, set

$$Z_\alpha^{\mathrm{m}} = \int_0^\sigma \mathrm{d}s \int_0^{H(s)} \sigma_{r,s}^\alpha\,\mathrm{d}r.$$

It follows from Lemma 1.6.11 that under $\mathbf{N}^{(x)}$, $(\mathfrak{h}, Z_\alpha^{\mathrm{m}})$ is distributed as $(x^{1-1/\gamma}\mathfrak{h}, x^{\alpha+2-1/\gamma}Z_\alpha^{\mathrm{m}})$ under $\mathbf{N}^{(1)}$. Recall that $\alpha > -1+1/\gamma$. Thus, using Bismut's decomposition as in (1.6.23), we have

$$\begin{aligned}
\mathbb{E}\left[F(\mathrm{T}(a))^2\right] &= 2\mathbb{E}\left[\int_0^\infty \mathsf{S}_t^\alpha \mathbf{1}_{\{\mathsf{H}_t<a\}}\mathrm{d}t\int_0^t \mathsf{S}_s^\alpha\,\mathrm{d}s\right] \\
&= 2\mathbf{N}\left[\sigma^{\alpha+1}\mathbf{1}_{\{\mathfrak{h}<a\}}\int_0^{H(U)}\sigma_{r,U}^\alpha\,\mathrm{d}r\right] \\
&= 2\mathbf{N}\left[\sigma^\alpha \mathbf{1}_{\{\mathfrak{h}<a\}}Z_\alpha^{\mathrm{m}}\right] \\
&= 2\mathfrak{g}(0)\int_0^\infty x^{-1-1/\gamma}\mathbf{N}^{(x)}\left[\sigma^\alpha\mathbf{1}_{\{\mathfrak{h}<a\}}Z_\alpha^{\mathrm{m}}\right]\mathrm{d}x \\
&= 2\mathfrak{g}(0)\int_0^\infty x^{-1-1/\gamma}\mathbf{N}^{(1)}\left[x^\alpha\mathbf{1}_{\{x^{1-1/\gamma}\mathfrak{h}<a\}}x^{\alpha+2-1/\gamma}Z_\alpha^{\mathrm{m}}\right]\mathrm{d}x \\
&= \frac{\mathfrak{g}(0)}{\alpha+1-1/\gamma}\mathbf{N}^{(1)}\left[\mathfrak{h}^{-2(1+\alpha\gamma/(\gamma-1))}Z_\alpha^{\mathrm{m}}\right]a^{2(1+\alpha\gamma/(\gamma-1))},
\end{aligned} \tag{1.6.25}$$

where the last term is finite by (1.6.12). Combining (1.6.24) and (1.6.25) and using Cauchy-Schwartz inequality, we deduce that there exists some finite constant $C > 0$ such that for all

$a, b > 0$

$$\mathbb{E}[F(\mathrm{T}(a))F(\mathrm{T}(b))] \le \mathbb{E}\left[F(\mathrm{T}(a))^2\right]^{1/2} \mathbb{E}\left[F(\mathrm{T}(b))^2\right]^{1/2} \le C\,\mathbb{E}[F(\mathrm{T}(a))]\,\mathbb{E}[F(\mathrm{T}(b))]. \tag{1.6.26}$$

For $i \in \mathbb{N}$, put $\mathrm{T}_i = \mathrm{T}(2^{-i})$, $h_i = h(2^{-i})$ and $\Delta h_i = h_{i+1} - h_i$. Notice that the sequence $(\mathrm{T}_i, i \in \mathbb{N})$ is nonincreasing and $\Delta h_i \ge 0$. Set $V_n = \sum_{i=1}^n F(\mathrm{T}_i)\Delta h_{i-1}$. Notice that $\mathbb{E}[V_n]$ is finite as $\mathbb{E}[F(\mathrm{T}(a))]$ is finite for all $a > 0$. By (1.6.26), we have

$$\begin{aligned}
\mathbb{E}\left[V_n^2\right] &= \sum_{i=1}^n \mathbb{E}\left[F(\mathrm{T}_i)^2\right](\Delta h_{i-1})^2 + 2\sum_{1\le i<j\le n} \mathbb{E}\left[F(\mathrm{T}_i)F(\mathrm{T}_j)\right]\Delta h_{i-1}\Delta h_{j-1} \\
&\le C\sum_{i=1}^n \mathbb{E}[F(\mathrm{T}_i)]^2(\Delta h_{i-1})^2 + 2C\sum_{1\le i<j\le n} \mathbb{E}[F(\mathrm{T}_i)]\,\mathbb{E}\left[F(\mathrm{T}_j)\right]\Delta h_{i-1}\Delta h_{j-1} \\
&= C\left(\sum_{i=1}^n \mathbb{E}[F(\mathrm{T}_i)]\,\Delta h_{i-1}\right)^2 = C\,\mathbb{E}[V_n]^2.
\end{aligned}$$

Therefore, we get that $\limsup_n \mathbb{E}[V_n]^2/\mathbb{E}\left[V_n^2\right] > 0$. By [108], it follows that

$$\mathbb{P}\left(\limsup_n \frac{V_n}{\mathbb{E}[V_n]} \ge 1\right) > 0. \tag{1.6.27}$$

Using (1.6.24), notice that for some finite constant $C > 0$, we have

$$\begin{aligned}
\int_0^1 x^{1+\alpha\gamma/(\gamma-1)}\,|\mathrm{d}h(x)| &\le \sum_{i=1}^\infty (2^{-i+1})^{1+\alpha\gamma/(\gamma-1)} \int_{2^{-i}}^{2^{-i+1}} |\mathrm{d}h(x)| \\
&= C\sum_{i=1}^\infty \mathbb{E}[F(\mathrm{T}_i)]\,\Delta h_{i-1} = C\lim_{n\to\infty} \mathbb{E}[V_n].
\end{aligned} \tag{1.6.28}$$

Since $\int_0^1 x^{1+\alpha\gamma/(\gamma-1)}\,|\mathrm{d}h(x)| \ge -h(1) + \left(1+\alpha\gamma/(\gamma-1)\right)\int_0^1 h(x)x^{\alpha\gamma/(\gamma-1)}\,\mathrm{d}x = \infty$ by assumption, it follows from (1.6.28) that $\lim_{n\to\infty}\mathbb{E}[V_n] = \infty$. Thus, using (1.6.27) and the fact that V_n is nondecreasing, we deduce that $\lim_{n\to\infty} V_n = \infty$ with positive probability, that is

$$\mathbb{P}\left(\sum_{i=1}^\infty F(\mathrm{T}_i)\Delta h_{i-1} = \infty\right) > 0. \tag{1.6.29}$$

Since h is nonincreasing, we have

$$\int_0^{\mathrm{T}_0} h(\mathrm{H}_t)\,\mathrm{d}F(t) \ge \sum_{i=0}^\infty h_{i-1}\left(F(\mathrm{T}_{i-1}) - F(\mathrm{T}_i)\right). \tag{1.6.30}$$

A summation by parts gives

$$\sum_{i=1}^n h_{i-1}\left(F(\mathrm{T}_{i-1}) - F(\mathrm{T}_i)\right) = F(\mathrm{T}_0)h_0 - F(\mathrm{T}_n)h_n + \sum_{i=1}^n F(\mathrm{T}_i)\Delta h_{i-1}. \tag{1.6.31}$$

But, notice that

$$F(\mathrm{T}_n)h_n = F(\mathrm{T}_n)h(2^{-n}) \leq \int_0^{\mathrm{T}_n} h(\mathsf{H}_t)\,\mathrm{d}F(t) \leq \int_0^{\mathrm{T}_0} h(\mathsf{H}_t)\,\mathrm{d}F(t).$$

Together with (1.6.30) and (1.6.31), this yields

$$F(\mathrm{T}_0)h_0 + \sum_{i=1}^{\infty} F(\mathrm{T}_i)\Delta h_{i-1} \leq 2\int_0^{\mathrm{T}_0} h(\mathsf{H}_t)\,\mathrm{d}F(t).$$

It follows from (1.6.29) that $\int_0^{\mathrm{T}_0} \mathsf{S}_t^{\alpha} h(\mathsf{H}_t)\,\mathrm{d}t = \int_0^{\mathrm{T}_0} h(\mathsf{H}_t)\,\mathrm{d}F(t)$ diverges with positive probability.

Finally, since the event $\left\{\int_0 \mathsf{S}_t^{\alpha} h(\mathsf{H}_t)\,\mathrm{d}t = \infty\right\}$ is $\mathscr{F}_{0+}$-measurable where $(\mathscr{F}_t)_{t\geq 0}$ is the filtration generated by the Poisson point process $\mathfrak{T}$, Blumenthal's zero-one law entails that $\int_0^1 \mathsf{S}_t^{\alpha} h(\mathsf{H}_t)\,\mathrm{d}t$ diverges with probability 1. □

Lemma 1.6.13. *Let $\beta > -1$ and $g \in \mathscr{B}_+([0,1])$ be nonincreasing such that $\int_0 g(x^{\gamma/(\gamma-1)})x^{\beta}\,\mathrm{d}x = \infty$. We have that a.s.*

$$\int_0 g(\mathsf{S}_t)\mathsf{H}_t^{\beta}\,\mathrm{d}t = \infty.$$

Proof. The proof is similar to that of Lemma 1.6.12 and we only highlight the major differences. Define the first passage time $\mathrm{T}(a) = \inf\{t > 0\colon \mathsf{S}_t > a\}$ for every $a > 0$. Since S is a stable subordinator, we have a.s. $\mathrm{T}(a) > 0$ for every $a > 0$. Set $F(t) = \int_0^t \mathsf{H}_s^{\beta}\,\mathrm{d}s$. Notice that $F(t) < \infty$ a.s. if $\beta \geq 0$. If $-1 < \beta < 0$, then using that $\mathsf{H}_s \geq s$, we have a.s. $F(t) \leq \int_0^t s^{\beta}\,\mathrm{d}s < \infty$. To compute the first moment of $F(\mathrm{T}(a))$, use Bismut's decomposition as in (1.6.23) to get

$$\begin{aligned}
\mathbb{E}[F(\mathrm{T}(a))] &= \mathbb{E}\left[\int_0^{\infty} \mathsf{H}_t^{\beta}\mathbf{1}_{\{\mathsf{S}_t<a\}}\,\mathrm{d}t\right] \\
&= \mathbf{N}\left[\sigma\mathbf{1}_{\{\sigma<a\}}\mathfrak{h}^{\beta}\right] \\
&= \mathfrak{g}(0)\int_0^a x^{(\beta+1)(1-1/\gamma)-1}\,\mathbf{N}^{(1)}\left[\mathfrak{h}^{\beta}\right]\mathrm{d}x \\
&= \frac{\mathfrak{g}(0)}{(\beta+1)(1-1/\gamma)}\mathbf{N}^{(1)}\left[\mathfrak{h}^{\beta}\right]a^{(\beta+1)(1-1/\gamma)}. \qquad (1.6.32)
\end{aligned}$$

Setting

$$Z_{\beta}^{\mathfrak{h}} = \int_0^{\sigma}\mathrm{d}s\int_0^{H(s)} H_{r,s}^{\beta}\,\mathrm{d}r$$

and using Bismut's decomposition as in (1.6.23) and the fact that under $\mathbf{N}^{(x)}$, $(\mathfrak{h}, Z_{\beta}^{\mathfrak{h}})$ is distributed as $(x^{1-1/\gamma}\mathfrak{h}, x^{(\beta+1)(1-1/\gamma)+1}Z_{\beta}^{\mathfrak{h}})$ under $\mathbf{N}^{(1)}$ by Lemma 1.6.11, we have

$$\begin{aligned}
\mathbb{E}\left[F(\mathrm{T}(a))^2\right] &= 2\mathbb{E}\left[\int_0^{\infty}\mathsf{H}_t^{\beta}\mathbf{1}_{\{\mathsf{S}_t<a\}}\,\mathrm{d}t\int_0^t \mathsf{H}_s^{\beta}\,\mathrm{d}s\right] \\
&= 2\mathbf{N}\left[\sigma\mathbf{1}_{\{\sigma<a\}}\mathfrak{h}^{\beta}\int_0^{H(U)} H_{r,U}^{\beta}\,\mathrm{d}r\right]
\end{aligned}$$

$$\begin{aligned}
&= 2\mathbf{N}\left[\mathbf{1}_{\{\sigma<a\}}\mathfrak{h}^{\beta} Z_{\beta}^{\mathfrak{h}}\right] \\
&= 2\mathfrak{g}(0)\int_0^a x^{-1-1/\gamma}\,\mathbf{N}^{(x)}\left[\mathfrak{h}^{\beta} Z_{\beta}^{\mathfrak{h}}\right]\mathrm{d}x \\
&= \frac{\mathfrak{g}(0)}{(\beta+1)(1-1/\gamma)}\mathbf{N}^{(1)}\left[\mathfrak{h}^{\beta} Z_{\beta}^{\mathfrak{h}}\right] a^{2(\beta+1)(1-1/\gamma)},
\end{aligned} \tag{1.6.33}$$

where $\mathbf{N}^{(1)}\left[\mathfrak{h}^{\beta} Z_{\beta}^{\mathfrak{h}}\right] < \infty$ by (1.6.12). Combining (1.6.32) and (1.6.33), we see that the estimate (1.6.26) holds. The rest of the proof is similar to that of Lemma 1.6.12 (with h_i replaced by $g_i = g(2^{-i})$). □

We can now finish the proof of Proposition 1.6.9. Let $f \in \mathscr{B}_+([0,1]\times\mathbb{R}_+)$ be of the form $f(x,u) = g(x)u^{\beta}$ or $f(x,u) = x^{\alpha}h(u)$ with g, h nonincreasing and such that $\int_0 f(x^{\gamma/(\gamma-1)}, x)\,\mathrm{d}x = \infty$. By Lemmas 1.6.12 and 1.6.13, we have that, in the cases $\alpha > -1+1/\gamma$ and $\beta > -1$, a.s.

$$\int_0 f(\mathsf{S}_t, \mathsf{H}_t)\,\mathrm{d}t = \infty. \tag{1.6.34}$$

Now suppose that $\alpha \le -1+1/\gamma$. Since h is nonincreasing and satisfies $\int_0 h(x)x^{\alpha\gamma/(\gamma-1)}\,\mathrm{d}x = \infty$, there exists a constant $C>0$ such that $h \ge C$ on some interval $(0,\varepsilon)$. Thus, we have

$$\int_0 \mathsf{S}_t^{\alpha} h(\mathsf{H}_t)\,\mathrm{d}t \ge C\int_0 \mathsf{S}_t^{\alpha}\,\mathrm{d}t,$$

where the last integral diverges a.s. by Lemma 1.6.13 as $\int_0 x^{\alpha\gamma/(\gamma-1)}\,\mathrm{d}x = \infty$. Similarly, if $\beta \le -1$, there exists a constant $C' > 0$ such that $g \ge C'$ on $(0,\varepsilon)$. Thus, we have

$$\int_0 g(\mathsf{S}_t)\mathsf{H}_t^{\beta}\,\mathrm{d}t \ge C'\int_0 \mathsf{H}_t^{\beta}\,\mathrm{d}t,$$

and the last integral diverges by Lemma 1.6.12 since $\int_0 x^{\beta}\,\mathrm{d}x = \infty$. This proves that (1.6.34) holds for all $\alpha, \beta \in \mathbb{R}$.

Combining (1.6.19) and (1.6.34), we deduce that

$$\begin{aligned}
\mathbf{N}[\sigma; Z_f < \infty] &= \mathbf{N}\left[\sigma; \sigma\int_0^{H(U)} f(\sigma_{r,U}, H_{r,U})\,\mathrm{d}r < \infty\right] \\
&= \int_0^{\infty} \mathbf{N}\left[\sigma; \sigma\int_0^{H(U)} f(\sigma_{r,U}, H_{r,U})\,\mathrm{d}r < \infty \,\middle|\, H(U) = t\right]\mathrm{d}t \\
&= \int_0^{\infty} \mathbb{P}\left(\mathsf{S}_t \int_0^t f(\mathsf{S}_r, \mathsf{H}_r)\,\mathrm{d}r < \infty\right)\mathrm{d}t = 0.
\end{aligned}$$

It follows that $\mathbf{N}$-a.e. $Z_f = \infty$. Disintegrating with respect to σ and using the scaling property from Lemma 1.6.11, we get

$$0 = \mathbf{N}[Z_f < \infty] = \int_0^{\infty} \mathbf{N}^{(x)}[Z_f < \infty]\,\pi_*(\mathrm{d}x) = \int_0^{\infty} \mathbf{N}^{(1)}[x^{2-1/\gamma} Z_{f_x} < \infty]\,\pi_*(\mathrm{d}x).$$

Consequently, dx-a.e. on $(0,\infty)$, we have $\mathbf{N}^{(1)}\left[Z_{f_x}<\infty\right]=0$. Suppose that $f(y,u)=g(y)u^\beta$ with g nonincreasing. Then, under $\mathbf{N}^{(1)}$, Z_{f_x} is equal to $x^{\beta(1-1/\gamma)}\int_0^1 \mathrm{d}s\int_0^{H(s)} g(x\sigma_{r,s})H_{r,s}^\beta\,\mathrm{d}r$ and we get that

$$x\longmapsto \mathbf{N}^{(1)}\left[\int_0^1 \mathrm{d}s\int_0^{H(s)} g(x\sigma_{r,s})H_{r,s}^\beta\,\mathrm{d}r<\infty\right]$$

vanishes dx-a.e. on $(0,\infty)$. Moreover, this function is nonincreasing in x as g is nonincreasing. Hence it is identically zero. In particular, taking $x=1$ yields $\mathbf{N}^{(1)}\left[Z_f<\infty\right]=0$, and thus $\Psi_{\mathscr{T}}^{\mathfrak{mh}}(f)=+\infty$ a.s. as Z_f under $\mathbf{N}^{(1)}$ is distributed as $\Psi_{\mathscr{T}}^{\mathfrak{mh}}(f)$. The same argument applies if we suppose that $f(y,u)=y^\alpha h(u)$ instead. This completes the proof.

1.7 Phase transition for functionals of the mass and height

Recall that τ^n is a BGW(ξ) conditioned to have n vertices (with $n\in\Delta$) and ξ satisfies (ξ1) and (ξ2)$'$, with the sequence $(b_n,n\in\mathbb{N}^*)$ in (1.4.1), and that $\mathscr{T}$ is a stable Lévy tree with branching mechanism $\psi(\lambda)=\kappa\lambda^\gamma$. In this section, we study the limit of

$$\mathscr{A}_n^\circ(f)=\frac{b_n}{n^2}\sum_{w\in\tau^{n,\circ}}|\tau_w^n|f\left(\frac{b_n}{n}\tau_w^n,\frac{b_n}{n}H(w)\right)$$

for functions $f\in\mathscr{B}(\mathbb{T}\times\mathbb{R}_+)$ continuous on $(\mathbb{T}\setminus\mathbb{T}_0)\times\mathbb{R}_+$ but that may blow up as either the mass or the height goes to 0.

1.7.1 A general convergence result

We now give a first convergence result for general functionals that may blow up. Recall from (1.2.5) the definition of $\mathbb{T}_0$. Notice that $\mathscr{A}_n^\circ(\mathbb{T}_0\times\mathbb{R}_+)=0$ and $\Psi_{\mathscr{T}}(\mathbb{T}_0\times\mathbb{R}_+)=0$.

Proposition 1.7.1. *Assume that ξ satisfies (ξ1) and (ξ2)$'$. Let $f\in\mathscr{B}(\mathbb{T}\times\mathbb{R}_+)$ be continuous on $(\mathbb{T}\setminus\mathbb{T}_0)\times\mathbb{R}_+$ and $\alpha,\beta\in\mathbb{R}$ with $\gamma\alpha+(\gamma-1)(\beta+1)>0$ be such that*

$$\left|f(T,r)\right|\le C\,\mathfrak{m}(T)^\alpha\mathfrak{h}(T)^\beta,\quad \textit{for all } T\in\mathbb{T}\setminus\mathbb{T}_0 \textit{ and } r\ge 0, \tag{1.7.1}$$

for some finite constant $C>0$. Then $\Psi_{\mathscr{T}}(|f|)$ is a.s. finite and we have the convergence in distribution

$$\mathscr{A}_n^\circ(f)=\frac{b_n}{n^2}\sum_{w\in\tau^{n,\circ}}|\tau_w^n|f\left(\frac{b_n}{n}\tau_w^n,\frac{b_n}{n}H(w)\right)\xrightarrow{(d)}\Psi_{\mathscr{T}}(f). \tag{1.7.2}$$

We also have the convergence of all moments of order $p\ge 1$ such that $p(\gamma\alpha+(\gamma-1)\beta)>1-\gamma$.

Proof. By Corollary 1.4.11, we know that $\mathscr{A}_n^\circ\xrightarrow{(d)}\Psi_{\mathscr{T}}$ in the space $\mathscr{M}(\mathbb{T}\times\mathbb{R}_+)$. In particular, the sequence $(\mathscr{A}_n^\circ,n\in\Delta)$ is tight (in distribution) in $\mathscr{M}(\mathbb{T}\times\mathbb{R}_+)$, and applying [99, Theorem 4.10], we have

$$\inf_{K\in\mathscr{K}}\sup_{n\in\Delta}\mathbb{E}\left[1\wedge\mathscr{A}_n^\circ(K^c)\right]=0, \tag{1.7.3}$$

where $\mathscr{K}$ is the set of all compact subsets of $\mathbb{T} \times \mathbb{R}_+$. We start by showing that

$$\inf_{K \in \mathscr{K}} \sup_{n \in \Delta} \mathbb{E}\left[\mathscr{A}_n^\circ(K^c)\right] = 0. \tag{1.7.4}$$

Let $K \in \mathscr{K}$. Using the inequality $x \le 1 \wedge x + x\sqrt{1 \wedge x}$ with $x = \mathscr{A}_n^\circ(K^c) \ge 0$ and the Cauchy–Schwartz inequality, we get that

$$\mathbb{E}\left[\mathscr{A}_n^\circ(K^c)\right] \le \mathbb{E}\left[1 \wedge \mathscr{A}_n^\circ(K^c)\right] + \sqrt{\mathbb{E}\left[\mathscr{A}_n^\circ(1)^2\right]\mathbb{E}\left[1 \wedge \mathscr{A}_n^\circ(K^c)\right]}. \tag{1.7.5}$$

Since $\mathscr{A}_n^\circ(1) \le \frac{b_n}{n}\mathfrak{h}(\tau^n)$ by (1.4.21), Lemma 1.4.4 implies that

$$\sup_{n \in \Delta} \mathbb{E}\left[\mathscr{A}_n^\circ(1)^2\right]^{1/2} \le \sup_{n \in \Delta} \mathbb{E}\left[\left(\frac{b_n}{n}\mathfrak{h}(\tau^n)\right)^2\right]^{1/2} < \infty.$$

This, in conjunction with (1.7.3) and (1.7.5), proves (1.7.4).

Let $\alpha, \beta \in \mathbb{R}$ such that $\gamma\alpha + (\gamma-1)(\beta+1) > 0$. We consider the space $S = \mathbb{T} \times \mathbb{R}_+$ with the metric $\rho((T,r),(T',r')) = d_{\mathrm{GHP}}(T,T') + |r-r'|$ and $S_0 = \mathbb{T}_0 \times \mathbb{R}_+$, so that (S,ρ) is a Polish metric space and S_0 is a closed subset of S. We shall consider $0_S = (\{\varnothing\}, 0) \in S_0$ as a distinguished point. We shall construct a family of functions $\mathfrak{F}$ on S satisfying assumptions (H1)–(H4) of Appendix 1.A in order to apply Proposition 1.A.10. Let $(\delta_k, k \in \mathbb{N})$ be a positive increasing sequence such that $(2\gamma-1)\delta_k < (\gamma-1) + \big(\gamma\alpha + (\gamma-1)\beta\big) \wedge 0$ for all $k \in \mathbb{N}$. Define for every $k \in \mathbb{N}$

$$f_k(T,r) = \left(\mathfrak{m}(T)^{\delta_k} \vee \mathfrak{m}(T)^{-\delta_k}\right)\left(\mathfrak{h}(T)^{\delta_k} \vee \mathfrak{h}(T)^{-\delta_k}\right) \quad \text{and} \quad g_k(T,r) = \mathfrak{m}(T)^\alpha \mathfrak{h}(T)^\beta f_k(T,r),$$

for all $T \in \mathbb{T} \setminus \mathbb{T}_0$ and $r \ge 0$ and $f_k = g_k = +\infty$ on $\mathbb{T}_0 \times \mathbb{R}_+$. The functions f_k and g_k are positive and continuous on $(\mathbb{T} \setminus \mathbb{T}_0) \times \mathbb{R}_+$. We define $\mathfrak{F} = \{\mathbf{1}\} \cup \{f_k, g_k \colon k \in \mathbb{N}\}$. Therefore assumptions (H1) and (H2) are satisfied. Notice that $\rho((T,r),S_0) = d_{\mathrm{GHP}}(T,\mathbb{T}_0)$. Let $\varepsilon > 0$ and $M > 0$. By (1.2.4), $d_{\mathrm{GHP}}(T,\{\varnothing\}) \le M$ implies that $\mathfrak{h}(T) \le 2M$ and $\mathfrak{m}(T) \le M$. Similarly, by Lemma 1.2.2, $d_{\mathrm{GHP}}(T,\mathbb{T}_0) \ge \varepsilon$ implies that $\mathfrak{h}(T) \ge \varepsilon$ and $\mathfrak{m}(T) \ge \varepsilon$. Therefore, we have the inclusion

$$\big\{(T,r) \in S \colon \rho((T,r),S_0) \ge \varepsilon,\, \rho((T,r),0_S) \le M\big\} \subset \{T \in \mathbb{T} \colon \mathfrak{h}(T) \in [\varepsilon, 2M],\, \mathfrak{m}(T) \in [\varepsilon, M]\} \times \mathbb{R}_+.$$

Since f_k and g_k are clearly bounded away from zero and infinity on the latter set, assumption (H3) is satisfied. Moreover, f_k/f_{k+1} and g_k/g_{k+1} are continuous and bounded on $S_0^c = (\mathbb{T} \setminus \mathbb{T}_0) \times \mathbb{R}_+$ for every $k \in \mathbb{N}$. Recall that $\rho((T,r),S_0) = d_{\mathrm{GHP}}(T,\mathbb{T}_0)$. Therefore, as $\rho((T,r),S_0) \to 0$, we have $\mathfrak{h}(T) \wedge \mathfrak{m}(T) \to 0$ by Lemma 1.2.2. It follows that $f_k(T,r)/f_{k+1}(T,r) \to 0$ and $g_k(T,r)/g_{k+1}(T,r) \to 0$ as $\rho((T,r),S_0) \to 0+$. Recall the notation $\mathfrak{F}^\star(f)$ from (H4). We deduce that $f_{k+1} \in \mathfrak{F}^\star(f_k)$ and $g_{k+1} \in \mathfrak{F}^\star(g_k)$ for $k \in \mathbb{N}^*$. We also have that $1/f_1$ is continuous and bounded on S_0^c and that $1/f_1(T,r) \to 0$ as $\rho((T,r),S_0) \to 0+$. This implies that $f_1 \in \mathfrak{F}^\star(\mathbf{1})$. Therefore, assumption (H4) is satisfied.

In order to apply Proposition 1.A.10 to the sequence of measures $(\mathscr{A}_n^\circ, n \in \Delta)$ and the family $\mathfrak{F}$, we shall check that the sequence $(\mathscr{A}_n^\circ, n \in \Delta)$ is tight (in distribution) in the space $\mathscr{M}_{\mathfrak{F}}$ (see

Appendix 1.A for the definition of $\mathscr{M}_{\mathfrak{F}}$). Thanks to Proposition 1.A.4, the sequence $(\mathscr{A}_n^\circ, n \in \Delta)$ is tight in the space $\mathscr{M}_{\mathfrak{F}}$ if and only if $(\mathfrak{f}\mathscr{A}_n^\circ, n \in \Delta)$ is tight in $\mathscr{M}(S)$ for all $\mathfrak{f} \in \mathfrak{F}$ (recall that the notation $\mathfrak{f}\mathscr{A}_n^\circ$ stands for the measure $\mathfrak{f}((T,r))\mathscr{A}_n^\circ(\mathrm{d}T, \mathrm{d}r)$). Let $\mathfrak{f} \in \mathfrak{F}$. Notice that for every $T \in \mathbb{T} \setminus \mathbb{T}_0$ and $r \geq 0$, we have

$$\mathfrak{f}((T,r)) \leq \sum_{1 \leq i,j \leq 2} \mathfrak{m}(T)^{\alpha_i} \mathfrak{h}(T)^{\beta_j}$$

for $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{R}$ such that $\gamma\alpha_i + (\gamma-1)(\beta_j+1) > 0$ holds for every $i, j \in \{1,2\}$. Therefore, by Lemma 1.5.2, we have for some $p > 1$ small enough

$$\sup_{n \in \Delta} \mathbb{E}\left[\mathscr{A}_n^\circ(\mathfrak{f})^p\right] < \infty \quad \text{and} \quad \sup_{n \in \Delta} \mathbb{E}\left[\mathscr{A}_n^\circ(\mathfrak{f}^p)\right] < \infty. \tag{1.7.6}$$

The first bound gives that (1.A.3) holds for all $\mathfrak{f} \in \mathfrak{F}$ by the Markov inequality. Recall that $\mathscr{K}$ denotes the set of compact subsets of $\mathbb{T} \times \mathbb{R}_+$. Moreover, with q such that $1/p + 1/q = 1$ and $K \in \mathscr{K}$, using Hölder's inequality, we get

$$\mathbb{E}\left[\mathscr{A}_n^\circ(\mathfrak{f}\mathbf{1}_{K^c})\right] \leq \mathbb{E}\left[\mathscr{A}_n^\circ(\mathbf{1}_{K^c})\right]^{1/q} \mathbb{E}\left[\mathscr{A}_n^\circ(\mathfrak{f}^p)\right]^{1/p}.$$

Using the second bound in (1.7.6) and (1.7.4), we deduce that

$$\inf_{K \in \mathscr{K}} \sup_{n \in \Delta} \mathbb{E}\left[\mathscr{A}_n^\circ(\mathfrak{f}\mathbf{1}_{K^c})\right] = 0.$$

Thus (1.A.4) holds for all $\mathfrak{f} \in \mathfrak{F}$. According to Proposition 1.A.4-(i), we get that the sequence $(\mathscr{A}_n^\circ, n \in \Delta)$ is tight (in distribution) in $\mathscr{M}_{\mathfrak{F}}(\mathbb{T} \times \mathbb{R}_+)$. Now apply Proposition 1.A.10 and Proposition 1.A.9 to get that

$$\mathscr{A}_n^\circ(\mathfrak{f}h) \xrightarrow[n\to\infty]{(d)} \Psi_{\mathscr{T}}(\mathfrak{f}h)$$

for every $h \in \mathscr{C}_b(\mathbb{T} \times \mathbb{R}_+)$ and every $\mathfrak{f} \in \mathfrak{F}$. Let $f \in \mathscr{B}(\mathbb{T} \times \mathbb{R}_+)$ satisfying the assumptions of Proposition 1.7.1. Consider $\mathfrak{f} = g_1$ and $h = f/g_1$. Notice that (1.7.1) implies that h is continuous on $\mathbb{T} \times \mathbb{R}_+$. Since $\mathfrak{f}h = g_1 h = f$ except possibly on $S_0 = \mathbb{T}_0 \times \mathbb{R}_+$ and $\mathscr{A}_n^\circ(S_0) = \Psi_{\mathscr{T}}(S_0) = 0$, we deduce that the convergence in distribution (1.7.2) holds.

Let $p > 1$ such that $p(\gamma\alpha + (\gamma-1)\beta) > 1 - \gamma$. There exists $q > p$ satisfying the same inequality. Since $|f(T,r)| \leq C\mathfrak{m}(T)^\alpha \mathfrak{h}(T)^\beta$, we get that

$$\sup_{n \in \Delta} \mathbb{E}\left[|\mathscr{A}_n^\circ(f)|^q\right] \leq C^q \sup_{n \in \Delta} \mathbb{E}\left[\left(\frac{b_n^{1+\beta}}{n^{2+\alpha+\beta}} \sum_{w \in \tau^{n,\circ}} |\tau_w^n|^{1+\alpha} \mathfrak{h}(\tau_w^n)^\beta\right)^q\right], \tag{1.7.7}$$

where the right-hand side is finite by Lemma 1.5.2. Thus, the sequence $(|\mathscr{A}_n^\circ(f)|^p, n \in \Delta)$ is uniformly integrable and the convergence of the moment of order p of $\mathscr{A}_n^\circ(f)$ towards the moment of order p of $\Psi_{\mathscr{T}}(f)$ readily follows from (1.7.2). $\square$

1.7.2 Phase transition for functionals of the mass and height

We refine the convergence result given in Proposition 1.7.1 for functionals depending only on the mass and height and describe a phase transition in that case.

We start with a technical lemma which is a consequence of the well-known de La Vallée Poussin criterion for uniform integrability.

Lemma 1.7.2. *Let ν be a nonnegative finite measure on $(0,1]$ and $f \in \mathscr{C}_+((0,1])$ be nonincreasing, belonging to $L^1(\nu)$ and such that $\lim_{x\to 0+} f(x) = +\infty$. Then there exists a positive function $f^\nu \in \mathscr{C}_+((0,1])$ which belongs to $L^1(\nu)$, such that f/f^ν is bounded on $(0,1]$ and $\lim_{x\to 0+} f(x)/f^\nu(x) = 0$.*

Proof. We may assume without loss of generality that f does not vanish anywhere in $(0,1]$ and that ν is a probability measure. By the de La Vallée Poussin criterion (see [50, §22]), there exists a convex nondecreasing function $F\colon \mathbb{R}_+ \to \mathbb{R}_+$ such that $\lim_{t\to\infty} F(t)/t = \infty$ and $F \circ f \in L^1(\nu)$. In fact, up to considering $F+1$ instead, we can and will assume that F does not vanish anywhere. Since F is convex on $\mathbb{R}_+$, it is continuous on $(0,\infty)$ and it follows that $F\circ f$ is continuous on $(0,1]$. Moreover, $F\circ f$ is clearly nonincreasing by composition. Further, since $\lim_{x\to 0} f(x) = \infty$ and $\lim_{t\to\infty} t/F(t) = 0$, we get $\lim_{x\to 0} f(x)/F\circ f(x) = 0$. The function $f/F\circ f$ being continuous on $(0,1]$ with a finite limit at 0, it is bounded on $(0,1]$. Setting $f^\nu = F\circ f$, the conclusion readily follows. $\square$

We now give the main result of this section. Recall that the notation $\Psi^{\mathfrak{mh}}_{\mathscr{T}}(g(x)h(u))$ stands for $\Psi^{\mathfrak{mh}}_{\mathscr{T}}(f)$ where $f(x,u) = g(x)h(u)$. For $g \in \mathscr{B}(\mathbb{R}_+)$, define

$$g^*(x) := \sup_{x\le y\le 1} |g(y)| \quad \text{for all } x \in (0,1]. \tag{1.7.8}$$

Theorem 1.7.3. *Assume that ξ satisfies (ξ 1) and (ξ 2)′.*

(i) Let $\beta \in \mathbb{R}$ and $g \in \mathscr{B}([0,1])$ be such that g is continuous on $(0,1]$ and satisfies

$$\int_0 g^*(x^{\gamma/(\gamma-1)})x^\beta \, \mathrm{d}x < \infty. \tag{1.7.9}$$

Then we have the convergence in distribution and of the first moment

$$\frac{b_n^{1+\beta}}{n^{2+\beta}} \sum_{w\in\tau^{n,\circ}} |\tau^n_w| \mathfrak{h}(\tau^n_w)^\beta g\left(\frac{|\tau^n_w|}{n}\right) \xrightarrow[n\to\infty]{(d)+\text{mean}} \Psi^{\mathfrak{mh}}_{\mathscr{T}}(g(x)u^\beta) \tag{1.7.10}$$

where $\Psi^{\mathfrak{mh}}_{\mathscr{T}}(|g(x)|u^\beta)$ is a.s. finite and integrable.

(ii) *Let $\alpha \in \mathbb{R}$ and $h \in \mathscr{B}(\mathbb{R}_+)$ be such that h is continuous on $(0,\infty)$ and satisfies $h(u) = O(\mathrm{e}^{u^\eta})$ as $u \to \infty$ for some $\eta \in (0,\gamma)$ and*

$$\int_0 x^{\alpha\gamma/(\gamma-1)} h^*(x)\,\mathrm{d}x < \infty. \tag{1.7.11}$$

Then we have the convergence in distribution and of the first moment

$$\frac{b_n}{n^{2+\alpha}} \sum_{w\in\tau^{n,\circ}} |\tau_w^n|^{1+\alpha} h\left(\frac{b_n}{n}\mathfrak{h}(\tau_w^n)\right) \xrightarrow[n\to\infty]{(d)+\text{mean}} \Psi_{\mathscr{T}}^{\mathfrak{m}\mathfrak{h}}(x^\alpha h(u)) \tag{1.7.12}$$

where $\Psi_{\mathscr{T}}^{\mathfrak{m}\mathfrak{h}}(x^\alpha|h(u)|)$ is a.s. finite and integrable.

(iii) *Let $f \in \mathscr{B}_+([0,1]\times\mathbb{R}_+)$ be such that*

$$\int_0 f(x^{\gamma/(\gamma-1)}, x)\,\mathrm{d}x = \infty. \tag{1.7.13}$$

Suppose that f is of the form $f(x,u) = g(x)u^\beta$ or $f(x,u) = x^\alpha h(u)$ where $\alpha,\beta \in \mathbb{R}$ and g,h are nonincreasing and continuous on $(0,1]$ and on $(0,\infty)$ respectively. Then we have

$$\frac{b_n}{n^2} \sum_{w\in\tau^{n,\circ}} |\tau_w^n| f\left(\frac{|\tau_w^n|}{n}, \frac{b_n}{n}\mathfrak{h}(\tau_w^n)\right) \xrightarrow[n\to\infty]{(d)+\text{mean}} \infty. \tag{1.7.14}$$

Proof. Notice that (1.7.9) implies that $\beta > -1$ as soon as g is not identically zero. To prove (i), we proceed in three steps.

Step 1 in the proof of (i). We first suppose that $g \in \mathscr{C}_+([0,1])$ is nonincreasing and nonzero. Let $(\beta_k,\ k\in\mathbb{N})$ be a decreasing sequence of nonpositive real numbers such that $\beta_0 = 0$ and $\lim_{k\to\infty}\beta_k = -1$. We define a set of functions $\mathfrak{F} = \{h_k\colon\ k\in\mathbb{N}\}$ where $h_k(u) = u^{\beta_k} \vee u^k$ for $u>0$ and $k\in\mathbb{N}$, and $h_0(0) = 1$ and $h_k(0) = +\infty$ for $k\in\mathbb{N}$. We shall prove that $\mathfrak{F}$ satisfies assumptions (H1)–(H5) of Appendix 1.A with $S = \mathbb{R}_+$ equipped with the Euclidean distance and $S_0 = \{0\}$. Notice that $h_0 \equiv 1$ and h_k is continuous on S_0^c for every $k\in\mathbb{N}$, so (H1) and (H2) are satisfied. Moreover, for every $k\in\mathbb{N}$, the function h_k/h_{k+1} is continuous on $(0,\infty)$ and we have

$$\lim_{u\to 0+}\frac{h_k(u)}{h_{k+1}(u)} = \lim_{u\to 0+} u^{\beta_k-\beta_{k+1}} = 0 \quad\text{and}\quad \lim_{u\to+\infty}\frac{h_k(u)}{h_{k+1}(u)} = \lim_{u\to+\infty}\frac{1}{u} = 0,$$

so that (H4) and (H5) are satisfied. Finally, since the set $\{x\in S\colon\ \rho(x,S_0)\geq\varepsilon,\ \rho(x,0)\leq M\} = [\varepsilon, M]$ is compact and h_k is continuous, it is bounded there and (H3) is satisfied. Define a (random) measure on $\mathbb{R}_+$ by setting

$$\zeta_n(h) = \frac{b_n}{n^2}\sum_{w\in\tau^{n,\circ}} |\tau_w^n| g\left(\frac{|\tau_w^n|}{n}\right) h\left(\frac{b_n}{n}\mathfrak{h}(\tau_w^n)\right) \tag{1.7.15}$$

for every $h\in\mathscr{B}_+(\mathbb{R}_+)$. By (1.4.23), ζ_n converges to ζ in distribution in $\mathscr{M}(\mathbb{R}_+)$ and $\mathbb{E}[\zeta_n(\bullet)]$ converges to $\mathbb{E}[\zeta(\bullet)]$ in $\mathscr{M}(\mathbb{R}_+)$ where ζ is defined by $\zeta(h) = \Psi_{\mathscr{T}}^{\mathfrak{m}\mathfrak{h}}(g(x)h(u))$. But, since we have

$\int_0 g(x)x^{(\beta_k+1)(1-1/\gamma)-1}\,\mathrm{d}x<\infty$ for every $k\in\mathbb{N}$, Lemma 1.5.1-(i) gives

$$\sup_{n\in\Delta}\mathbb{E}[\zeta_n(h_k)]\le\sup_{n\in\Delta}\mathbb{E}\left[\zeta_n(u^{\beta_k})\right]+\sup_{n\in\Delta}\mathbb{E}\left[\zeta_n(u^k)\right]<\infty\quad\text{for all }k\in\mathbb{N}.$$

Thus, Corollary 1.A.11 yields the convergence in distribution $\zeta_n\xrightarrow{(d)}\zeta$ in $\mathcal{M}_{\mathfrak{F}}$ as well as the convergence of the first moment $\mathbb{E}[\zeta_n(\bullet)]\to\mathbb{E}[\zeta(\bullet)]$ in $\mathcal{M}_{\mathfrak{F}}$. By Proposition 1.A.9, this implies that for every $g\in\mathscr{C}_+([0,1])$ nonincreasing and every $\beta>-1$, we have

$$\frac{b_n^{1+\beta}}{n^{2+\beta}}\sum_{w\in\tau^{n,\circ}}|\tau_w^n|\mathfrak{h}(\tau_w^n)^\beta g\left(\frac{|\tau_w^n|}{n}\right)\xrightarrow[n\to\infty]{(d)+\text{mean}}\Psi_{\mathcal{T}}^{\mathfrak{m}\mathfrak{h}}(g(x)u^\beta).\tag{1.7.16}$$

Step 2 in the proof of (i). Now fix $\beta>-1$ and define the (random) measure ξ_n on $[0,1]$ by

$$\xi_n(g)=\frac{b_n^{1+\beta}}{n^{2+\beta}}\sum_{w\in\tau^{n,\circ}}|\tau_w^n|\mathfrak{h}(\tau_w^n)^\beta g\left(\frac{|\tau_w^n|}{n}\right),\tag{1.7.17}$$

for every $g\in\mathscr{B}_+([0,1])$. Notice that (1.7.16) can be rewritten as

$$\xi_n(g)\xrightarrow[n\to\infty]{(d)+\text{mean}}\xi(g)\tag{1.7.18}$$

for every $g\in\mathscr{C}_+([0,1])$ nonincreasing, where the measure ξ is defined by $\xi(g)=\Psi_{\mathcal{T}}^{\mathfrak{m}\mathfrak{h}}(g(x)u^\beta)$. Moreover, Lemma 1.5.1-(i) applied with $g\equiv1$ gives $\sup_{n\in\Delta}\mathbb{E}[\xi_n(1)]<\infty$. As a consequence, by the Markov inequality, we have $\lim_{r\to\infty}\sup_{n\in\Delta}\mathbb{P}(\xi_n(1)>r)=0$. Since $[0,1]$ is compact, this means that the sequence of random measures $(\xi_n,\,n\in\Delta)$ is tight in distribution in $\mathcal{M}([0,1])$, see [99, Theorem 4.10]. Hence, it is relatively compact by Prokhorov's theorem as the space $\mathcal{M}([0,1])$ is Polish for the weak topology. Let $\hat{\xi}$ be a limit point. Then we have $\xi(g)\overset{(d)}{=}\hat{\xi}(g)$ for every $g\in\mathscr{C}_+([0,1])$ nonincreasing. Therefore, we get that $\xi\overset{(d)}{=}\hat{\xi}$ and the sequence $(\xi_n,\,n\in\Delta)$ has only one limit point ξ. Since it is relatively compact, we deduce that ξ_n converges to ξ in distribution in $\mathcal{M}([0,1])$. A similar deterministic argument shows that $\mathbb{E}[\xi_n(\bullet)]$ converges to $\mathbb{E}[\xi(\bullet)]$ in $\mathcal{M}([0,1])$.

Step 3 in the proof of (i). Let $\beta>-1$ and $g\in\mathscr{B}([0,1])$ be continuous on $(0,1]$, nonzero and such that $\int_0 g^*(x)x^{(\beta+1)(1-1/\gamma)-1}\,\mathrm{d}x<\infty$. Set $g_0\equiv1$. If $\lim_{x\to0}g^*(x)=\infty$, set $g_1=g^*+1$. If g^* has a finite limit at 0 (which is then positive), then there exists $\varepsilon>0$ such that $\int_0 x^{-\varepsilon}g^*(x)x^{(\beta+1)(1-1/\gamma)-1}\,\mathrm{d}x<\infty$. We also have $\lim_{x\to0+}x^{-\varepsilon}g^*(x)=\infty$ and the function $x\mapsto x^{-\varepsilon}g^*(x)$ is continuous on $(0,1]$ and nonincreasing. In that case, we set $g_1(x)=x^{-\varepsilon}g^*(x)+1$ for $x\in(0,1]$.

Define a set of functions $\mathfrak{F}=\{g_k:\ k\in\mathbb{N}\}$ as follows: for every $k\ge1$, set $g_{k+1}=g_k^\nu$ which is given by Lemma 1.7.2 applied with the finite measure $\nu(\mathrm{d}x)=x^{(\beta+1)(1-1/\gamma)-1}\mathrm{d}x$. By construction, the sequence $\mathfrak{F}$ satisfies assumptions (H1)–(H4) of Appendix 1.A with $S=[0,1]$, $S_0=\{0\}$ and $\mathfrak{F}^\star(g_k)=\{g_j:\ j>k\}$ (notice (H3) is automatically satisfied as $[0,1]$ is compact). Notice that, by Lemma 1.7.2, for every $k\in\mathbb{N}$, the function g_k is continuous and nonincreasing on $(0,1]$ and satisfies $\int_0 g_k(x)x^{(\beta+1)(1-1/\gamma)-1}\,\mathrm{d}x < \infty$. So, by Lemma 1.5.1, we get that

$$\sup_{n\in\Delta}\mathbb{E}\left[\xi_n(g_k)\right] < \infty \quad \text{for all } k\in\mathbb{N}.$$

Now, Corollary 1.A.11 applies and yields, in conjunction with Proposition 1.A.9, the convergence in distribution and of the first moment

$$\xi_n(g_k\ell) \xrightarrow[n\to\infty]{(d)+\text{mean}} \xi(g_k\ell)$$

for every $k\in\mathbb{N}$ and $\ell\in\mathscr{C}([0,1])$. Now apply this with $k=1$ and $\ell = g/g_1$. Notice that $g_1\ell = g$ except possibly on $S_0 = \{0\}$. Since $\xi_n(S_0) = \xi(S_0) = 0$, we deduce that

$$\xi_n(g) \xrightarrow[n\to\infty]{(d)+\text{mean}} \xi(g).$$

This, together with Proposition 1.6.9, proves (i).

The proof of (ii) is quite similar so we only indicate the changes compared with (i). Again notice that (1.7.11) implies that $\alpha > -1+1/\gamma$ as soon as h is not identically zero.
Step 1 in the proof of (ii). Let $h\in\mathscr{C}_+(\mathbb{R}^+)$ be nonincreasing and nonzero.

Taking a decreasing sequence $(\alpha_k,\, k\in\mathbb{N})$ of nonpositive real numbers such that $\alpha_0 = 0$ and $\lim_{k\to\infty}\alpha_k = -1+1/\gamma$ and defining a set of functions $\mathfrak{F} = \{g_k\colon\ k\in\mathbb{N}\}$ by $g_k(x) = x^{\alpha_k}$, we can show that for every $h\in\mathscr{C}_+(\mathbb{R}_+)$ nonincreasing and every $\alpha > -1+1/\gamma$, we have

$$\frac{b_n}{n^{2+\alpha}}\sum_{w\in\tau^{n,\circ}}|\tau^n_w|^{1+\alpha}h\left(\frac{b_n}{n}\mathfrak{h}(\tau^n_w)\right) \xrightarrow[n\to\infty]{(d)+\text{mean}} \Psi^{\mathfrak{mh}}_{\mathscr{T}}(x^\alpha h(u)). \tag{1.7.19}$$

Step 2 in the proof of (ii). Fix $\alpha > -1+1/\gamma$ and define the (random) measure ξ_n on $\mathbb{R}_+$ by

$$\xi_n(h) = \frac{b_n}{n^{2+\alpha}}\sum_{w\in\tau^{n,\circ}}|\tau^n_w|^{1+\alpha}h\left(\frac{b_n}{n}\mathfrak{h}(\tau^n_w)\right), \tag{1.7.20}$$

for every $h\in\mathscr{B}_+(\mathbb{R}_+)$. Notice that (1.7.19) can be rewritten as

$$\xi_n(h) \xrightarrow[n\to\infty]{(d)+\text{mean}} \xi(h) \tag{1.7.21}$$

for every $h\in\mathscr{C}_+(\mathbb{R}_+)$ nonincreasing, where the measure ξ is defined by $\xi(h) = \Psi^{\mathfrak{mh}}_{\mathscr{T}}(x^\alpha h(u))$. Moreover, Lemma 1.5.1-(ii) applied with $h\equiv 1$ gives $\sup_{n\in\Delta}\mathbb{E}[\xi_n(1)] < \infty$. As a consequence, by the Markov inequality, we have $\lim_{r\to\infty}\sup_{n\in\Delta}\mathbb{P}(\xi_n(1) > r) = 0$. Fix $\beta > 0$ and let $r > 0$. Then, using the inequality $\mathbf{1}_{[r,\infty)}(u) \le (u/r)^\beta$ for every $u\ge 0$, we get

$$\sup_{n\in\Delta}\mathbb{E}[\xi_n([r,\infty))] \le \frac{1}{r^\beta}\sup_{n\in\Delta}\mathbb{E}\left[\mathscr{A}^{\mathfrak{mh},\circ}_n(x^\alpha u^\beta)\right].$$

Notice that the right-hand side is finite by Lemma 1.5.2 since $\gamma\alpha + (\gamma-1)(\beta+1) > 0$. We deduce that

$$\inf_{K\subset\mathbb{R}_+} \sup_{n\in\Delta} \mathbb{E}\left[\xi_n(K^c)\right] = 0,$$

where the infimum is taken over all compact subsets $K \subset \mathbb{R}_+$. By [99, Theorem 4.10], this means that the sequence of random measures $(\xi_n,\, n\in\Delta)$ is tight in distribution in $\mathscr{M}(\mathbb{R}_+)$. Following the end of step 2 for property (i), we are then able to show that ξ_n converges to ξ in distribution in $\mathscr{M}([0,\infty))$ and $\mathbb{E}[\xi_n(\bullet)]$ converges to $\mathbb{E}[\xi(\bullet)]$ in $\mathscr{M}([0,\infty))$.

Step 3 in the proof of (ii). Let $h \in \mathscr{B}(\mathbb{R}_+)$ be continuous on $(0,\infty)$ such that h^* is non-zero, $\int_0 h^*(u) u^{\alpha\gamma/(\gamma-1)}\,\mathrm{d}u < \infty$ and $h(u) = O(\mathrm{e}^{u^\eta})$ as $u \to \infty$ for some $\eta \in (0,\gamma)$. Set $h_0 \equiv 1$ and define a positive function $h_1 \in \mathscr{B}_+((0,\infty))$ in the following way. If $\lim_{u\to 0} h^*(u) = \infty$, set $h_1 = h^* + 1$ on $(0,1]$. If h^* has a finite limit at 0 (which is positive as h^* is non-zero), then $\alpha > -1 + 1/\gamma$, and thus there exists $\varepsilon > 0$ such that $\int_0 u^{-\varepsilon} h^*(u) u^{\alpha\gamma/(\gamma-1)}\,\mathrm{d}u < \infty$. Moreover, we have $\lim_{u\to 0} u^{-\varepsilon} h^*(u) = \infty$ and the function $u \mapsto u^{-\varepsilon} h^*(u)$ is continuous and nonincreasing. In that case, we set $h_1(u) = u^{-\varepsilon} h^*(u) + 1$ for $u \in (0,1]$. Now extend h_1 to a continuous function on $(0,\infty)$ such that $h_1(u) = \exp(u^{\eta_1})$ for $u \geq 2$ for some $\eta_1 \in (\eta,\gamma)$. Define a set of functions $\mathfrak{F} = \{h_k\colon\ k \in \mathbb{N}\}$ as follows. Let $(\eta_k,\, k \geq 2)$ be an increasing sequence in (η_1,γ). Recall that $\alpha > -1 + 1/\gamma$ so that the measure $\nu(\mathrm{d}u) = \mathbf{1}_{(0,1]}(u)\, u^{\alpha\gamma/(\gamma-1)}\mathrm{d}u$ is finite. For every $k \geq 1$, define $h_{k+1} \in \mathscr{B}_+([0,\infty))$ continuous and positive on $(0,\infty)$ and such that $h_{k+1} = h_k^\nu$ on $(0,1]$, with h_k^ν defined in Lemma 1.7.2, and $h_{k+1}(u) = \exp(u^{\eta_{k+1}})$ for $u \geq 2$. In particular, we have $\lim_{x\to 0+} h_k(x)/h_{k+1}(x) = \lim_{x\to+\infty} h_k(x)/h_{k+1}(x) = 0$. Then, it is easy to check that the sequence $\mathfrak{F}$ satisfies assumptions (H1)–(H5) of Appendix 1.A with $S = \mathbb{R}_+$, $S_0 = \{0\}$ and $\mathfrak{F}^\star(h_k) = \{h_j\colon\ j > k\}$ for $k \in \mathbb{N}$. Notice that, by Lemma 1.7.2, for every $k \in \mathbb{N}$, the function h_k is continuous and nonincreasing on $(0,1]$ and satisfies $\int_0 h_k(u) u^{\alpha\gamma/(\gamma-1)}\,\mathrm{d}u < \infty$. So, by Lemma 1.5.1 (i) and (ii), we get that for all $k \in \mathbb{N}$, there exists a finite constant $C_k > 0$ such that

$$\sup_{n\in\Delta} \mathbb{E}[\xi_n(h_k)] \leq \sup_{n\in\Delta} \mathbb{E}\left[\xi_n(h_k \mathbf{1}_{(0,1]})\right] + C_k \sup_{n\in\Delta} \mathbb{E}\left[\xi_n(\exp(u^{\eta_k})\mathbf{1}_{\{u\geq 1\}})\right] < \infty.$$

Now, Corollary 1.A.11 applies and yields, in conjunction with Proposition 1.A.9, the convergence in distribution and of the first moment

$$\xi_n(h_k f) \xrightarrow[n\to\infty]{(d)+\text{mean}} \xi(h_k f)$$

for every $k \in \mathbb{N}$ and every $f \in \mathscr{C}(\mathbb{R}_+)$. Taking $k = 1$ and $f = h/h_1$ proves (1.7.12) as $\xi_n(S_0) = \xi(S_0) = 0$. This, together with Proposition 1.6.9, proves (ii).

To prove (iii), notice that by (1.4.23) we have the convergence in distribution $\mathscr{A}_n^{\mathfrak{m}\mathfrak{h},\circ} \xrightarrow{(d)} \Psi_{\mathscr{T}}^{\mathfrak{m}\mathfrak{h}}$ in the space $\mathscr{M}([0,1]\times\mathbb{R}_+)$. Thanks to Skorokhod's representation theorem, we may assume that we have a.s. convergence. Thus, we get that a.s. for every $k \in \mathbb{N}$,

$$\lim_{n\to\infty} \frac{b_n}{n^2} \sum_{w\in\tau^{n,\circ}} |\tau_w^n| \left(f\left(\frac{|\tau_w^n|}{n}, \frac{b_n}{n}\mathfrak{h}(\tau_w^n)\right) \wedge k \right) = \Psi_{\mathscr{T}}^{\mathfrak{m}\mathfrak{h}}(f \wedge k).$$

Therefore, we have for $k \in \mathbb{N}$

$$\liminf_{n\to\infty} \frac{b_n}{n^2} \sum_{w\in\tau^{n,\circ}} |\tau^n_w| f\left(\frac{|\tau^n_w|}{n}, \frac{b_n}{n}\mathfrak{h}(\tau^n_w)\right) \geq \Psi^{\mathfrak{mh}}_{\mathcal{T}}(f \wedge k). \tag{1.7.22}$$

But by the monotone convergence theorem and Proposition 1.6.9, we have that a.s. $\lim_{k\to\infty} \Psi^{\mathfrak{mh}}_{\mathcal{T}}(f \wedge k) = \Psi^{\mathfrak{mh}}_{\mathcal{T}}(f) = \infty$. Thus, (1.7.14) follows from (1.7.22) by letting k go to infinity. □

Recall from (1.4.16) that we excluded the leaves to be able to consider functions taking infinite values on trees whose height vanishes. In the particular case where the function only blows up as the mass goes to zero, one can get rid of this restriction.

Remark 1.7.4. Recall the definition of the random measure $\mathscr{A}^{\mathfrak{mh},\circ}_n \in \mathcal{M}([0,1]\times\mathbb{R}_+)$:

$$\mathscr{A}^{\mathfrak{mh},\circ}_n(f) = \frac{b_n}{n^2} \sum_{w\in\tau^{n,\circ}} |\tau^n_w| f\left(\frac{|\tau^n_w|}{n}, \frac{b_n}{n}\mathfrak{h}(\tau^n_w)\right).$$

Similarly to the measure $\mathscr{A}^{\mathfrak{mh},\circ}_n$, we define the measure $\mathscr{A}^{\mathfrak{mh}}_n \in \mathcal{M}([0,1]\times\mathbb{R}_+)$, where the sum is over all the vertices (the internal vertices and the leaves): for $f \in \mathscr{B}_+([0,1]\times\mathbb{R}_+)$

$$\mathscr{A}^{\mathfrak{mh}}_n(f) = \frac{b_n}{n^2} \sum_{w\in\tau^{n}} |\tau^n_w| f\left(\frac{|\tau^n_w|}{n}, \frac{b_n}{n}\mathfrak{h}(\tau^n_w)\right).$$

Let $\beta \geq 0$ and $g \in \mathscr{B}([0,1])$ such that g is continuous on $(0,1]$ and $\int_0 g^*(x^{\gamma/(\gamma-1)})x^\beta\,\mathrm{d}x < \infty$. By Theorem 1.7.3-(i), we have

$$\mathscr{A}^{\mathfrak{mh},\circ}_n(g(x)u^\beta) \xrightarrow[n\to\infty]{(d)+\text{mean}} \Psi^{\mathfrak{mh}}_{\mathcal{T}}(g(x)u^\beta). \tag{1.7.23}$$

Now note that

$$\mathscr{A}^{\mathfrak{mh}}_n(g(x)u^\beta) = \frac{b_n^{1+\beta}}{n^{2+\beta}} \sum_{w\in\tau^n} |\tau^n_w| \mathfrak{h}(\tau^n_w)^\beta g\left(\frac{|\tau^n_w|}{n}\right)$$

makes sense when the function g blows up at 0. If $\beta > 0$, we have $\mathscr{A}^{\mathfrak{mh}}_n(g(x)u^\beta) = \mathscr{A}^{\mathfrak{mh},\circ}_n(g(x)u^\beta)$ since $\mathfrak{h}(\tau^n_w) = 0$ for every leaf $w \in \mathrm{Lf}(\tau^n)$. Thus we only need to consider the case $\beta = 0$. Then, using (1.4.2) and the fact that $|\mathrm{Lf}(\tau^n)| \leq n$ and that $|\tau^n_w| = 1$ for every $w \in \mathrm{Lf}(\tau^n)$, we have

$$\left|\mathscr{A}^{\mathfrak{mh}}_n(g(x)) - \mathscr{A}^{\mathfrak{mh},\circ}_n(g(x))\right| = \frac{b_n}{n^2}\left|\sum_{w\in\mathrm{Lf}(\tau^n)} |\tau^n_w| g\left(\frac{|\tau^n_w|}{n}\right)\right| \leq \bar{b} n^{-1+1/\gamma} g^*\left(\frac{1}{n}\right).$$

Since g^* is nonincreasing and satisfies $\int_0 g^*(x^{\gamma/(\gamma-1)})\,\mathrm{d}x < \infty$, it is straightforward to check that $g^*(x) = o(x^{1/\gamma-1})$ as $x \to 0$. Thus, we deduce that $\lim_{n\to\infty} \mathscr{A}^{\mathfrak{mh}}_n(g(x)u^\beta) - \mathscr{A}^{\mathfrak{mh},\circ}_n(g(x)u^\beta) = 0$ a.s. and in $L^1(\mathbb{P})$. As a consequence, the convergence (1.7.23) still holds if we replace $\mathscr{A}^{\mathfrak{mh},\circ}_n(g(x)u^\beta)$ by $\mathscr{A}^{\mathfrak{mh}}_n(g(x)u^\beta)$.

Similarly, let $\alpha > -1 + 1/\gamma$ and $h \in \mathscr{C}(\mathbb{R}_+)$ such that $h(u) = O(\mathrm{e}^{u^\eta})$ as $u \to \infty$ for some $\eta \in (0,\gamma)$. Then h^* is bounded near 0 and necessarily $\int_0 x^{\alpha\gamma/(\gamma-1)} h^*(x)\,\mathrm{d}x < \infty$. Thus, by Theorem 1.7.3, we have

$$\mathscr{A}_n^{\mathfrak{mh},\circ}(x^\alpha h(u)) \xrightarrow[n\to\infty]{(d)+\text{mean}} \Psi_{\mathscr{T}}^{\mathfrak{mh}}(x^\alpha h(u)). \tag{1.7.24}$$

Furthermore, using (1.4.2) we have

$$\left|\mathscr{A}_n^{\mathfrak{mh}}(x^\alpha h(u)) - \mathscr{A}_n^{\mathfrak{mh},\circ}(x^\alpha h(u))\right| = \frac{b_n}{n^{2+\alpha}}\left|\mathrm{Lf}(\tau^n)\right||h(0)| \le \overline{b}n^{-\alpha-1+1/\gamma}|h(0)|.$$

Thus, we deduce that $\lim_{n\to\infty}\mathscr{A}_n^{\mathfrak{mh}}(x^\alpha h(u)) - \mathscr{A}_n^{\mathfrak{mh},\circ}(x^\alpha h(u)) = 0$ a.s. and in $L^1(\mathbb{P})$ and the convergence (1.7.24) holds for $\mathscr{A}_n^{\mathfrak{mh}}(x^\alpha h(u))$.

Example 1.7.5. Fix $\alpha > -1+1/\gamma$ and set $g(x) = |\log(x)|x^\alpha$. It is clear that $\int_0 g(x^{\gamma/(\gamma-1)})\,\mathrm{d}x < \infty$, so by Theorem 1.7.3 we have the convergence in distribution

$$\mathscr{A}_n^{\mathfrak{mh},\circ}(g(x)) \xrightarrow[n\to\infty]{(d)} \Psi_{\mathscr{T}}^{\mathfrak{mh}}(g(x)).$$

But notice that

$$\begin{aligned}\mathscr{A}_n^{\mathfrak{mh},\circ}(g(x)) &= \frac{b_n\log(n)}{n^{2+\alpha}}\sum_{w\in\tau^{n,\circ}}\left|\tau_w^n\right|^{1+\alpha} - \frac{b_n}{n^{2+\alpha}}\sum_{w\in\tau^{n,\circ}}\left|\tau_w^n\right|^{1+\alpha}\log\left|\tau_w^n\right| \\ &= \log(n)\,\mathscr{A}_n^{\mathfrak{mh},\circ}(x^\alpha) - \frac{b_n}{n^{2+\alpha}}\sum_{w\in\tau^{n,\circ}}\left|\tau_w^n\right|^{1+\alpha}\log\left|\tau_w^n\right|.\end{aligned}$$

Again Theorem 1.7.3 gives the convergence in distribution $\mathscr{A}_n^{\mathfrak{mh},\circ}(x^\alpha) \xrightarrow{(d)} \Psi_{\mathscr{T}}^{\mathfrak{mh}}(x^\alpha)$. Therefore, we get the following asymptotic expansion in distribution

$$\frac{b_n}{n^{2+\alpha}}\sum_{w\in\tau^{n,\circ}}\left|\tau_w^n\right|^{1+\alpha}\log\left|\tau_w^n\right| \overset{(d)}{=} \log(n)\,\Psi_{\mathscr{T}}^{\mathfrak{mh}}(x^\alpha) - \Psi_{\mathscr{T}}^{\mathfrak{mh}}(|\log(x)|x^\alpha) + o(1).$$

Furthermore, since

$$\lim_{n\to\infty}\mathbb{E}\left[\mathscr{A}_n^{\mathfrak{mh},\circ}(g(x))\right] = \mathbb{E}\left[\Psi_{\mathscr{T}}^{\mathfrak{mh}}(g(x))\right] \quad\text{and}\quad \lim_{n\to\infty}\mathbb{E}\left[\mathscr{A}_n^{\mathfrak{mh},\circ}(x^\alpha)\right] = \mathbb{E}\left[\Psi_{\mathscr{T}}^{\mathfrak{mh}}(x^\alpha)\right],$$

we get the corresponding asymptotic expansion for the first moment

$$\frac{b_n}{n^{2+\alpha}}\mathbb{E}\left[\sum_{w\in\tau^{n,\circ}}\left|\tau_w^n\right|^{1+\alpha}\log\left|\tau_w^n\right|\right] = \log(n)\,\mathbb{E}\left[\Psi_{\mathscr{T}}^{\mathfrak{mh}}(x^\alpha)\right] - \mathbb{E}\left[\Psi_{\mathscr{T}}^{\mathfrak{mh}}(|\log(x)|x^\alpha)\right] + o(1).$$

1.A A space of measures

Let (S,ρ) be a Polish metric space, $S_0 \subset S$ be a closed set in S and $0 \in S_0$ be a distinguished point. Denote by $\mathscr{K}$ the class of compact sets $K \subset S$. For any $x \in S$ and $A \subset S$, the distance from x to A is defined by $\rho(x,A) = \inf\{\rho(x,y) : y \in A\}$. Let $\mathfrak{F}$ be a countable set of measurable $[0,+\infty]$-valued functions on S satisfying the following assumptions:

(H1) The constant function $\mathbf{1}$ belongs to $\mathfrak{F}$.

(H2) All $f \in \mathfrak{F}$ are continuous on S_0^c.

(H3) All $f \in \mathfrak{F}$ are bounded away from zero and infinity on $\{x \in S\colon \rho(x,S_0) \geq \varepsilon, \rho(x,0) \leq M\}$ for every $0 < \varepsilon < M < +\infty$.

(H4) For all $f \in \mathfrak{F}$, the set $\mathfrak{F}^\star(f) \subset \mathfrak{F}$ of functions $f^\star \in \mathfrak{F}$ such that $f/f^\star$ is bounded on S_0^c and $\lim_{\rho(x,S_0)\to 0+} f(x)/f^\star(x) = 0$ is non-empty.

Note that assumption (H3) is automatically satisfied when S is compact and every $f \in \mathfrak{F}$ is positive on S_0^c. Notice that (H4) implies that $\mathfrak{F}^\star(f)$ is countably infinite for any $f \in \mathfrak{F}$. We shall write $f^\star$ for any element of $\mathfrak{F}^\star(f)$. By (H1) and (H4), we have $\lim_{\rho(x,S_0)\to 0+} \mathbf{1}^\star(x) = +\infty$. By convention, we take $\mathbf{1}^\star \equiv +\infty$ on S_0 and $f/f^\star \equiv 0$ on S_0 for every $f \in \mathfrak{F}$. We will occasionally need the following additional assumption:

(H5) S is compact or $\inf_{K\in\mathcal{K}} \sup_{x\in K^c} f(x)/f^\star(x) = 0$ for every $f \in \mathfrak{F}$ (and some $f^\star \in \mathfrak{F}^\star(f)$).

Denote by $\mathcal{M} = \mathcal{M}(S)$ the space of nonnegative finite measures on S endowed with the weak topology. Recall that $(\mathcal{M}, d_{\mathrm{BL}})$, with d_{BL} the bounded Lipschitz distance is a Polish metric space. Recall that, for $\mu \in \mathcal{M}$ and $f \in \mathcal{B}_+(S)$, the notation $f\mu$ stands for the measure $f(x)\mu(\mathrm{d}x)$. Set

$$\mathcal{M}_{\mathfrak{F}} = \mathcal{M}_{\mathfrak{F}}(S) := \left\{\mu \in \mathcal{M}\colon\ \mu(f) < \infty \text{ for all } f \in \mathfrak{F}\right\}. \tag{1.A.1}$$

For $\mu \in \mathcal{M}_{\mathfrak{F}}$, we have $\mu(S_0) = 0$ (as $\mathbf{1}^\star \equiv +\infty$ on S_0) and $f\mu \in \mathcal{M}$ for every $f \in \mathfrak{F}$. In particular, since $(f/f^\star)f^\star = f$ on S_0^c, we have $(f/f^\star)f^\star\mu = f\mu$ for every $f \in \mathfrak{F}$ (and $f^\star \in \mathfrak{F}^\star(f)$). We say a sequence $(\mu_n, n \in \mathbb{N})$ of elements of $\mathcal{M}_{\mathfrak{F}}$ converges to $\mu \in \mathcal{M}_{\mathfrak{F}}$ if and only if $(f\mu_n, n \in \mathbb{N})$ converges to $f\mu$ in $\mathcal{M}$ for every $f \in \mathfrak{F}$. We consider the following distance $d_{\mathfrak{F}}$ on $\mathcal{M}_{\mathfrak{F}}$ which defines the same topology:

$$d_{\mathfrak{F}}(\mu,\nu) = \sum_{k\in\mathbb{N}} \frac{1}{2^k}\left(1 \wedge d_{\mathrm{BL}}\left(f_k\mu, f_k\nu\right)\right) \quad \text{for} \quad \mu,\nu \in \mathcal{M}_{\mathfrak{F}}, \tag{1.A.2}$$

where $\{f_k\colon\ k \in \mathbb{N}\}$ is an enumeration of $\mathfrak{F}$. (The choice of the enumeration is unimportant, as the corresponding distances all define the same topology on $\mathcal{M}_{\mathfrak{F}}$.) Notice that the mapping $\mu \mapsto f\mu$ is continuous from $\mathcal{M}_{\mathfrak{F}}$ to $\mathcal{M}$. In particular, taking $f = \mathbf{1}$ gives that every sequence which converges in $\mathcal{M}_{\mathfrak{F}}$ also converges in $\mathcal{M}$ to the same limit.

We shall see that the space $(\mathcal{M}_{\mathfrak{F}}, d_{\mathfrak{F}})$ is complete and separable (Proposition 1.A.1) and give a complete description of its compact subsets (Proposition 1.A.2). The main goal of this section is to give conditions which allow to strengthen a convergence in $\mathcal{M}$ to a convergence in $\mathcal{M}_{\mathfrak{F}}$ for deterministic measures (Corollary 1.A.3) and then to extend this result to random measures (Proposition 1.A.10 and Corollary 1.A.11).

Proposition 1.A.1. *The space $(\mathcal{M}_{\mathfrak{F}}, d_{\mathfrak{F}})$ is complete and separable.*

Proof. Let $(\mu_n, n \in \mathbb{N})$ be a Cauchy sequence in $\mathcal{M}_{\mathfrak{F}}$. Then, by definition of $d_{\mathfrak{F}}$, the sequence $(f\mu_n, n \in \mathbb{N})$ is Cauchy in $\mathcal{M}$ for every $f \in \mathfrak{F}$. By completeness of $\mathcal{M}$, for every $f \in \mathfrak{F}$, there exists a measure $\nu_f \in \mathscr{M}$ such that $\lim_{n\to\infty} f\mu_n = \nu_f$ in $\mathscr{M}$. We claim that $\nu_f(S_0) = 0$ for every $f \in \mathfrak{F}$. Indeed, fix $f \in \mathfrak{F}$ and $f^\star \in \mathfrak{F}^\star(f)$. As $f^\star \in \mathfrak{F}$, we have $\lim_{n\to\infty} f^\star\mu_n = \nu_{f^\star}$ in $\mathscr{M}$. By (H4), the function $f/f^\star$ is continuous and bounded on S, so that the mapping $\pi \mapsto (f/f^\star)\pi$ is continuous on $\mathscr{M}$. In particular, we have $\lim_{n\to\infty} f\mu_n = (f/f^\star)\nu_{f^\star}$ in $\mathscr{M}$. On the other hand, we have $\lim_{n\to\infty} f\mu_n = \nu_f$ in $\mathscr{M}$. We deduce that $\nu_f = (f/f^\star)\nu_{f^\star}$. It follows that $\nu_f(S_0) = 0$ since $f/f^\star = 0$ on S_0.

We set $\mu = \nu_{\mathbf{1}}$ so that $\lim_{n\to\infty}\mu_n = \mu$ in $\mathscr{M}$. Let $f \in \mathfrak{F}$. We shall prove that $f\mu = \nu_f$. Consider the closed set $F_k = \{f \geq 1/k\}$ for $k \in \mathbb{N}^*$. Notice that $F_k \subset \mathrm{int}(F_{k+1})$. Therefore, by Urysohn's lemma, there exists, for $k \in \mathbb{N}^*$, a continuous function $\chi_k \colon S \to [0,1]$ such that $\chi_k = 1$ on F_k and $\mathrm{supp}(\chi_k) \subset \mathrm{int}(F_{k+1})$. Notice that $(\chi_k f/f)\mu_n = \chi_k\mu_n$ since $(f/f) = 1$ on S_0^c and $\mu_n(S_0) = 0$. Since χ_k and χ_k/f are continuous and bounded, the mappings $\nu \mapsto \chi_k\nu$ and $\nu \mapsto (\chi_k/f)\nu$ are continuous from $\mathscr{M}$ to itself. We deduce that $\chi_k\mu = \lim_{n\to\infty}\chi_k\mu_n = \lim_{n\to\infty}(\chi_k/f)f\mu_n = (\chi_k/f)\nu_f$ in $\mathscr{M}$. Letting k go to infinity, as $\chi_k \uparrow \mathbf{1}$ on S_0^c since f is positive on S_0^c, and $\mu(S_0) = \nu_f(S_0) = 0$, we deduce (using the monotone convergence theorem) that $\mu = (1/f)\nu_f$ and thus $f\mu = \nu_f$. Since this holds for all $f \in \mathfrak{F}$, this proves that $\mu \in \mathscr{M}_{\mathfrak{F}}$ and that $\lim_{n\to\infty} f\mu_n = f\mu$ in $\mathscr{M}$ for every $f \in \mathfrak{F}$. Thus $\mathscr{M}_{\mathfrak{F}}$ is complete.

Next, define $F'_n = \{x \in S \colon \rho(x, S_0) \geq 1/n,\ \rho(x,0) \leq n\}$. We will identify the space $\mathscr{M}(F'_n)$ with the subset of $\mathscr{M}$ consisting of the measures whose support lies in F'_n. Notice that F'_n is a Polish space (when endowed with the topology induced by ρ) as a closed subset of the Polish space S. In particular, the set $\mathscr{M}(F'_n)$ endowed with the bounded Lipschitz distance is a Polish space. Let $f \in \mathfrak{F}$. By (H3), the functions f and $1/f$ are both continuous and bounded on F'_n, so it is easy to check that the topology induced by $d_{\mathfrak{F}}$ on $\mathscr{M}(F'_n)$ coincides with the topology of weak convergence, i.e. the one induced by d_{BL}. Therefore, the space $(\mathscr{M}(F'_n), d_{\mathfrak{F}})$ is separable. To prove that $\mathscr{M}_{\mathfrak{F}}$ is separable, it suffices to show that $\mathscr{M}_{\mathfrak{F}}$ is equal to the completion of $\bigcup_{n\geq 1}\mathscr{M}(F'_n)$ with respect to d_{BL}. Notice that $F'_n \subset \mathrm{int}(F'_{n+1})$. Therefore, by Urysohn's lemma, there exists a continuous function $\chi'_n \colon S \to [0,1]$ such that $\chi'_n = 1$ on F'_n and $\mathrm{supp}(\chi'_n) \subset \mathrm{int}(F'_{n+1})$. Let $\mu \in \mathscr{M}_{\mathfrak{F}}$ and set $\mu_n = \chi'_n\mu$. Then it is clear that μ_n has support in F'_{n+1} and thus $\mu_n \in \mathscr{M}(F'_{n+1})$. Moreover, for every $f \in \mathfrak{F}$ and every nonnegative $h \in \mathscr{C}_b(S)$, we have

$$\mu_n(hf) = \mu(hf\chi'_n) \xrightarrow[n\to\infty]{} \mu(hf)$$

by the monotone convergence theorem, since $\chi'_n \uparrow \mathbf{1}_{S_0^c}$ and $\mu(S_0) = 0$. This proves that $(f\mu_n, n \in \mathbb{N})$ converges to $f\mu$ in $\mathscr{M}$ for every $f \in \mathfrak{F}$, thus $d_{\mathfrak{F}}(\mu_n, \mu) \to 0$. This concludes the proof. $\square$

A set of measures $A \subset \mathscr{M}$ is said to be bounded if $\sup_{\mu\in A}\mu(\mathbf{1}) < \infty$. We now give a characterization of compactness in $\mathscr{M}_{\mathfrak{F}}$.

Proposition 1.A.2. *Let $A \subset \mathscr{M}_{\mathfrak{F}}$.*

(i) A is relatively compact if and only if for every $f \in \mathfrak{F}$, the family $\{f\mu \colon \mu \in A\}$ of finite measures is bounded and tight.

(ii) If (H5) holds, then A is relatively compact if and only if for every $f \in \mathfrak{F}$, the family $\{f\mu\colon \mu \in A\}$ is bounded.

Proof. To prove (i), start by assuming that A is relatively compact. For every $\mu \in \mathscr{M}_{\mathfrak{F}}$ and every $f \in \mathfrak{F}$, set $F_f(\mu) = f\mu$. This defines a continuous mapping $F_f\colon \mathscr{M}_{\mathfrak{F}} \to \mathscr{M}$. It follows that the set

$$F_f(A) = \{f\mu\colon \mu \in A\}$$

is relatively compact in $\mathscr{M}$, i.e. it is bounded and tight by Prokhorov's theorem.

Conversely, let us assume that $\{f\mu\colon \mu \in A\}$ is bounded and tight in $\mathscr{M}$ for all $f \in \mathfrak{F}$. Let $(\mu_n, n \in \mathbb{N})$ be a sequence in A. Since the sequence of measures $(f\mu_n, n \in \mathbb{N})$ is bounded and tight, it is relatively compact in $\mathscr{M}$ for every $f \in \mathfrak{F}$. Therefore, by diagonal extraction, there exists a subsequence still denoted by $(f\mu_n, n \in \mathbb{N})$ which converges in $\mathscr{M}$ for every $f \in \mathfrak{F}$. By the same argument as in the proof of Proposition 1.A.1, it follows that $(\mu_n, n \in \mathbb{N})$ converges in $\mathscr{M}_{\mathfrak{F}}$. This proves that A is relatively compact.

To prove (ii), assume that (H5) holds. The statement for a compact S follows immediately since a family of finite measures on a compact space is always tight. Now assume that S is not compact and let $A \subset \mathscr{M}_{\mathfrak{F}}$ such that the family $\{f\mu\colon \mu \in A\}$ is bounded for every $f \in \mathfrak{F}$. To prove that $A \subset \mathscr{M}_{\mathfrak{F}}$ is relatively compact, it is enough to show that $\{f\mu\colon \mu \in A\}$ is tight and to apply the first point. Let $f^\star \in \mathfrak{F}^\star(f)$ be the one appearing in (H5) and $K \subset S$ be a compact subset. For every $\mu \in A$, since $\mu(S_0) = 0$, we have

$$\begin{aligned}\int_{K^c} f(x)\,\mu(\mathrm{d}x) &= \int_{K^c} f(x)\mathbf{1}_{S_0^c}(x)\,\mu(\mathrm{d}x)\\ &= \int_{K^c} \frac{f(x)}{f^\star(x)}\mathbf{1}_{S_0^c}(x) f^\star(x)\,\mu(\mathrm{d}x)\\ &\le \mu(f^\star)\sup_{K^c}\frac{f}{f^\star}.\end{aligned}$$

It follows that

$$\sup_{\mu\in A}\int_{K^c} f(x)\,\mu(\mathrm{d}x) \le \sup_{\mu\in A}\mu(f^\star)\sup_{K^c}\frac{f}{f^\star},$$

and taking the infimum over all compact subsets $K \in \mathscr{K}$ yields, thanks to (H5)

$$\inf_{K\in\mathscr{K}}\sup_{\mu\in A}\int_{K^c} f(x)\,\mu(\mathrm{d}x) = 0,$$

i.e. the family $\{f\mu\colon \mu \in A\}$ is tight. This completes the proof. □

The next result gives sufficient conditions allowing to strengthen convergence in $\mathscr{M}$ to convergence in $\mathscr{M}_{\mathfrak{F}}$.

Corollary 1.A.3. *Let $(\mu_n, n \in \mathbb{N})$ be a sequence of elements of $\mathscr{M}_{\mathfrak{F}}$ converging in $\mathscr{M}$ to some $\mu \in \mathscr{M}$. Then $\mu \in \mathscr{M}_{\mathfrak{F}}$ and $\lim_{n\to\infty}\mu_n = \mu$ in $\mathscr{M}_{\mathfrak{F}}$ under either of the following conditions:*

(i) *$(f\mu_n, n \in \mathbb{N})$ is bounded and tight for every $f \in \mathfrak{F}$.*

(ii) *(H5) holds and $(f\mu_n, n \in \mathbb{N})$ is bounded for every $f \in \mathfrak{F}$.*

Proof. Either condition guarantees that the sequence $(\mu_n, n \in \mathbb{N})$ is relatively compact in $\mathcal{M}_{\mathfrak{F}}$ by Proposition 1.A.2. Let $\hat{\mu} \in \mathcal{M}_{\mathfrak{F}}$ be a limit point of $(\mu_n, n \in \mathbb{N})$. Then there exists a subsequence, still denoted by $(\mu_n, n \in \mathbb{N})$ such that $\lim_{n\to\infty} \mu_n = \hat{\mu}$ in $\mathcal{M}_{\mathfrak{F}}$. In particular, we have $\lim_{n\to\infty} \mu_n = \hat{\mu}$ in $\mathcal{M}$. Since $\lim_{n\to\infty} \mu_n = \mu$ in $\mathcal{M}$ by assumption, it follows that $\hat{\mu} = \mu$. This proves that $\mu \in \mathcal{M}_{\mathfrak{F}}$ and that $\lim_{n\to\infty} \mu_n = \mu$ in $\mathcal{M}_{\mathfrak{F}}$ since the sequence $(\mu_n, n \in \mathbb{N})$ is relatively compact in $\mathcal{M}_{\mathfrak{F}}$ and has only one limit point μ. □

The compactness criterion of Proposition 1.A.2 yields a tightness criterion for random measures in $\mathcal{M}_{\mathfrak{F}}$.

Proposition 1.A.4. *Let Ξ be a family of $\mathcal{M}_{\mathfrak{F}}$-valued random variables.*

(i) *The family Ξ is tight (in distribution) in $\mathcal{M}_{\mathfrak{F}}$ if and only if for every $f \in \mathfrak{F}$, the family $\{f\xi\colon \xi \in \Xi\}$ is tight (in distribution) in $\mathcal{M}$, i.e. if and only if*

$$\lim_{r\to\infty} \sup_{\xi\in\Xi} \mathbb{P}\big(\xi(f) > r\big) = 0 \tag{1.A.3}$$

and

$$\inf_{K\in\mathcal{K}} \sup_{\xi\in\Xi} \mathbb{E}\left[1 \wedge \int_{K^c} f(x)\xi(\mathrm{d}x)\right] = 0. \tag{1.A.4}$$

(ii) *If (H5) holds, then Ξ is tight (in distribution) in $\mathcal{M}_{\mathfrak{F}}$ if and only if (1.A.3) holds for every $f \in \mathfrak{F}$.*

Proof. To prove (i), assume that Ξ is tight in $\mathcal{M}_{\mathfrak{F}}$. Since the mapping $F_f\colon \mu \mapsto f\mu$ is continuous from $\mathcal{M}_{\mathfrak{F}}$ to $\mathcal{M}$ for every $f \in \mathfrak{F}$ and since tightness is preserved by continuous mappings, it follows that the family $F_f(\Xi) = \{f\xi\colon \xi \in \Xi\}$ is tight in $\mathcal{M}$ for every $f \in \mathfrak{F}$. The result now follows from Theorem 4.10 in [99].

Conversely, assume that (1.A.3) and (1.A.4) hold for all $f \in \mathfrak{F}$ and let $\varepsilon > 0$. Let $\{f_k\colon k \in \mathbb{N}^*\}$ be an enumeration of $\mathfrak{F}$. We set for $k \in \mathbb{N}^*$:

$$C_k = k\left(1 + \sup_{j\le k,\, f_k\in\mathfrak{F}^\star(f_j)} \|f_j/f_k\|_\infty\right),$$

with the convention that $\sup\emptyset = 0$. For every $k \in \mathbb{N}^*$, there exists $r_k > 0$ and a compact set $K_k \in \mathcal{K}$ such that

$$\sup_{\xi\in\Xi} \mathbb{P}\big(\xi(f_k) > r_k\big) \le \frac{\varepsilon}{2^k} \quad \text{and} \quad \sup_{\xi\in\Xi} \mathbb{E}\left[1 \wedge \int_{K_k^c} f_k(x)\xi(\mathrm{d}x)\right] \le \frac{\varepsilon}{C_k 2^k}.$$

Set

$$A_\varepsilon = \bigcap_{k\in\mathbb{N}^*} \left\{ \mu \in \mathscr{M}_{\mathfrak{F}} \colon \ \mu(f_k) \le r_k \text{ and } \int_{K_k^c} f_k(x)\mu(\mathrm{d}x) \le \frac{1}{C_k} \right\}.$$

Then for every $\xi \in \Xi$, we have

$$\begin{aligned} \mathbb{P}\left(\xi \in A_\varepsilon^c\right) &= \mathbb{P}\left(\exists k \in \mathbb{N}^*, \xi(f_k) > r_k \text{ or } \int_{K_k^c} f_k(x)\xi(\mathrm{d}x) > \frac{1}{C_k}\right) \\ &\le \sum_{k\in\mathbb{N}^*} \mathbb{P}\left(\xi(f_k) > r_k\right) + \sum_{k\in\mathbb{N}^*} \mathbb{P}\left(\int_{K_k^c} f_k(x)\xi(\mathrm{d}x) > \frac{1}{C_k}\right) \le 2\varepsilon, \end{aligned}$$

where in the last inequality we used that

$$\mathbb{P}\left(\int_{K_k^c} f(x)\xi(\mathrm{d}x) > \frac{1}{C_k}\right) = \mathbb{P}\left(1 \wedge \int_{K_k^c} f_k(x)\xi(\mathrm{d}x) > \frac{1}{C_k}\right) \le C_k \mathbb{E}\left[1 \wedge \int_{K_k^c} f_k(x)\xi(\mathrm{d}x)\right] \le \frac{\varepsilon}{2^k}.$$

Thus, to prove that Ξ is tight in $\mathscr{M}_{\mathfrak{F}}$, it remains to show that $A_\varepsilon \subset \mathscr{M}_{\mathfrak{F}}$ is relatively compact. We have $\sup_{\mu\in A_\varepsilon} \mu(f_k) \le r_k < \infty$ so that the family $\{f_k\mu \colon \ \mu \in A_\varepsilon\}$ is bounded for every $k \in \mathbb{N}^*$. Moreover, for every $i \ge k$ such that $f_i \in \mathfrak{F}^\star(f_k)$, we have

$$\sup_{\mu\in A_\varepsilon} \int_{K_i^c} f_k(x)\mu(\mathrm{d}x) \le \left\| f_k/f_i \right\|_\infty \sup_{\mu\in A_\varepsilon} \int_{K_i^c} f_i(x)\mu(\mathrm{d}x) \le \frac{1}{i}.$$

This implies that $\inf_{K\in\mathscr{K}} \sup_{\mu\in A_\varepsilon} \int_{K^c} f_k(x)\mu(\mathrm{d}x) \le 1/i$ for $i \ge k$ such that $f_i \in \mathfrak{F}^\star(f_k)$. Since there are infinitely many such i, we deduce that

$$\inf_{K\in\mathscr{K}} \sup_{\mu\in A_\varepsilon} \int_{K^c} f_k(x)\mu(\mathrm{d}x) = 0,$$

i.e. the family $\{f_k\mu \colon \ \mu \in A_\varepsilon\}$ is tight. As this holds for all $k \in \mathbb{N}^*$, we get by Proposition 1.A.2 that A_ε is relatively compact in $\mathscr{M}_{\mathfrak{F}}$ (in fact, A_ε is compact as it is closed). This proves (i). The proof of (ii) is similar. □

We now give a sufficient condition for tightness in the space $\mathscr{M}_{\mathfrak{F}}$.

Corollary 1.A.5. *Assume that (H5) holds. Let Ξ be a family of $\mathscr{M}_{\mathfrak{F}}$-valued random variables such that for every $f \in \mathfrak{F}$,*

$$\sup_{\xi\in\Xi} \mathbb{E}\left[\xi(f)\right] < \infty. \tag{1.A.5}$$

Then Ξ is tight (in distribution) in $\mathscr{M}_{\mathfrak{F}}$.

Proof. By the Markov inequality, we have for every $f \in \mathfrak{F}$,

$$\sup_{\xi\in\Xi} \mathbb{P}\left(\xi(f) > r\right) \le \frac{1}{r} \sup_{\xi\in\Xi} \mathbb{E}\left[\xi(f)\right] \xrightarrow[r\to\infty]{} 0.$$

This proves that Ξ is tight in $\mathcal{M}_{\mathfrak{F}}$ by Proposition 1.A.4-(ii). $\square$

We denote by $\mathscr{B}$ (resp. $\mathscr{B}_{\mathfrak{F}}$) the Borel σ-field on $(\mathcal{M}, d_{\mathrm{BL}})$ (resp. on $(\mathcal{M}_{\mathfrak{F}}, d_{\mathfrak{F}})$). We also denote by $\mathscr{B}_{\mathrm{tr}} = \{A \cap \mathcal{M}_{\mathfrak{F}} \colon A \in \mathscr{B}\}$ the trace σ-field of $\mathscr{B}$ on $\mathcal{M}_{\mathfrak{F}}$.

Lemma 1.A.6. *We have $\mathscr{B}_{\mathfrak{F}} = \mathscr{B}_{\mathrm{tr}}$.*

Proof. **Step 1.** We first prove that $\mathcal{M}_{\mathfrak{F}}$ is a Borel subset in $\mathcal{M}$. For $g \in \mathscr{B}_+(S)$, we consider the function Θ_g defined on $\mathcal{M}$ by $\Theta_g(\mu) = g\mu$. Denote $\mathscr{B}_{b+} = \mathscr{B}_b(S) \cap \mathscr{B}_+(S)$ the set of bounded nonnegative measurable functions defined on S. We follow the proof of [21, Theorem 15.13] to prove that, for every $g \in \mathscr{B}_{b+}$, Θ_g is a measurable function from $\mathcal{M}$ to $\mathcal{M}$. Denote by $\mathscr{F} = \{g \in \mathscr{B}_{b+} \colon \Theta_g \text{ is measurable}\}$. The function Θ_g is continuous for g belonging to $\mathscr{C}_{b+} = \mathscr{C}_b(S) \cap \mathscr{C}_+(S)$. Furthermore, the set $\mathscr{F}$ is closed under bounded pointwise convergence: if $g_n \to g$ pointwise, with $g \in \mathscr{B}_{b+}$ and $(g_n, n \in \mathbb{N})$ a bounded sequence of elements of $\mathscr{F}$ (i.e. $\sup_{n\in\mathbb{N}} \|g_n\|_\infty < \infty$), then $\Theta_g(\mu) = \lim_{n\to\infty} \Theta_{g_n}(\mu)$ by dominated convergence and thus g belongs to $\mathscr{F}$. An immediate extension of [21, Theorem 4.33] gives that $\mathscr{B}_{b+} \subset \mathscr{F}$.

We then deduce that the function $\theta_g \colon \mathcal{M} \to [0, +\infty]$ defined by $\theta_g(\mu) = g\mu(\mathbf{1}) = \mu(g)$ is measurable for every $g \in \mathscr{B}_{b+}$, and as $g \in \mathscr{B}_+(S)$ is the limit of $g \wedge n \in \mathscr{B}_{b+}$ as n goes to infinity, we deduce by monotone convergence that $\theta_g = \lim_{n\to\infty} \theta_{g\wedge n}$, and thus θ_g is measurable for every $g \in \mathscr{B}_+(S)$. By definition of $\mathcal{M}_{\mathfrak{F}}$, we have that $\mathcal{M}_{\mathfrak{F}} = \bigcap_{f\in\mathfrak{F}} \theta_f^{-1}(\mathbb{R}_+)$, and thus $\mathcal{M}_{\mathfrak{F}}$ is a Borel subset in $\mathcal{M}$.

Step 2. We prove that for every $\mu \in \mathcal{M}_{\mathfrak{F}}$, the mapping $\nu \mapsto d_{\mathfrak{F}}(\mu, \nu)$ defined on $\mathcal{M}_{\mathfrak{F}}$ is $\mathscr{B}_{\mathrm{tr}}$-measurable. Let $g \in \mathscr{B}_{b+}$. Since the function Θ_g is measurable from $\mathcal{M}$ to itself by step 1, it is $\mathscr{B}/\mathscr{B}$-measurable. By definition of the trace σ-field, it follows that the mapping Θ_g from $\mathcal{M}_{\mathfrak{F}}$ to $\mathcal{M}$ is $\mathscr{B}_{\mathrm{tr}}/\mathscr{B}$-measurable. Let $f \in \mathfrak{F}$. By monotone convergence we get that $\Theta_f = \lim_{n\to\infty} \Theta_{f\wedge n}$, and thus Θ_f is $\mathscr{B}_{\mathrm{tr}}/\mathscr{B}$-measurable.

Since $\mu \in \mathcal{M}_{\mathfrak{F}}$, we have $f\mu \in \mathcal{M}$ and the mapping $\pi \mapsto d_{\mathrm{BL}}(f\mu, \pi)$ from $\mathcal{M}$ to $\mathbb{R}$ is continuous hence $\mathscr{B}$-measurable. Thus, by composition we get that the mapping $\nu \mapsto d_{\mathrm{BL}}(f\mu, f\nu)$ from $\mathcal{M}_{\mathfrak{F}}$ to $\mathbb{R}$ is $\mathscr{B}_{\mathrm{tr}}$-measurable. Finally, the mapping $\nu \mapsto d_{\mathfrak{F}}(\mu, \nu)$ from $\mathcal{M}_{\mathfrak{F}}$ to $\mathbb{R}$ is $\mathscr{B}_{\mathrm{tr}}$-measurable as a sum of $\mathscr{B}_{\mathrm{tr}}$-measurable mappings.

Step 3. We conclude the proof of the lemma. For every $\mu \in \mathcal{M}_{\mathfrak{F}}$ and every $\varepsilon > 0$, we have

$$B(\mu, \varepsilon) = \{\nu \in \mathcal{M}_{\mathfrak{F}} \colon d_{\mathfrak{F}}(\mu, \nu) < \varepsilon\} \in \mathscr{B}_{\mathrm{tr}}$$

by Step 2. Since $\mathcal{M}_{\mathfrak{F}}$ is a Polish space, every open set is the countable union of open balls and it follows that every open set lies in $\mathscr{B}_{\mathrm{tr}}$. Hence we get $\mathscr{B}_{\mathfrak{F}} \subset \mathscr{B}_{\mathrm{tr}}$.

Conversely, notice that the identity mapping from $(\mathcal{M}_{\mathfrak{F}}, d_{\mathfrak{F}})$ to $(\mathcal{M}_{\mathfrak{F}}, d_{\mathrm{BL}})$ is continuous. Therefore, if $V \subset \mathcal{M}$ is an open set, $V \cap \mathcal{M}_{\mathfrak{F}}$ is open in $(\mathcal{M}_{\mathfrak{F}}, d_{\mathrm{BL}})$ hence also in $(\mathcal{M}_{\mathfrak{F}}, d_{\mathfrak{F}})$. In particular, we have $V \cap \mathcal{M}_{\mathfrak{F}} \in \mathscr{B}_{\mathfrak{F}}$. Since this is true for every open set $V \subset \mathcal{M}$, we deduce that $\mathscr{B}_{\mathrm{tr}} \subset \mathscr{B}_{\mathfrak{F}}$. $\square$

The following two results are a direct consequence of Lemma 1.A.6.

Corollary 1.A.7. *Let ξ be a $\mathcal{M}$-valued random variable such that a.s. $\xi(f) < \infty$ for every $f \in \mathfrak{F}$. Then ξ is a $\mathcal{M}_{\mathfrak{F}}$-valued random variable. Conversely, if ξ is a $\mathcal{M}_{\mathfrak{F}}$-valued random variable then ξ is also a $\mathcal{M}$-valued random variable.*

Corollary 1.A.8. *Let ξ and ζ be $\mathcal{M}_{\mathfrak{F}}$-valued random variables. Then the following conditions are equivalent:*

(i) $\xi \overset{(d)}{=} \zeta$ *when viewed as $\mathcal{M}_{\mathfrak{F}}$-valued random variables.*

(ii) $\xi \overset{(d)}{=} \zeta$ *when viewed as $\mathcal{M}$-valued random variables.*

(iii) $\xi(h) \overset{(d)}{=} \zeta(h)$ *for every $h \in \mathscr{C}_b(S)$.*

(iv) $\xi(fh) \overset{(d)}{=} \zeta(fh)$ *for every $h \in \mathscr{C}_b(S)$ and $f \in \mathfrak{F}$.*

We now characterize convergence in distribution of random measures in $\mathcal{M}_{\mathfrak{F}}$. Recall that (H1)–(H4) are in force.

Proposition 1.A.9. *Let ξ_n and ξ be $\mathcal{M}_{\mathfrak{F}}$-valued random variables. Then ξ_n converges in distribution to ξ in $\mathcal{M}_{\mathfrak{F}}$ if and only if $\xi_n(fh) \xrightarrow[n\to\infty]{(d)} \xi(fh)$ for every $h \in \mathscr{C}_b(S)$ and every $f \in \mathfrak{F}$.*

Proof. Assume that ξ_n converges in distribution to ξ in $\mathcal{M}_{\mathfrak{F}}$. Let $f \in \mathfrak{F}$. Since $F\colon \mu \mapsto f\mu$ is continuous from $\mathcal{M}_{\mathfrak{F}}$ to $\mathcal{M}$ and $\nu \mapsto \nu(h)$ is continuous from $\mathcal{M}$ to $\mathbb{R}$ for every $h \in \mathscr{C}_b(S)$, it follows that the mapping $\mu \mapsto \mu(fh)$ is continuous from $\mathcal{M}_{\mathfrak{F}}$ to $\mathbb{R}$. By the continuous mapping theorem, we get $\xi_n(fh) \xrightarrow{(d)} \xi(fh)$.

Conversely, for every $f \in \mathfrak{F}$, $f\xi_n$ and $f\xi$ are $\mathcal{M}$-valued random variables, and we have $\xi_n(fh) \xrightarrow{(d)} \xi(fh)$ for every $h \in \mathscr{C}_b(S)$. By [99, Theorem 4.11], this implies that $f\xi_n \xrightarrow[n\to\infty]{(d)} f\xi$ in the space $\mathcal{M}$. In particular, $(f\xi_n, n \in \mathbb{N})$ is tight (in distribution) in $\mathcal{M}$ for every $f \in \mathfrak{F}$. By Proposition 1.A.4, it follows that $(\xi_n, n \in \mathbb{N})$ is tight in $\mathcal{M}_{\mathfrak{F}}$. Since $\mathcal{M}_{\mathfrak{F}}$ is Polish, Prokhorov's theorem ensures that $(\xi_n, n \in \mathbb{N})$ is relatively compact (in distribution) in $\mathcal{M}_{\mathfrak{F}}$. Let $\hat{\xi}$ be a limit point (in distribution) of $(\xi_n, n \in \mathbb{N})$. There exists a subsequence, still denoted by ξ_n, such that $\xi_n \xrightarrow{(d)} \hat{\xi}$ in $\mathcal{M}_{\mathfrak{F}}$. Let $h \in \mathscr{C}_b(S)$. Applying the first part of the proof, we get that $\xi_n(fh) \xrightarrow[n\to\infty]{(d)} \hat{\xi}(fh)$ for every $f \in \mathfrak{F}$. Therefore, we have $\hat{\xi}(fh) \overset{(d)}{=} \xi(fh)$ for every $h \in \mathscr{C}_b(S)$. It follows from Corollary 1.A.8 that $\hat{\xi} \overset{(d)}{=} \xi$ in $\mathcal{M}_{\mathfrak{F}}$. Thus the sequence $(\xi_n, n \in \mathbb{N})$ is relatively compact and has only one limit point ξ in $\mathcal{M}_{\mathfrak{F}}$. This proves the result. □

We state now the main result of this section. Recall that (H1)–(H4) are in force.

Proposition 1.A.10. *Let $(\xi_n, n \in \mathbb{N})$ be a sequence of $\mathcal{M}_{\mathfrak{F}}$-valued random variables and ξ be a $\mathcal{M}$-valued random variable such that $\xi_n \xrightarrow{(d)} \xi$ in $\mathcal{M}$ and $(\xi_n, n \in \mathbb{N})$ is tight (in distribution) in $\mathcal{M}_{\mathfrak{F}}$. Then ξ is a $\mathcal{M}_{\mathfrak{F}}$-valued random variable and we have the convergence in distribution $\xi_n \xrightarrow{(d)} \xi$ in $\mathcal{M}_{\mathfrak{F}}$.*

Proof. By assumption, the sequence $(\xi_n, n \in \mathbb{N})$ is relatively compact (in distribution) in the space $\mathcal{M}_{\mathfrak{F}}$. Let $\hat{\xi} \in \mathcal{M}_{\mathfrak{F}}$ be a limit point in distribution and let $h \in \mathscr{C}_b(S)$. On the one hand, Proposition 1.A.9 applied with $f = \mathbf{1}$ yields the convergence $\xi_n(h) \xrightarrow{(d)} \hat{\xi}(h)$. On the other hand, since $\xi_n \xrightarrow{(d)} \xi$ in $\mathcal{M}$ it follows that $\xi_n(h) \xrightarrow{(d)} \xi(h)$. Therefore $\hat{\xi}(h) \stackrel{(d)}{=} \xi(h)$ for every $h \in \mathscr{C}_b(S)$, i.e. $\hat{\xi} \stackrel{(d)}{=} \xi$ in $\mathcal{M}$. Since the distribution of $\hat{\xi}$ is concentrated on $\mathcal{M}_{\mathfrak{F}}$, the same is true for ξ. In other words $\xi \in \mathcal{M}_{\mathfrak{F}}$ a.s., and so ξ is a $\mathcal{M}_{\mathfrak{F}}$-valued random variable by Corollary 1.A.7. Now, applying Corollary 1.A.8 we get $\hat{\xi} \stackrel{(d)}{=} \xi$ in the space $\mathcal{M}_{\mathfrak{F}}$. Thus the sequence $(\xi_n, n \in \mathbb{N})$ is relatively compact in $\mathcal{M}_{\mathfrak{F}}$ and has only one limit point ξ, so $\xi_n \xrightarrow{(d)} \xi$ in $\mathcal{M}_{\mathfrak{F}}$. □

The following special case is particularly useful. Recall that (H1)–(H4) are in force.

Corollary 1.A.11. *Assume that (H5) holds. Let $(\xi_n, n \in \mathbb{N})$ and ξ be $\mathcal{M}$-valued random variables such that $\xi_n \xrightarrow{(d)} \xi$ in $\mathcal{M}$ and for every $f \in \mathfrak{F}$,*

$$\sup_n \mathbb{E}\left[\xi_n(f)\right] < \infty. \tag{1.A.6}$$

Then $(\xi_n, n \in \mathbb{N})$ and ξ are $\mathcal{M}_{\mathfrak{F}}$-valued random variables and we have the convergence in distribution $\xi_n \xrightarrow{(d)} \xi$ in $\mathcal{M}_{\mathfrak{F}}$. Moreover, for every $f \in \mathfrak{F}$, we have

$$\mathbb{E}\left[\xi(f)\right] \le \liminf_{n\to\infty} \mathbb{E}\left[\xi_n(f)\right] < \infty.$$

Furthermore, if $(\mathbb{E}[\xi_n(\bullet)], n \in \mathbb{N})$ converges to $\mathbb{E}[\xi(\bullet)]$ in $\mathcal{M}$ then the convergence actually holds in $\mathcal{M}_{\mathfrak{F}}$.

Proof. The random variable ξ_n is $\mathcal{M}$-valued and satisfies $\xi_n(f) < \infty$ a.s. since $\mathbb{E}\left[\xi_n(f)\right] < \infty$ for every $f \in \mathfrak{F}$, so by Corollary 1.A.7, ξ_n is a $\mathcal{M}_{\mathfrak{F}}$-valued random variable. By Corollary 1.A.5, the assumption (1.A.6) implies that $(\xi_n, n \in \mathbb{N})$ is tight (in distribution) in $\mathcal{M}_{\mathfrak{F}}$. Therefore Proposition 1.A.10 applies and gives the convergence in distribution $\xi_n \xrightarrow{(d)} \xi$ in $\mathcal{M}_{\mathfrak{F}}$. Moreover, Skorokhod's representation theorem in conjunction with Fatou's lemma implies that for every $f \in \mathfrak{F}$,

$$\mathbb{E}\left[\xi(f)\right] \le \liminf_{n\to\infty} \mathbb{E}\left[\xi_n(f)\right] < \infty.$$

Now set $\mu_n = \mathbb{E}[\xi_n(\bullet)]$ and $\mu = \mathbb{E}[\xi(\bullet)]$ and assume that $\mu_n \to \mu$ in $\mathcal{M}$. Notice that the assumption (1.A.6) implies that $\mu_n \in \mathcal{M}_{\mathfrak{F}}$ for every $n \in \mathbb{N}$ and that the sequence of measures $(f\mu_n, n \in \mathbb{N})$ is bounded for every $f \in \mathfrak{F}$. Thus Corollary 1.A.3 gives the convergence $\lim_{n\to\infty} \mu_n = \mu$ in $\mathcal{M}_{\mathfrak{F}}$. □

1.B Sub-exponential tail bounds for the height of conditioned BGW trees

Assume that ξ satisfies (ξ1) and (ξ2) and denote by τ^n a BGW(ξ) tree conditioned to have n vertices. Then by [112, Theorem 1] which is stated for the aperiodic case but is trivially extended to the general case, for every $\alpha \in (0, \gamma/(\gamma-1))$, there exist two constants $C_0, c_0 > 0$ such that for every $y \geq 0$ and every $n \in \Delta$

$$\mathbb{P}\left(\frac{b_n}{n}\mathfrak{h}(\tau^n) \leq y\right) \leq C_0 \exp\left(-c_0 y^{-\alpha}\right). \tag{1.B.1}$$

We will show that under the stronger assumption $(\xi 2)'$, the previous inequality holds with $\alpha = \gamma/(\gamma-1)$. Since the finite variance case has already been treated in [14], we assume henceforth that ξ has infinite variance.

Recall that L is a slowly varying function such that $\mathbb{E}\left[\xi^2 \mathbf{1}_{\{\xi \leq n\}}\right] = n^{2-\gamma} L(n)$. On the other hand, the slowly varying function appearing in the appendix of [112], which we denote by K, satisfies $\mathrm{Var}\left(\xi \mathbf{1}_{\{\xi \leq n\}}\right) = n^{2-\gamma} K(n)$. Since $\mathrm{Var}(\xi) = +\infty$, we have as n goes to infinity that

$$\mathbb{E}\left[\xi^2 \mathbf{1}_{\{\xi \leq n\}}\right] \sim n^{2-\gamma} K(n) + 1 \sim n^{2-\gamma} K(n),$$

see the appendix in [112]. Therefore, we get $K(n) \sim L(n)$ and K is bounded above.

Following the proof of [112, Theorem 1] to get (1.B.1) holds for $\alpha = \gamma/(\gamma-1)$, it is enough to prove the analogue of Proposition 8 therein with $\alpha = \gamma/(\gamma-1)$, that is Proposition 1.B.1 below. Let $(W_n, n \in \mathbb{N})$ be a random walk with starting point $W_0 = 0$ and jump distribution $\xi - 1$.

Proposition 1.B.1. *Assume that ξ satisfies $(\xi 1)$ and $(\xi 2)'$. There exist two constants $C_0, c_0 > 0$ such that for every $u \geq 0$ and every $n \geq 1$,*

$$\mathbb{P}\left(\min_{1 \leq i \leq n} W_i \leq -u b_n\right) \leq C_0 \exp\left(-c_0 u^{\gamma/(\gamma-1)}\right). \tag{1.B.2}$$

Proof. Note that $\mathbb{P}\left(\min_{1 \leq i \leq n} W_i \leq -u b_n\right) = 0$ if $u b_n > n$, so that it is enough to prove (1.B.2) for $1 \leq u \leq n/b_n$. Write, for $h > 0$

$$\mathbb{P}\left(\min_{1 \leq i \leq n} W_i \leq -u b_n\right) = \mathbb{P}\left(\max_{1 \leq i \leq n} \mathrm{e}^{-hW_i} \geq \mathrm{e}^{hub_n}\right) \leq \mathrm{e}^{-hub_n}\, \mathbb{E}\left[\mathrm{e}^{-hW_n}\right] = \mathrm{e}^{-hub_n}\, \mathbb{E}\left[\mathrm{e}^{-hW_1}\right]^n, \tag{1.B.3}$$

where the inequality follows from Doob's maximal inequality applied to the submartingale $(\mathrm{e}^{-hW_n}, n \in \mathbb{N})$. We shall apply (1.B.3) with $h = \varepsilon u^\eta / b_n$ where $\eta = 1/(\gamma-1)$ and $\varepsilon > 0$ is a constant to be chosen later. Note that $\gamma/(\gamma-1) = \eta\gamma = 1 + \eta$. Observe that $\varepsilon u^\eta / b_n$ is bounded uniformly in $1 \leq u \leq n/b_n$ and $n \geq 1$. Indeed, since $b_n \geq \underline{b} n^{1/\gamma}$, we have

$$\frac{u^\eta}{b_n} \leq \left(\frac{n}{b_n}\right)^\eta \frac{1}{b_n} \leq \frac{1}{\underline{b}^{1+\eta}}.$$

Therefore, using [112, Eq. (42)] together with the inequality $1 + x \leq \mathrm{e}^x$, we have for every $n \geq 1$ and every $1 \leq u \leq n/b_n$

$$\mathbb{E}\left[\mathrm{e}^{-\varepsilon \frac{u^\eta}{b_n} W_1}\right]^n \leq \exp\left\{Cn\left(\varepsilon \frac{u^\eta}{b_n}\right)^\gamma K\left(\frac{b_n}{\varepsilon u^\eta}\right)\right\} \leq \exp\left(C' \varepsilon^\gamma u^{\eta\gamma}\right),$$

as K is bounded from above and $b_n \geq \underline{b} n^{1/\gamma}$. Thus, we deduce from (1.B.3) that for $1 \leq u \leq n/b_n$

$$\mathbb{P}\left(\min_{1 \leq i \leq n} W_i \leq -u b_n\right) \leq \exp\left(-\left(\varepsilon - C' \varepsilon^{\gamma}\right) u^{1+\eta}\right).$$

The conclusion readily follows by choosing $\varepsilon > 0$ small enough such that $\varepsilon - C'\varepsilon^{\gamma} > 0$. □

Remark 1.B.2. In fact, this proof is valid if we only assume that the slowly varying function L of $(\xi 2)'$ is bounded from above, in which case $n^{-1/\gamma} b_n$ is bounded below.

2 Zooming in at the root of the stable tree

This chapter is based on the paper [128], published in *Electronic Journal of Probability*.

We study the shape of the normalized stable Lévy tree $\mathcal{T}$ near its root. We show that, when zooming in at the root at the proper speed with a scaling depending on the index of stability, we get the unnormalized Kesten tree. In particular the limit is described by a tree-valued Poisson point process which does not depend on the initial normalization. We apply this to study the asymptotic behavior of additive functionals of the form

$$\mathbf{Z}_{\alpha,\beta} = \int_{\mathcal{T}} \mu(\mathrm{d}x) \int_0^{H(x)} \sigma_{r,x}^{\alpha} \mathfrak{h}_{r,x}^{\beta} \, \mathrm{d}r$$

as $\max(\alpha, \beta) \to \infty$, where μ is the mass measure on $\mathcal{T}$, $H(x)$ is the height of x and $\sigma_{r,x}$ (resp. $\mathfrak{h}_{r,x}$) is the mass (resp. height) of the subtree of $\mathcal{T}$ above level r containing x. Such functionals arise as scaling limits of additive functionals of the size and height on conditioned Bienaymé-Galton-Watson trees.

2.1 Introduction

Stable trees are special instances of Lévy trees which were introduced by Le Gall and Le Jan [119] in order to generalize Aldous' Brownian tree [16]. More precisely, stable trees are compact weighted rooted real trees depending on a parameter $\gamma \in (1,2]$, with $\gamma = 2$ corresponding to the Brownian tree, which encode the genealogical structure of continuous-state branching processes with branching mechanism $\psi(\lambda) = \lambda^{\gamma}$. As such, they are the possible scaling limits of Bienaymé-Galton-Watson trees with critical offspring distribution belonging to the domain of attraction of a stable distribution with index $\gamma \in (1,2]$, see Duquesne [54] and Kortchemski [110]. They also appear as scaling limits of various models of trees and graphs, see e.g. Haas and Miermont [82], and are intimately related to fragmentation and coalescence processes, see Miermont [124, 125] and Berestycki, Berestycki and Schweinsberg [27]. Stable trees can be defined via the normalized excursion of the so-called height process which is a local time functional of a spectrally positive Lévy process. We refer to Duquesne and Le Gall [57] for a detailed account. See also Duquesne and Winkel [62], Goldschmidt and Haas [77], Marchal [122] for alternative constructions.

In the present paper, we study the shape of the normalized stable tree $\mathscr{T}$ (*i.e.* the stable tree conditioned to have total mass 1) near its root. More precisely we show that, after zooming in at the root of $\mathscr{T}$ and rescaling, one gets the continuous analogue of the Kesten tree, that is a random real tree consisting of an infinite branch on which subtrees are grafted according to a Poisson point process. In particular, the (rescaled) subtrees near the root of $\mathscr{T}$ are independent and the conditioning for the total mass to be equal to 1 disappears when zooming in. This idea to zoom in at the root of the stable tree is closely related to the small time asymptotics – present in the works of Miermont [124] and Haas [80] – of the self-similar fragmentation process $F^-(t)$ obtained from the stable tree by removing vertices located under height t. See Remark 2.4.5 in this direction. As a consequence, we obtain the asymptotic behavior of additive functionals on $\mathscr{T}$ of the form

$$\mathbf{Z}_{\alpha,\beta} = \int_{\mathscr{T}} Z_{\alpha,\beta}(x)\,\mu(\mathrm{d}x) \qquad \text{with} \quad \forall x \in \mathscr{T},\ Z_{\alpha,\beta}(x) = \int_0^{H(x)} \sigma_{r,x}^{\alpha}\mathfrak{h}_{r,x}^{\beta}\,\mathrm{d}r, \tag{2.1.1}$$

where μ is the mass measure on $\mathscr{T}$ which is a uniform measure supported by the set of leaves, $H(x)$ is the height of $x \in \mathscr{T}$, that is its distance to the root, and $\sigma_{r,x}$ (resp. $\mathfrak{h}_{r,x}$) is the mass (resp. height) of the subtree of $\mathscr{T}$ above level r containing x.

Before stating our results, we first introduce some notations. Let $\mathbb{T}$ be the space of weighted rooted compact real trees, that is the set of compact real trees (T,d) endowed with a distinguished vertex $\emptyset$ called the root and with a nonnegative finite measure μ. We equip the set $\mathbb{T}$ with the Gromov-Hausdorff-Prokhorov topology, see Section 2.2 for a precise definition.

Define a rescaling map $R_\gamma\colon \mathbb{T}\times(0,\infty)\to\mathbb{T}$ by

$$R_\gamma\big((T,\emptyset,d,\mu),a\big) = \big(T,\emptyset,ad,a^{\gamma/(\gamma-1)}\mu\big). \tag{2.1.2}$$

In words, $R_\gamma((T,\emptyset,d,\mu),a)$ is the tree obtained from $(T,\emptyset,d,\mu)$ by multiplying all distances by a and all masses by $a^{\gamma/(\gamma-1)}$. Moreover, define for every $(T,\emptyset,d,\mu)\in\mathbb{T}$

$$\mathsf{norm}_\gamma(T) = R_\gamma(T,\mu(T)^{-1+1/\gamma}), \tag{2.1.3}$$

which is the tree T normalized to have total mass 1 and where distances are rescaled accordingly. Denote by $\mathbf{N}^{(1)}$ the distribution of the normalized stable tree with total mass 1, see Section 2.3 for a precise definition. Under $\mathbf{N}^{(1)}$, let U be a uniformly chosen leaf, that is U is a $\mathscr{T}$-valued random variable with distribution μ. Denote by $\mathscr{T}_i$, $i\in I_U$ the trees grafted on the branch $[\![\emptyset,U]\!]$ joining the root $\emptyset$ to the leaf U, each one at height h_i and with total mass $\sigma_i=\mu(\mathscr{T}_i)$, see Figure 2.1. Fix $\mathfrak{f}\colon(0,\infty)\to(0,\infty)$ (this represents the speed at which we zoom in) and define for every $\varepsilon>0$ a point measure on $[0,\infty)^2\times\mathbb{T}$ by

$$\mathscr{N}_\varepsilon^{\mathfrak{f}}(U) = \sum_{h_i\le\mathfrak{f}(\varepsilon)H(U)} \delta_{(\varepsilon^{-1}h_i,\varepsilon^{-\gamma/(\gamma-1)}\sigma_i,\mathsf{norm}_\gamma(\mathscr{T}_i))}. \tag{2.1.4}$$

Finally, for any metric space X, we denote by $\mathscr{M}_p(X)$ the space of point measures on X equipped with the topology of vague convergence.

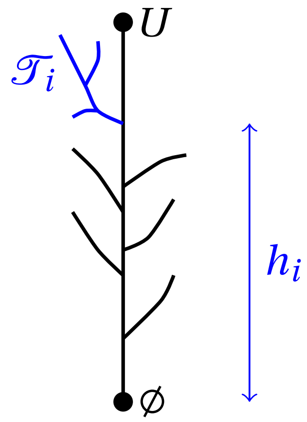

Figure 2.1 – The subtrees $\mathscr{T}_i$ grafted on the branch $[\![\emptyset, U]\!]$ at height h_i.

Our first main result states that the measure $\mathscr{N}_\varepsilon^{\mathfrak{f}}(U)$ converges to a Poisson point process which is independent of the underlying tree $\mathscr{T}$ and of $H(U)$.

Theorem 2.1.1. *Let $\mathscr{T}$ be the normalized stable tree with branching mechanism $\psi(\lambda) = \lambda^\gamma$ where $\gamma \in (1,2]$. Conditionally on $\mathscr{T}$, let U be a $\mathscr{T}$-valued random variable with distribution μ under $\mathbf{N}^{(1)}$. Let $(\mathsf{T}'_s, s \geq 0)$ be a Poisson point process with intensity $\mathbf{N}_{\mathrm{B}}$ given by (2.4.1), independent of $(\mathscr{T}, H(U))$. Let $\Phi\colon [0,\infty)^2 \times \mathbb{T} \to [0,\infty)$ be a measurable function such that there exists $C > 0$ such that for every $h \geq 0$ and $T \in \mathbb{T}$, we have*

$$|\Phi(h,b,T) - \Phi(h,a,T)| \leq C|b-a|.$$

(i) If $\lim_{\varepsilon\to 0} \varepsilon^{-1/2}\mathfrak{f}(\varepsilon) = 0$ and $\lim_{\varepsilon\to 0}\varepsilon^{-1}\mathfrak{f}(\varepsilon) = \infty$, then we have the following convergence in distribution

$$\left(\mathscr{T}, H(U), \langle \mathscr{N}_\varepsilon^{\mathfrak{f}}(U), \Phi\rangle\right) \xrightarrow[\varepsilon\to 0]{(d)} \left(\mathscr{T}, H(U), \sum_{s\geq 0} \Phi\left(s, \mu(\mathsf{T}'_s), \mathrm{norm}_\gamma\left(\mathsf{T}'_s\right)\right)\right) \tag{2.1.5}$$

in the space $\mathbb{T} \times [0,\infty) \times [0,\infty]$. In particular, we have the following convergence in distribution in $\mathbb{T} \times [0,\infty) \times \mathscr{M}_p([0,\infty) \times \mathbb{T})$

$$\left(\mathscr{T}, H(U), \sum_{h_i \leq \mathfrak{f}(\varepsilon) H(U)} \delta_{(\varepsilon^{-1}h_i, R_\gamma(\mathscr{T}_i, \varepsilon^{-1}))}\right) \xrightarrow[\varepsilon\to 0]{(d)} \left(\mathscr{T}, H(U), \sum_{s\geq 0} \delta_{(s, \mathsf{T}'_s)}\right). \tag{2.1.6}$$

(ii) If $\mathfrak{f}(\varepsilon) = \varepsilon$, then we have the following convergence in distribution

$$\left(\mathscr{T}, H(U), \langle \mathscr{N}_\varepsilon^{\mathfrak{f}}(U), \Phi\rangle\right) \xrightarrow[\varepsilon\to 0]{(d)} \left(\mathscr{T}, H(U), \sum_{s\leq H(U)} \Phi\left(s, \mu(\mathsf{T}'_s), \mathrm{norm}_\gamma\left(\mathsf{T}'_s\right)\right)\right) \tag{2.1.7}$$

in the space $\mathbb{T} \times [0,\infty) \times [0,\infty]$.

In other words, zooming in at the speed $\mathfrak{f}(\varepsilon) = \varepsilon$ gives a *finite* branch on which subtrees are grafted in a Poissonian manner, whereas zooming in at a slower speed gives an *infinite* branch at the limit. Notice that the convergence (2.1.5) is stronger than convergence in distribution for the vague topology (2.1.6) as it holds for functions Φ with very few regularity assumptions: $\Phi(h,a,T)$ is only Lipschitz-continuous with respect to a instead of (Lipschitz-)continuous with respect to (h,a,T) with bounded support. In particular, this could allow to consider local time functionals of the tree.

As an application of this result, we study the asymptotic behavior as $\max(\alpha,\beta)\to\infty$ of additive functionals $\mathbf{Z}_{\alpha,\beta}$ on the stable tree $\mathscr{T}$. Such functionals arise as scaling limits of additive functionals of the size and height on conditioned Bienaymé-Galton-Watson trees, see Delmas, Dhersin and Sciauveau [52] or Abraham, Delmas and Nassif [7] where it is shown that $\mathbf{Z}_{\alpha,\beta}<\infty$ a.s. if (and only if) $\gamma\alpha+(\gamma-1)(\beta+1)>0$, see Corollary 6.10 therein. In the present paper, we only consider $\alpha,\beta\geq 0$ which guarantees in particular the finiteness of $\mathbf{Z}_{\alpha,\beta}$. For example, let us mention the total path length and the Wiener index which when properly scaled converge respectively to $\mathbf{Z}_{0,0}$ and $\mathbf{Z}_{1,0}$. Fill and Janson [73] considered the case $\gamma=2$ and $\beta=0$ (*i.e.* functionals of the mass on the Brownian tree) and proved that there is convergence in distribution as $\alpha\to\infty$ of $\mathbf{Z}_{\alpha,0}$ properly normalized to

$$\int_0^\infty \mathrm{e}^{-S_t}\,\mathrm{d}t,$$

where $(S_t,\ t\geq 0)$ is a $1/2$-stable subordinator. Their proof relies on the connection between the normalized Brownian excursion which codes the Brownian tree and the three-dimensional Bessel bridge. Our aim is twofold: we extend their result to the non-Brownian stable case $\gamma\in(1,2)$ while also considering polynomial functionals depending on both the mass and the height. We use a different approach relying on the Bismut decomposition of the stable tree.

Going back to the connection with the fragmentation process $F^-(t)=(F_1^-(t),F_2^-(t),\ldots)$, it is not hard to see that the additive functional $\mathbf{Z}_{\alpha,0}$ can be expressed in terms of F^- as

$$\mathbf{Z}_{\alpha,0}=\sum_{i\geq 1}\int_0^\infty F_i^-(t)^{\alpha+1}\,\mathrm{d}t.$$

Once this is established, one can argue that only the largest fragment F_1^- contributes to the limit, the others being negligible, then use [80, Corollary 17] which implies that $1-F_1^-$ properly normalized converges in distribution to a $(1-1/\gamma)$-stable subordinator S, to get the convergence of $\mathbf{Z}_{\alpha,0}$ to $\int_0^\infty \mathrm{e}^{-S_t}\,\mathrm{d}t$. In the present paper, we do not adopt this approach as it does not allow to consider functionals of the height (that is $\beta\neq 0$).

We distinguish two regimes according to the behavior of $\beta/\alpha^{1-1/\gamma}$. The regime $\beta/\alpha^{1-1/\gamma}\to c\in[0,\infty)$ is related to Theorem 2.1.1 and the result in that case can be stated as follows, see Theorem 2.5.4 for a more general statement.

Theorem 2.1.2. *Assume that $\alpha \to \infty$, $\beta \geq 0$ and $\beta/\alpha^{1-1/\gamma} \to c \in [0,\infty)$. Let $\mathscr{T}$ be the normalized stable tree with branching mechanism $\psi(\lambda) = \lambda^\gamma$ where $\gamma \in (1,2]$ and denote by $\mathfrak{h}$ its height. Then we have the following convergence in distribution under $\mathbf{N}^{(1)}$*

$$\alpha^{1-1/\gamma}\mathfrak{h}^{-\beta}\mathbf{Z}_{\alpha,\beta} \xrightarrow[\alpha\to\infty]{(d)} \int_0^\infty \mathrm{e}^{-S_t - ct/\mathfrak{h}}\,\mathrm{d}t, \tag{2.1.8}$$

where $(S_t, t \geq 0)$ is a stable subordinator with Laplace exponent $\varphi(\lambda) = \gamma\lambda^{1-1/\gamma}$, independent of $\mathscr{T}$.

Let us briefly explain why we get a subordinator S at the limit. It is well known that μ is supported on the set of leaves of $\mathscr{T}$. Let $x \in \mathscr{T}$ be a leaf and recall that $\sigma_{r,x}$ is the mass of the subtree above level r containing x. Since the total mass of the stable tree is 1, the main contribution to $Z_{\alpha,\beta}(x)$ as $\alpha \to \infty$ comes from large subtrees $\mathscr{T}_{r,x}$ with r close to 0. The height $\mathfrak{h}_{r,x}$ of such subtrees is approximately $\mathfrak{h} - r$. On the other hand, their mass is equal to 1 minus the mass we discarded from the subtrees grafted on the branch $[\![\emptyset, x]\!]$ at height less than r. By Theorem 2.1.1, subtrees are grafted on $[\![\emptyset, x]\!]$ according to a point process which is approximately Poissonian, at least close to the root $\emptyset$. Thus the mass $\sigma_{r,x}$ is approximately $1 - S_r$.

Theorem 2.5.4 is slightly more general: we prove joint convergence in distribution of $\alpha^{1-1/\gamma}\mathfrak{h}^{-\beta}$ $\mathbf{Z}_{\alpha,\beta}$ and $\alpha^{1-1/\gamma}\mathfrak{h}^{-\beta}Z_{\alpha,\beta}(U)$, where $U \in \mathscr{T}$ is a leaf chosen uniformly at random (*i.e.* according to the measure μ), to the same random variable. In other words, taking the average of $Z_{\alpha,\beta}(x)$ over all leaves yields the same asymptotic behavior as taking a leaf uniformly at random. This is due to the following observations: a) a uniform leaf U is not too close to the root with high probability in the sense that its most recent common ancestor with x^* has height greater than ε, where x^* is the heighest leaf of $\mathscr{T}$, b) when taking the average over all leaves, the contribution of those leaves whose most recent common ancestor with x^* has height less than ε is negligible, and c) for those $x \in \mathscr{T}$ whose most recent common ancestor with x^* has height greater than ε, the main contribution to $Z_{\alpha,\beta}(x)$ comes from large subrees $\mathscr{T}_{r,x}$ with $r \leq \varepsilon$, these subtrees are common to all such leaves as $\mathscr{T}_{r,x} = \mathscr{T}_{r,x^*}$. This is made rigorous in Lemma 2.5.3.

Let us make a connection with Theorem 1.18 of Fill and Janson [73]. Recall that the normalized Brownian tree with branching mechanism $\psi(\lambda) = \lambda^2$ is coded by $\sqrt{2}B^{\mathrm{ex}}$ where B^{ex} is the normalized Brownian excursion, see [57]. Thanks to the representation formula of [52, Lemma 8.6], we see that Fill and Janson's $Y(\alpha) = \sqrt{2}\mathbf{Z}_{\alpha-1,0}$. Thus, we recover their result in the Brownian case $\gamma = 2$ when $\beta = 0$ (in which case $c = 0$).

Notice that as long as the exponent β of the height does not grow too quickly, viz. $\beta/\alpha^{1-1/\gamma} \to 0$, the additional dependence on the height makes no contribution at the limit. On the other hand, in the regime $\beta/\alpha^{1-1/\gamma} \to \infty$, the height $\mathfrak{h}_{r,x}^\beta$ dominates the mass $\sigma_{r,x}^\alpha$ so we get the convergence in probability of $\mathbf{Z}_{\alpha,\beta}$ with a different scaling and there is no longer a subordinator at the limit. See Theorem 2.6.1 for a more general statement.

Theorem 2.1.3. *Assume that $\beta \to \infty$, $\alpha \geq 0$ and $\alpha^{1-1/\gamma}/\beta \to 0$. Let $\mathcal{T}$ be the normalized stable tree with branching mechanism $\psi(\lambda) = \lambda^\gamma$ where $\gamma \in (1,2]$. Then we have the following convergence in $\mathbf{N}^{(1)}$-probability*

$$\lim_{\beta\to\infty} \beta\mathfrak{h}^{-\beta}\mathbf{Z}_{\alpha,\beta} = \mathfrak{h}. \tag{2.1.9}$$

Remark 2.1.4. Assume that $\alpha, \beta \to \infty$ and $\beta/\alpha^{1-1/\gamma} \to c \in (0,\infty)$ so that Theorem 2.1.2 applies. Then we have the convergence in distribution under $\mathbf{N}^{(1)}$

$$\beta\mathfrak{h}^{-\beta}\mathbf{Z}_{\alpha,\beta} = \frac{\beta}{\alpha^{1-1/\gamma}}\alpha^{1-1/\gamma}\mathfrak{h}^{-\beta}\mathbf{Z}_{\alpha,\beta} \xrightarrow[\beta\to\infty]{(d)} c\int_0^\infty \mathrm{e}^{-S_t - ct/\mathfrak{h}}\,\mathrm{d}t = \mathfrak{h}\int_0^\infty \mathrm{e}^{-S_{\mathfrak{h}t/c}-t}\,\mathrm{d}t.$$

Now letting $c \to \infty$, the right-hand side converges to $\mathfrak{h}\int_0^\infty \mathrm{e}^{-t}\,\mathrm{d}t = \mathfrak{h}$. Thus, one may view Theorem 2.1.3 as a special case of Theorem 2.1.2 by saying that, if $\beta \to \infty$ and $\beta/\alpha^{1-1/\gamma} \to c \in (0,\infty]$, then we have the convergence in distribution under $\mathbf{N}^{(1)}$

$$\beta\mathfrak{h}^{-\beta}\mathbf{Z}_{\alpha,\beta} \xrightarrow[\beta\to\infty]{(d)} c\int_0^\infty \mathrm{e}^{-S_t - ct/\mathfrak{h}}\,\mathrm{d}t,$$

where the measure $c\mathrm{e}^{-ct/\mathfrak{h}}\,\mathrm{d}t$ on $[0,\infty)$ should be understood as $\mathfrak{h}\delta_0$ if $c = \infty$.

We conclude the introduction by giving a decomposition of a general (compact) Lévy tree used in the proof of Theorem 2.1.2 which is of independent interest. Consider a Lévy tree $\mathcal{T}$ under its excursion measure $\mathbf{N}$ associated with a branching mechanism $\psi(\lambda) = a\lambda + b\lambda^2 + \int_0^\infty (\mathrm{e}^{-\lambda r} - 1 + \lambda r)\,\pi(\mathrm{d}r)$ where $a, b \geq 0$ and π is a σ-finite measure on $(0,\infty)$ satisfying $\int_0^\infty (r \wedge r^2)\,\pi(\mathrm{d}r) < \infty$. We further assume that the Grey condition holds $\int^\infty \mathrm{d}\lambda/\psi(\lambda) < \infty$ which is equivalent to the compactness of the Lévy tree. We refer to [57, Section 1] for a complete presentation of the subject. For every $x \in \mathcal{T}$ and every $0 \leq r < r' \leq H(x)$, we let $\mathcal{T}_{[r,r'),x} = (\mathcal{T}_{r,x} \setminus \mathcal{T}_{r',x}) \cup \{x_{r'}\}$ where $x_{r'}$ is the unique ancestor of x at height $H(x_{r'}) = r'$ and $\mathcal{T}_{r,x}$ is the subtree of $\mathcal{T}$ above level r containing x. The following result states that, when $x \in \mathcal{T}$ and $0 =: r_0 < r_1 < \ldots < r_n < r_{n+1} := H(x)$ are chosen "uniformly" at random under $\mathbf{N}$, then the random trees $\mathcal{T}_{[r_{i-1},r_i),x}$, $1 \leq i \leq n+1$ are independent and distributed as $\mathcal{T}$ under $\mathbf{N}[\sigma\bullet]$, see Figure 2.2. In particular, this generalizes [7, Lemma 6.1] which corresponds to $n = 1$.

Theorem 2.1.5. *Let $\mathcal{T}$ be the Lévy tree with a general branching mechanism ψ satisfying the Grey condition $\int^\infty \mathrm{d}\lambda/\psi(\lambda) < \infty$ under its excursion measure $\mathbf{N}$. Then for every $n \geq 1$ and all nonnegative measurable functions f_i, $1 \leq i \leq n+1$ defined on $[0,\infty) \times \mathbb{T}$, we have with $r_0 = 0$ and $r_{n+1} = H(x)$*

$$\mathbf{N}\left[\int_{\mathcal{T}} \mu(\mathrm{d}x) \int_{0<r_1<\ldots<r_n<H(x)} \prod_{i=1}^{n+1} f_i\left(r_i - r_{i-1}, \mathcal{T}_{[r_{i-1},r_i),x}\right) \prod_{i=1}^{n} \mathrm{d}r_i\right]$$
$$= \prod_{i=1}^{n+1} \mathbf{N}\left[\int_{\mathcal{T}} \mu(\mathrm{d}x) f_i(H(x),\mathcal{T})\right].$$

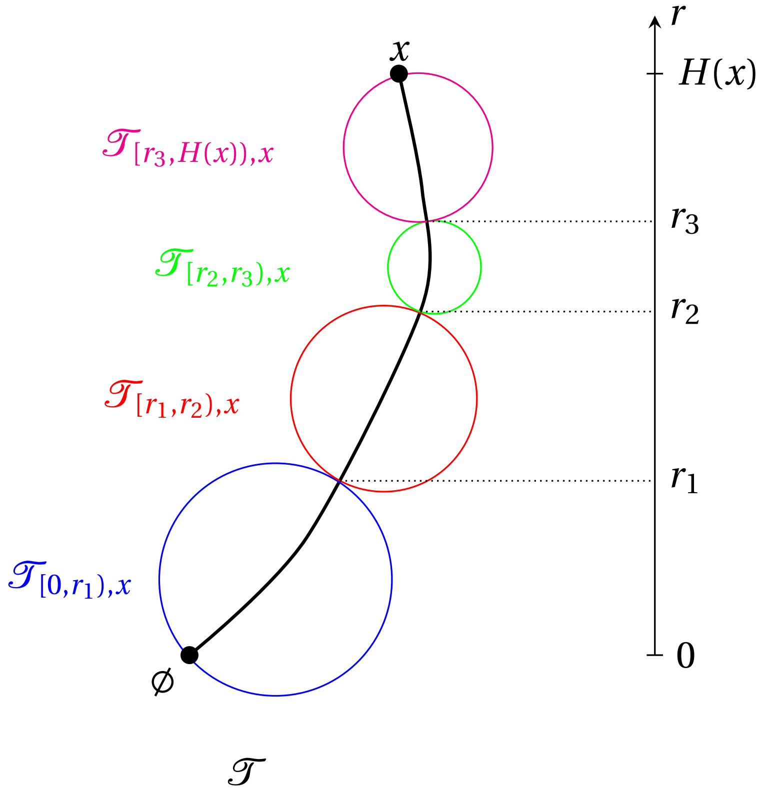

Figure 2.2 – The decomposition of $\mathscr{T}$ under $\mathbf{N}$ into $n+1$ (with $n=3$) subtrees along the ancestral line of a uniformly chosen leaf x.

In particular, for every nonnegative measurable functions g_i, $1 \le i \le n+1$ defined on $\mathbb{T}$, we have

$$\mathbf{N}\left[\int_{\mathscr{T}} \mu(\mathrm{d}x) \int_{0<r_1<\ldots<r_n<H(x)} \prod_{i=1}^{n+1} g_i\left(\mathscr{T}_{[r_{i-1},r_i),x}\right) \prod_{i=1}^{n} \mathrm{d}r_i\right] = \prod_{i=1}^{n+1} \mathbf{N}\left[\sigma g_i(\mathscr{T})\right].$$

A consequence of this decomposition is the following result giving the joint distribution of $\mathscr{T}_y$, the subtree of $\mathscr{T}$ above vertex $y \in \mathscr{T}$, and $H(y)$ when y is chosen according to the length measure $\ell(\mathrm{d}y)$ on the stable tree $\mathscr{T}$ (which roughly speaking is the Lebesgue measure on the branches of $\mathscr{T}$). In particular, this generalizes [7, Proposition 1.6].

Corollary 2.1.6. *Let $\mathscr{T}$ be the normalized stable tree with branching mechanism $\psi(\lambda) = \lambda^\gamma$ where $\gamma \in (1,2]$. Let f and g be nonnegative measurable functions defined on $\mathbb{T}$ and $[0,\infty)$ respectively. We have*

$$\mathbf{N}^{(1)}\left[\int_{\mathscr{T}} f(\mathscr{T}_y) g(H(y))\, \ell(\mathrm{d}y)\right] = \mathbf{N}\left[\mathbf{1}_{\{\sigma<1\}}(1-\sigma)^{-1/\gamma} G(1-\sigma) f(\mathscr{T})\right] \tag{2.1.10}$$

where

$$G(a) = \mathbf{N}^{(1)}\left[\int_{\mathscr{T}} \mu(\mathrm{d}x) g\left(a^{1-1/\gamma} H(x)\right)\right], \quad \forall a > 0.$$

The paper is organized as follows. In Section 2.2 we define the space of real trees and the Gromov-Hausdorff-Prokhorov topology. In Section 2.3, we introduce the stable tree, recall some of its properties and prove Theorem 2.1.5 as well as some other useful results. In Section 2.4, we prove Theorem 2.1.1. Sections 2.5 and 2.6 deal with the asymptotic behavior of $\mathbf{Z}_{\alpha,\beta}$ when $\beta/\alpha^{1-1/\gamma} \to c \in [0,\infty)$ and $\beta/\alpha^{1-1/\gamma} \to \infty$ respectively. Finally, we gather some technical proofs in Section 2.7.

2.2 Real trees and the Gromov-Hausdorff-Prokhorov topology

2.2.1 Real trees

We recall the formalism of real trees, see [69]. A metric space (T,d) is a real tree if the following two properties hold for every $x, y \in T$.

(i) (Unique geodesics). There exists a unique isometric map $f_{x,y}\colon [0,d(x,y)] \to T$ such that $f_{x,y}(0) = x$ and $f_{x,y}(d(x,y)) = y$.

(ii) (Loop-free). If φ is a continuous injective map from $[0,1]$ into T such that $\varphi(0) = x$ and $\varphi(1) = y$, then we have

$$\varphi([0,1]) = f_{x,y}\left([0,d(x,y)]\right).$$

A weighted rooted real tree $(T,\emptyset,d,\mu)$ is a real tree (T,d) with a distinguished vertex $\emptyset \in T$ called the root and equipped with a nonnegative finite measure μ. Let us consider a weighted rooted real tree $(T,\emptyset,d,\mu)$. The range of the mapping $f_{x,y}$ described above is denoted by $[\![x,y]\!]$ (this is the line segment between x and y in the tree). In particular, $[\![\emptyset,x]\!]$ is the path going from the root to x which we will interpret as the ancestral line of vertex x. We define a partial order on the tree by setting $x \preccurlyeq y$ (x is an ancestor of y) if and only if $x \in [\![\emptyset,y]\!]$. If $x, y \in T$, there is a unique $z \in T$ such that $[\![\emptyset,x]\!] \cap [\![\emptyset,y]\!] = [\![\emptyset,z]\!]$. We write $z = x \wedge y$ and call it the most recent common ancestor to x and y. For every vertex $x \in T$, we define its height by $H(x) = d(\emptyset,x)$. The height of the tree is defined by $\mathfrak{h}(T) = \sup_{x\in T} H(x)$. Note that if (T,d) is compact, then $\mathfrak{h}(T) < \infty$.

Let $x \in T$ be a vertex. For every $r \in [0,H(x)]$, we denote by $x_r \in T$ the unique ancestor of x at height r. Furthermore, we define the subtree $T_{r,x}$ of T above level r containing x as

$$T_{r,x} = \left\{y \in T\colon\; H(x\wedge y) \geq r\right\}. \tag{2.2.1}$$

Equivalently, $T_{r,x} = \{y \in T\colon\; x_r \preccurlyeq y\}$ is the subtree of T above x_r. Then $T_{r,x}$ can be naturally viewed as a weighted rooted real tree, rooted at x_r and endowed with the distance d and the measure $\mu_{|T_{r,x}}$ (the restriction of μ to $T_{r,x}$). Note that $T_{0,x} = T$. We also define the subtree of T above x by $T_x := T_{H(x),x}$. Denote by

$$\sigma_{r,x}(T) = \mu(T_{r,x}) \quad \text{and} \quad \mathfrak{h}_{r,x}(T) = \mathfrak{h}(T_{r,x}) \tag{2.2.2}$$

the total mass and the height of $T_{r,x}$. For every $\alpha, \beta \geq 0$, we define

$$Z^T_{\alpha,\beta}(x) = \int_0^{H(x)} \sigma_{r,x}(T)^\alpha \mathfrak{h}_{r,x}(T)^\beta \, \mathrm{d}r, \quad \forall x \in T. \tag{2.2.3}$$

We shall omit the dependence on T when there is no ambiguity, simply writing $\sigma_{r,x}$, $\mathfrak{h}_{r,x}$ and $Z_{\alpha,\beta}(x)$. For every $0 \leq r < r' \leq H(x)$, we also introduce the notation

$$T_{[r,r'),x} = \left(T_{r,x} \setminus T_{r',x}\right) \cup \{x_{r'}\} = \{y \in T \colon\ r \leq H(x \wedge y) < r'\} \cup \{x_{r'}\}, \tag{2.2.4}$$

which defines a weighted rooted real tree, equipped with the distance and the measure it inherits from T and naturally rooted at x_r.

The next lemma, whose proof is elementary, relates $\mathfrak{h}_{r,x}(T)$, the height of the subtree of T above level r containing x, to the total height $\mathfrak{h}(T)$.

Lemma 2.2.1. *Let T be a compact real tree. For every $x \in T$ and $r \in [0, H(x)]$, we have*

$$\mathfrak{h}(T) \geq \mathfrak{h}_{r,x}(T) + r. \tag{2.2.5}$$

Furthermore, if $x^* \in T$ is such that $H(x^*) = \mathfrak{h}(T)$, then for every $r \in [0, H(x \wedge x^*)]$, we have

$$\mathfrak{h}(T) = \mathfrak{h}_{r,x}(T) + r. \tag{2.2.6}$$

2.2.2 The Gromov-Hausdorff-Prokhorov topology

We denote by $\mathbb{T}$ the set of (measure-preserving, root-preserving isometry classes of) compact real trees. We will often identify a class with an element of this class. So we shall write $(T, \emptyset, d, \mu) \in \mathbb{T}$ for a weighted rooted compact real tree.

Let us define the Gromov-Hausdorff-Prokhorov (GHP) topology on $\mathbb{T}$. Take two compact real trees $(T, \emptyset, d, \mu), (T', \emptyset', d', \mu') \in \mathbb{T}$. Recall that a correspondence between T and T' is a subset $\mathscr{R} \subset T \times T'$ such that for every $x \in T$, there exists $x' \in T'$ such that $(x, x') \in \mathscr{R}$, and conversely, for every $x' \in T'$, there exists $x \in T$ such that $(x, x') \in \mathscr{R}$. In other words, if we denote by $p \colon T \times T' \to T$ (resp. $p' \colon T \times T' \to T'$) the canonical projection on T (resp. on T'), a correspondence is a subset $\mathscr{R} \subset T \times T'$ such that $p(\mathscr{R}) = T$ and $p'(\mathscr{R}) = T'$. If $\mathscr{R}$ is a correspondence between T and T', its distortion is defined by

$$\operatorname{dis}(\mathscr{R}) = \sup\left\{\left|d(x,y) - d'(x',y')\right| \colon\ (x,x'), (y,y') \in \mathscr{R}\right\}.$$

Next, for any nonnegative finite measure m on $T \times T'$, we define its discrepancy with respect to μ and μ' by

$$\mathrm{D}(m; \mu, \mu') = d_{\mathrm{TV}}(m \circ p^{-1}, \mu) + d_{\mathrm{TV}}(m \circ p'^{-1}, \mu'),$$

where d_{TV} denotes the total variation distance. Then the GHP distance between T and T' is defined as

$$d_{\mathrm{GHP}}(T,T') = \inf\left\{\frac{1}{2}\operatorname{dis}(\mathscr{R}) \vee \mathrm{D}(m;\mu,\mu') \vee m(\mathscr{R}^c)\right\}, \tag{2.2.7}$$

where the infimum is taken over all correspondences $\mathscr{R}$ between T and T' such that $(\emptyset,\emptyset') \in \mathscr{R}$ and all nonnegative finite measures m on $T \times T'$. It can be verified that d_{GHP} is indeed a distance on $\mathbb{T}$ and that the space $(\mathbb{T}, d_{\mathrm{GHP}})$ is complete and separable, see e.g. [13].

The next lemma gives an upper bound for the GHP distance between a tree $(T,\emptyset,d,\mu) \in \mathbb{T}$ and the tree $(T,\emptyset,ad,b\mu)$ obtained from T by multiplying all distances by $a>0$ and the measure μ by $b>0$. The proof is elementary and is left to the reader.

Lemma 2.2.2. *For every $T \in \mathbb{T}$ and $a,b>0$, we have*

$$d_{\mathrm{GHP}}\big((T,\emptyset,d,\mu),(T,\emptyset,ad,b\mu)\big) \le 2|a-1|\mathfrak{h}(T) + |b-1|\mu(T). \tag{2.2.8}$$

2.3 Preleminary results on general compact Lévy trees and stable trees

2.3.1 Two decompositions of the general Lévy tree

Although in this paper we are only interested in the stable case $\psi(\lambda) = \lambda^\gamma$, we state the results of this section in the general Lévy case. Let $\mathscr{T}$ denote a Lévy tree under its excursion measure $\mathbf{N}$ associated with a branching mechanism

$$\psi(\lambda) = a\lambda + b\lambda^2 + \int_0^\infty (\mathrm{e}^{-\lambda r} - 1 + \lambda r)\,\pi(\mathrm{d}r) \tag{2.3.1}$$

where $a,b \ge 0$ and π is a σ-finite measure on $(0,\infty)$ satisfying $\int_0^\infty (r \wedge r^2)\,\pi(\mathrm{d}r) < \infty$. We further assume that $\int^\infty \mathrm{d}\lambda/\psi(\lambda) < \infty$ so that the Lévy tree is compact.

Remark 2.3.1. The Brownian case $\psi(\lambda) = \lambda^2$ corresponds to $a=0$, $b=1$ and $\pi=0$ while the non-Brownian stable case $\psi(\lambda) = \lambda^\gamma$ with $\gamma \in (1,2)$ corresponds to $a=b=0$ and

$$\pi(\mathrm{d}r) = \frac{\gamma(\gamma-1)}{\Gamma(2-\gamma)}\frac{\mathrm{d}r}{r^{1+\gamma}}. \tag{2.3.2}$$

We shall need Bismut's decomposition of the stable tree on several occasions. This is a decomposition of the tree along the ancestral line of a uniformly chosen leaf. We refer the reader to [58, Theorem 4.5] and [3, Theorem 2.1] for more details. We will also need the probability measure $\mathbb{P}_r$ on $\mathbb{T}$ which is the distribution of the Lévy tree starting from an initial mass $r>0$. More precisely, take $\sum_{i\in I}\delta_{\mathscr{T}_i}$ a Poisson point measure on $\mathbb{T}$ with intensity $r\,\mathbf{N}$ and define $\mathbb{P}_r$ as the distribution of the random tree $\mathscr{T}$ obtained by gluing together the trees $\mathscr{T}_i$ at their root. See [3, Section 2.6] for further details.

Before stating the result, we first introduce some notations. Let $(T,\emptyset,d,\mu)$ be a (class representative of a)compact real tree and let $x\in T$. Denote by $(x_i,\, i\in I_x)$ the branching points of T which lie on the branch $[\![\emptyset,x]\!]$, that is those points $y\in[\![\emptyset,x]\!]$ such that $T\setminus\{y\}$ has at least three connected components. For every $i\in I_x$, define the tree grafted on the branch $[\![\emptyset,x]\!]$ at x_i by $T_i=\{y\in T:\, x\wedge y=x_i\}$. We consider T_i as an element of $\mathbb{T}$ in the obvious way. Let $h_i=H(x_i)$ and define a point measure on $[0,\infty)\times\mathbb{T}$ by

$$\mathscr{M}_x^T=\sum_{i\in I_x}\delta_{(h_i,T_i)}.$$

We can now state Bismut's decomposition, see [58, thm 4.5] or [3, Theorem 2.1].

Theorem 2.3.2. *Let $\mathscr{T}$ be the Lévy tree with a general branching mechanism* (2.3.1) *satisfying the Grey condition $\int^\infty \mathrm{d}\lambda/\psi(\lambda)<\infty$ under its excursion measure* $\mathbf{N}$. *For every $\lambda\geq 0$ and every nonnegative measurable function Φ on $[0,\infty)\times\mathbb{T}$, we have*

$$\mathbf{N}\left[\int_{\mathscr{T}}\mu(\mathrm{d}x)\mathrm{e}^{-\lambda H(x)-\langle\mathscr{M}_x^{\mathscr{T}},\Phi\rangle}\right]=\int_0^\infty \mathrm{d}t\,\mathrm{e}^{-(\lambda+a)t}\,\mathbb{E}\left[\mathrm{e}^{-\sum_{0\leq s\leq t}\Phi(s,\mathsf{T}_s)}\right],\tag{2.3.3}$$

where $(\mathsf{T}_s,\, 0\leq s\leq t)$ is a Poisson point process with intensity $\mathbf{N}_{\mathrm{B}}[\mathrm{d}\mathscr{T}]=2b\mathbf{N}[\mathrm{d}\mathscr{T}]+\int_0^\infty r\,\pi(\mathrm{d}r)\,\mathbb{P}_r(\mathrm{d}\mathscr{T})$.

Remark 2.3.3. Bismut's decomposition states the following: let $\mathscr{T}$ be the Lévy tree under its excursion measure $\mathbf{N}$ and, conditionally on $\mathscr{T}$, let U be a leaf chosen uniformly at random, *i.e.* according to the distribution $\sigma^{-1}\mu$. Then, under $\mathbf{N}[\sigma\bullet]$, the random variable $H(U)$ has "distribution" $\mathrm{e}^{-at}\mathrm{d}t$ on $(0,\infty)$ and, conditionally on $H(U)=t$, the point measure $\mathscr{M}_U^{\mathscr{T}}$ is distributed as $\sum_{s\leq t}\delta_{(s,\mathsf{T}_s)}$. One can make this claim rigorous by introducing the space of compact weighted rooted real trees with an additional marked vertex and considering the semidirect product measure $\mathbf{N}\times\sigma^{-1}\mu$ on it which corresponds to the distribution of the pair $(\mathscr{T},U)$. Under this measure, the distribution of the random pair $(H(U),\mathscr{M}_U^{\mathscr{T}})$ does not depend on the particular choice of representative in the class of $\mathscr{T}$.

As a first application of Bismut's decomposition, we give a decomposition of the Lévy tree into $n+1$ subtrees which generalizes [7, Lemma 6.1].

Theorem 2.3.4. *Let $\mathscr{T}$ be the Lévy tree with a general branching mechanism* (2.3.1) *under its excursion measure* $\mathbf{N}$. *Then for every $n\geq 1$ and all nonnegative measurable functions f_i, $1\leq i\leq n+1$ defined on $[0,\infty)\times\mathbb{T}$, we have with $r_0=0$ and $r_{n+1}=H(x)$*

$$\mathbf{N}\left[\int_{\mathscr{T}}\mu(\mathrm{d}x)\int_{0<r_1<\ldots<r_n<H(x)}\prod_{i=1}^{n+1}f_i\left(r_i-r_{i-1},\mathscr{T}_{[r_{i-1},r_i),x}\right)\prod_{i=1}^{n}\mathrm{d}r_i\right]$$
$$=\prod_{i=1}^{n+1}\mathbf{N}\left[\int_{\mathscr{T}}\mu(\mathrm{d}x)f_i(H(x),\mathscr{T})\right].\tag{2.3.4}$$

Proof. Recall from (2.3.7) the definition of $\mathsf{T}^{\downarrow}$. By Theorem 2.3.2, we have

$$\mathbf{N}\left[\int_{\mathscr{T}}\mu(\mathrm{d}x)\int_{0<r_1<\ldots<r_n<H(x)}\prod_{i=1}^{n+1}f_i\left(r_i-r_{i-1},\mathscr{T}_{[r_{i-1},r_i),x}\right)\prod_{i=1}^{n}\mathrm{d}r_i\right]$$
$$=\int_0^{\infty}\mathrm{d}r_{n+1}\,\mathrm{e}^{-ar_{n+1}}\,\mathbb{E}\left[\int_{0<r_1<\ldots<r_n<r_{n+1}}\prod_{i=1}^{n+1}f_i\left(r_i-r_{i-1},\mathsf{T}_{[r_{i-1},r_i)}\right)\prod_{i=1}^{n}\mathrm{d}r_i\right],$$

where we set $\mathsf{T}_{[r,r')}=(\mathsf{T}^{\downarrow}_{t-r}\setminus\mathsf{T}^{\downarrow}_{t-r'})\cup\{t-r'\}$ for every $0<r<r'<t$. Since $(\mathsf{T}_s,0\le s\le t)$ is a Poisson point process, we get that the $\mathsf{T}_{[r_{i-1},r_i)}$ are independent and distributed as $\mathsf{T}_{[0,r_i-r_{i-1})}$. We deduce that

$$\mathbf{N}\left[\int_{\mathscr{T}}\mu(\mathrm{d}x)\int_{0<r_1<\ldots<r_n<H(x)}\prod_{i=1}^{n+1}f_i\left(r_i-r_{i-1},\mathscr{T}_{[r_{i-1},r_i),x}\right)\prod_{i=1}^{n}\mathrm{d}r_i\right]$$
$$=\int_{0<r_1<\ldots<r_n<r_{n+1}}\prod_{i=1}^{n+1}\mathrm{e}^{-a(r_i-r_{i-1})}\,\mathbb{E}\left[f_i\left(r_i-r_{i-1},\mathsf{T}_{[0,r_i-r_{i-1})}\right)\right]\mathrm{d}r_i$$
$$=\int_{[0,\infty)^{n+1}}\prod_{i=1}^{n+1}\mathrm{e}^{-as_i}\,\mathbb{E}\left[f_i\left(s_i,\mathsf{T}_{[0,s_i)}\right)\right]\mathrm{d}s_i$$
$$=\prod_{i=1}^{n+1}\mathbf{N}\left[\int_{\mathscr{T}}\mu(\mathrm{d}x)f_i(H(x),\mathscr{T})\right],$$

where we made the change of variables $(s_1,s_2,\ldots,s_{n+1})=(r_1,r_2-r_1,\ldots,r_{n+1}-r_n)$ for the second equality and used Bismut's decomposition (2.3.12) together with the fact that $\mathsf{T}_{[0,t)}=\mathsf{T}^{\downarrow}_t$ $\mathbb{P}$-a.s. for the last. □

2.3.2 The stable tree and its scaling property

Here, we define the stable tree and recall some of its properties. We refer to [58] for background. We shall work with the stable tree $\mathscr{T}$ with branching mechanism $\psi(\lambda)=\lambda^{\gamma}$ where $\gamma\in(1,2]$ under its excursion measure $\mathbf{N}$: more explicitly, using the coding of compact real trees by height functions, one can define a σ-finite measure $\mathbf{N}$ on $\mathbb{T}$ with the following properties.

(i) **Mass measure.** $\mathbf{N}$-a.e. the mass measure μ is supported by the set of leaves $\mathrm{Lf}(\mathscr{T}):=\{x\in\mathscr{T}:\mathscr{T}\setminus\{x\}\text{ is connected}\}$ and the distribution on $(0,\infty)$ of the total mass $\sigma:=\mu(\mathscr{T})$ is given by

$$\mathbf{N}[\sigma\in\mathrm{d}a]=\frac{1}{\gamma\Gamma(1-1/\gamma)}\frac{\mathrm{d}a}{a^{1+1/\gamma}}.$$

(ii) **Height.** $\mathbf{N}$-a.e. there exists a unique leaf $x^*\in\mathscr{T}$ realizing the height, that is $H(x^*)=\mathfrak{h}(\mathscr{T})$, and the distribution on $(0,\infty)$ of the height $\mathfrak{h}:=\mathfrak{h}(\mathscr{T})$ is given by

$$\mathbf{N}[\mathfrak{h}\in\mathrm{d}a]=(\gamma-1)^{-\gamma/(\gamma-1)}\frac{\mathrm{d}a}{a^{\gamma/(\gamma-1)}}.$$

We will make extensive use of the scaling property of the stable tree under $\mathbf{N}$. Recall from (2.1.2) the definition of R_γ and note that if T has total mass σ and height $\mathfrak{h}$ then $R_\gamma(T,a)$ has total mass $a^{\gamma/(\gamma-1)}\sigma$ and height $a\mathfrak{h}$. Furthermore, it is straightforward to show that for all $x \in T$, $r \in [0, H(x)]$ and $a > 0$:

$$\begin{aligned} \sigma_{ar,x}(R_\gamma(T,a)) &= a^{\gamma/(\gamma-1)}\sigma_{r,x}(T), \\ \mathfrak{h}_{ar,x}(R_\gamma(T,a)) &= a\mathfrak{h}_{r,x}(T), \\ Z_{\alpha,\beta}^{R_\gamma(T,a)}(x) &= a^{\alpha\gamma/(\gamma-1)+\beta+1} Z_{\alpha,\beta}^{T}(x). \end{aligned} \tag{2.3.5}$$

The scaling property of the stable tree can be written as follows:

$$R_\gamma(\mathscr{T}, a) \quad \text{under } \mathbf{N} \quad \overset{(d)}{=} \quad \mathscr{T} \quad \text{under } a^{1/(\gamma-1)}\mathbf{N}, \tag{2.3.6}$$

see e.g. [60, Eq. (40)]. Using this, one can define a regular conditional probability measure $\mathbf{N}^{(a)} = \mathbf{N}[\bullet|\sigma = a]$ such that $\mathbf{N}^{(a)}$-a.s. $\sigma = a$ and

$$\mathbf{N}[\bullet] = \frac{1}{\gamma\Gamma(1-1/\gamma)}\int_0^\infty \mathbf{N}^{(a)}[\bullet]\,\frac{\mathrm{d}a}{a^{1+1/\gamma}}.$$

Informally, $\mathbf{N}^{(a)}$ can be seen as the distribution of the stable tree $\mathscr{T}$ with total mass a.

The next result is a restatement of [76, Proposition 5.7] in terms of trees which gives a version of the scaling property for the stable tree conditioned on its total mass. Recall from (2.1.3) the definition of norm_γ.

Lemma 2.3.5. *Let $\mathscr{T}$ be the stable tree with branching mechanism $\psi(\lambda) = \lambda^\gamma$ where $\gamma \in (1,2]$.*

(i) For every measurable function $F\colon \mathbb{T} \to [0,\infty]$, we have

$$\mathbf{N}^{(1)}\left[F(\mathscr{T})\right] = \Gamma(1-1/\gamma)\,\mathbf{N}\left[\mathbf{1}_{\{\sigma>1\}}F(\mathsf{norm}_\gamma(\mathscr{T}))\right].$$

(ii) Under $\mathbf{N}^{(a)}$, the random tree $\mathscr{T}$ is distributed as $R_\gamma(\mathscr{T}, a^{1-1/\gamma})$ under $\mathbf{N}^{(1)}$ for every $a > 0$.

2.3.3 Preliminary results on the stable tree

Let $(\mathsf{T}_s, 0 \le s \le t)$ be a Poisson point process on $\mathbb{T}$ with intensity $\mathbf{N}_\mathrm{B}$ given by

$$\mathbf{N}_\mathrm{B}[\mathrm{d}\mathscr{T}] = \begin{cases} 2\,\mathbf{N}[\mathrm{d}\mathscr{T}] & \text{if } \gamma = 2, \\ \int_0^\infty r\pi(\mathrm{d}r)\,\mathbb{P}_r(\mathrm{d}\mathscr{T}) & \text{if } \gamma \in (1,2), \end{cases}$$

and denote by

$$\mathsf{T}_r^{\downarrow} := [t-r, t] \circledast_{t-r\le s\le t} (\mathsf{T}_s, s), \quad \forall 0 \le r \le t \tag{2.3.7}$$

the random real tree obtained by grafting T_s on a branch $[t-r,t]$ at height s for every $t-r \le s \le t$ and rooted at $t-r$, see Figure 2.3. We refer the reader to [3, Section 2.4] for a precise definition of the grafting procedure. Let

$$\tau_r := \mu(\mathsf{T}_r^{\downarrow}) = \sum_{t-r\le s\le t} \mu(\mathsf{T}_s) \quad \text{and} \quad \eta_r := \mathfrak{h}(\mathsf{T}_r^{\downarrow}) = \max_{t-r\le s\le t} (\mathfrak{h}(\mathsf{T}_s) + s - (t-r)) \tag{2.3.8}$$

denote its mass and height. Finally, let

$$S_r := \sum_{s\le r} \mu(\mathsf{T}_s). \tag{2.3.9}$$

It is shown in the proof of [52, Lemma 4.6], see Section 8.6 and more precisely (8.20) therein, that in the stable case $\psi(\lambda) = \lambda^\gamma$, both τ and S are subordinators defined on $[0,t]$ with Laplace exponent

$$\varphi(\lambda) = \gamma\lambda^{1-1/\gamma}. \tag{2.3.10}$$

In particular, thanks to [154, Section 4] or [155, Eq. (2.1.8)], we have for every $p \in (-\infty, 1-1/\gamma)$,

$$\mathbb{E}\left[\tau_1^p\right] < \infty. \tag{2.3.11}$$

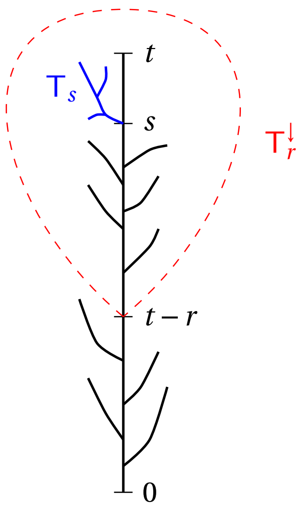

Figure 2.3 – The real tree $\mathsf{T}_r^{\downarrow}$ obtained by grafting the atoms T_s of a Poisson point process on a branch $[t-r,t]$ at height s.

We now give the following form of Bismut's decomposition which we will use throughout the paper. Denote by $D[0,\infty)$ the space of cadlag functions on $[0,\infty)$ endowed with the Skorokhod $J1$ topology. By Theorem 2.3.2 we have, for every measurable function $F\colon [0,\infty)^3 \times \mathbb{T} \times D[0,\infty)^2 \to [0,\infty]$,

$$\mathbf{N}\left[\int_{\mathcal{T}} \mu(\mathrm{d}x) F\big(H(x), \sigma, \mathfrak{h}, \mathcal{T}, \big(\sigma_{H(x)-r,x}, 0 \le r \le H(x)\big), \big(\mathfrak{h}_{H(x)-r,x}, 0 \le r \le H(x)\big)\big)\right]$$
$$= \int_0^\infty \mathrm{d}t\, \mathbb{E}\left[F\Big(t, \tau_t, \eta_t, \mathrm{T}_t^{\downarrow}, (\tau_r, 0 \le r \le t), \big(\eta_r, 0 \le r \le t\big)\Big)\right]. \tag{2.3.12}$$

Notice that by definition $\tau_t = S_t$ and $S_{r-} = \tau_t - \tau_{t-r}$ for every $r \in [0, t]$. This will be used implicitly in the sequel. In particular, the following computation will be useful

$$\int_0^\infty \mathbb{E}\left[\frac{1}{S_t}\mathbf{1}_{\{S_t > 1\}}\right] \mathrm{d}t = \int_0^\infty \mathbb{E}\left[\frac{1}{\tau_t}\mathbf{1}_{\{\tau_t > 1\}}\right] \mathrm{d}t = \mathbf{N}[\sigma > 1] = \frac{1}{\Gamma(1 - 1/\gamma)}, \tag{2.3.13}$$

where in the last equality we used Lemma 2.3.5-(i) with $F \equiv 1$.

Next, as an application of Theorem 2.3.4, we give the decomposition of the *normalized* stable tree into $n+1$ subtrees. For functions f, g defined on $(0, \infty)$, we denote by $f * g$ their convolution defined by

$$f * g(t) = \int_0^t f(s) g(t-s)\, \mathrm{d}s, \quad \forall t > 0.$$

Proposition 2.3.6. *Let $\mathcal{T}$ be the stable tree with branching mechanism $\psi(\lambda) = \lambda^\gamma$ where $\gamma \in (1, 2]$. For every $n \ge 1$ and all nonnegative measurable functions f_i, $1 \le i \le n+1$ defined on $[0, \infty) \times \mathbb{T}$, we have with $r_0 = 0$ and $r_{n+1} = H(x)$*

$$\mathbf{N}^{(1)}\left[\int_{\mathcal{T}} \mu(\mathrm{d}x) \int_{0 < r_1 < \ldots < r_n < H(x)} \prod_{i=1}^{n+1} f_i\left(r_i - r_{i-1}, \mathcal{T}_{[r_{i-1}, r_i), x}\right) \prod_{i=1}^{n} \mathrm{d}r_i\right]$$
$$= \frac{1}{\gamma^n \Gamma(1 - 1/\gamma)^n} F_1 * \cdots * F_{n+1}(1), \tag{2.3.14}$$

where R_γ is defined in (2.1.2) and

$$F_i(a) = a^{-1/\gamma} \mathbf{N}^{(1)}\left[\int_{\mathcal{T}} \mu(\mathrm{d}x) f_i\left(a^{1-1/\gamma} H(x), R_\gamma\left(\mathcal{T}, a^{1-1/\gamma}\right)\right)\right], \quad \forall a > 0.$$

In particular, for every $n \ge 1$ and all nonnegative measurable functions g_i, $1 \le i \le n+1$ defined on $[0, \infty) \times [0, 1]$, we have

$$\mathbf{N}^{(1)}\left[\int_{\mathcal{T}} \mu(\mathrm{d}x) \int_{0 < r_1 < \ldots < r_n < H(x)} \prod_{i=1}^{n+1} g_i\left(r_i - r_{i-1}, \sigma_{r_{i-1}, x} - \sigma_{r_i, x}\right) \prod_{i=1}^{n} \mathrm{d}r_i\right]$$
$$= \frac{1}{\gamma^n \Gamma(1 - 1/\gamma)^n} G_1 * \cdots * G_{n+1}(1), \tag{2.3.15}$$

where

$$G_i(a) = a^{-1/\gamma} \mathbf{N}^{(1)}\left[\int_{\mathcal{T}} \mu(\mathrm{d}x) g_i\left(a^{1-1/\gamma} H(x), a\right)\right], \quad \forall a > 0.$$

Proof. Let $f_i : [0, \infty) \times \mathbb{T} \to \mathbb{R}$ be continuous and bounded for $1 \le i \le n+1$. By Theorem 2.3.4, we have for $\lambda > 0$

$$
\begin{aligned}
&\prod_{i=1}^{n+1} \mathbf{N}\left[\mathrm{e}^{-\lambda\sigma}\int_{\mathcal{T}} \mu(\mathrm{d}x) f_i(H(x),\mathcal{T})\right] \\
&\quad = \mathbf{N}\left[\int_{\mathcal{T}} \mu(\mathrm{d}x) \int_{0<r_1<\ldots<r_n<H(x)} \prod_{i=1}^{n+1} \mathrm{e}^{-\lambda\mu(\mathcal{T}_{[r_{i-1},r_i),x})} f_i\left(r_i-r_{i-1},\mathcal{T}_{[r_{i-1},r_i),x}\right) \prod_{i=1}^{n} \mathrm{d}r_i\right] \\
&\quad = \mathbf{N}\left[\mathrm{e}^{-\lambda\sigma}\int_{\mathcal{T}} \mu(\mathrm{d}x) \int_{0<r_1<\ldots<r_n<H(x)} \prod_{i=1}^{n+1} f_i\left(r_i-r_{i-1},\mathcal{T}_{[r_{i-1},r_i),x}\right) \prod_{i=1}^{n} \mathrm{d}r_i\right]. \qquad (2.3.16)
\end{aligned}
$$

Disintegrating with respect to σ and using the scaling property from Lemma 2.3.5-(ii), we have

$$
\begin{aligned}
&\mathbf{N}\left[\mathrm{e}^{-\lambda\sigma}\int_{\mathcal{T}} \mu(\mathrm{d}x) f_i(H(x),\mathcal{T})\right] \\
&\qquad = \frac{1}{\gamma\Gamma(1-1/\gamma)}\int_0^\infty \mathrm{e}^{-\lambda a}\mathbf{N}^{(a)}\left[\int_{\mathcal{T}} \mu(\mathrm{d}x) f_i(H(x),\mathcal{T})\right]\frac{\mathrm{d}a}{a^{1+1/\gamma}} \\
&\qquad = \frac{1}{\gamma\Gamma(1-1/\gamma)}\mathscr{L}F_i(\lambda), \qquad (2.3.17)
\end{aligned}
$$

where $\mathscr{L}$ denotes the Laplace transform on $[0,\infty)$.

On the other hand, again disintegrating with respect to σ, we have

$$
\begin{aligned}
&\gamma\Gamma(1-1/\gamma)\mathbf{N}\left[\mathrm{e}^{-\lambda\sigma}\int_{\mathcal{T}} \mu(\mathrm{d}x) \int_{0<r_1<\ldots<r_n<H(x)} \prod_{i=1}^{n+1} f_i\left(r_i-r_{i-1},\mathcal{T}_{[r_{i-1},r_i),x}\right) \prod_{i=1}^{n} \mathrm{d}r_i\right] \\
&= \int_0^\infty \frac{\mathrm{d}a}{a^{1+1/\gamma}}\mathrm{e}^{-\lambda a}\mathbf{N}^{(a)}\left[\int_{\mathcal{T}} \mu(\mathrm{d}x) \int_{0<r_1<\ldots<r_n<H(x)} \prod_{i=1}^{n+1} f_i\left(r_i-r_{i-1},\mathcal{T}_{[r_{i-1},r_i),x}\right) \prod_{i=1}^{n} \mathrm{d}r_i\right] \\
&= \int_0^\infty \mathrm{d}a\, a^{(n+1)(1-1/\gamma)-1}\mathrm{e}^{-\lambda a}F(a), \qquad (2.3.18)
\end{aligned}
$$

where we set

$$
\begin{aligned}
F(a) = \mathbf{N}^{(1)}\Bigg[\int_{\mathcal{T}} \mu(\mathrm{d}x) &\\
&\int_{0<r_1<\ldots<r_n<H(x)} \prod_{i=1}^{n+1} f_i\left(a^{1-1/\gamma}(r_i-r_{i-1}), R_\gamma\left(\mathcal{T}_{[r_{i-1},r_i),x}, a^{1-1/\gamma}\right)\right) \prod_{i=1}^{n} \mathrm{d}r_i\Bigg].
\end{aligned}
$$

Putting together (2.3.16)–(2.3.18) yields

$$
\begin{aligned}
\frac{1}{\gamma^n\Gamma(1-1/\gamma)^n}\mathscr{L}(F_1*\ldots*F_{n+1})(\lambda) &= \frac{1}{\gamma^n\Gamma(1-1/\gamma)^n}\prod_{i=1}^{n+1}\mathscr{L}F_i(\lambda) \\
&= \int_0^\infty \mathrm{d}a\, a^{(n+1)(1-1/\gamma)-1}\mathrm{e}^{-\lambda a}F(a).
\end{aligned}
$$

Since this holds for every $\lambda > 0$, we deduce that da-a.e. on $(0,\infty)$,

$$\frac{1}{\gamma^n \Gamma(1-1/\gamma)^n} F_1 * \ldots * F_{n+1}(a) = a^{(n+1)(1-1/\gamma)-1} F(a). \tag{2.3.19}$$

Thanks to Lemma 2.2.2, the mapping $a \mapsto R_\gamma(T, a^{1-1/\gamma})$ is continuous on $(0,\infty)$ for every $T \in \mathbb{T}$. We deduce from the dominated convergence theorem that the F_i are continuous on $(0,\infty)$ and thus $F_1 * \ldots * F_{n+1}$ too. Similarly, the right-hand side of (2.3.19) is continuous with respect to a. Therefore the equality holds for every $a \in (0,\infty)$. In particular, taking $a = 1$ proves (2.3.14) for continuous bounded functions $f_i\colon [0,\infty) \times \mathbb{T} \to \mathbb{R}$. This extends to measurable functions $f_i\colon [0,\infty) \times \mathbb{T} \to \mathbb{R}$ thanks to the monotone class theorem. Finally, (2.3.15) is a direct consequence of (2.3.14). □

In particular, the following corollary will be useful.

Corollary 2.3.7. *We have*

$$\sup_{\alpha \geq 0} \alpha^{2-2/\gamma} \mathbf{N}^{(1)} \left[\int_{\mathcal{T}} \mu(\mathrm{d}x) \left(\int_0^{H(x)} \sigma_{r,x}^\alpha \, \mathrm{d}r \right)^2 \right] < \infty. \tag{2.3.20}$$

Proof. Applying (2.3.15) with $n = 2$, $g_1(r,a) = g(1-a)$, $g_2(r,a) = 1$ and $g_3(r,a) = g(a)$ yields, for every measurable function $g\colon [0,1] \to [0,\infty]$,

$$\mathbf{N}^{(1)} \left[\int_{\mathcal{T}} \mu(\mathrm{d}x) \left(\int_0^{H(x)} g(\sigma_{r,x}) \, \mathrm{d}r \right)^2 \right] \\ = \frac{2}{\gamma^2 \Gamma(1-1/\gamma)^2} \int_0^1 g(y)(1-y)^{-1/\gamma} \, \mathrm{d}y \int_0^y g(z) z^{-1/\gamma} (y-z)^{-1/\gamma} \, \mathrm{d}z. \tag{2.3.21}$$

Taking $g(a) = a^\alpha$, we get

$$\begin{aligned} \alpha^{2-2/\gamma} \mathbf{N}^{(1)} &\left[\int_{\mathcal{T}} \mu(\mathrm{d}x) \left(\int_0^{H(x)} \sigma_{r,x}^\alpha \, \mathrm{d}r \right)^2 \right] \\ &= \frac{2\alpha^{2-2/\gamma}}{\gamma^2 \Gamma(1-1/\gamma)^2} \int_0^1 y^\alpha (1-y)^{-1/\gamma} \, \mathrm{d}y \int_0^y z^{\alpha-1/\gamma} (y-z)^{-1/\gamma} \, \mathrm{d}z \\ &= \frac{2\alpha^{2-2/\gamma}}{\gamma^2 \Gamma(1-1/\gamma)^2} \mathrm{B}\big(2\alpha + 2 - 2/\gamma, 1 - 1/\gamma\big) \mathrm{B}\big(\alpha + 1 - 1/\gamma, 1 - 1/\gamma\big), \end{aligned}$$

where B is the Beta function. Using that $\mathrm{B}(x, 1-1/\gamma) \sim \Gamma(1-1/\gamma) x^{-1+1/\gamma}$ as $x \to \infty$, (2.3.20) readily follows. □

As a consequence of Proposition 2.3.6, we are able to compute the intensity measure of the random measure $\Psi_{\mathcal{T}}$ appearing in [7], see Proposition 6.3 therein.

Corollary 2.3.8. *Let $\mathcal{T}$ be the normalized stable tree with branching mechanism $\psi(\lambda) = \lambda^\gamma$ where $\gamma \in (1,2]$. Let f and g be nonnegative measurable functions defined on $\mathbb{T}$ and $[0,\infty)$ respectively. We have*

$$\gamma\Gamma(1-1/\gamma)\mathbf{N}^{(1)}\left[\int_{\mathcal{T}} \mu(\mathrm{d}x)\int_0^{H(x)} f(\mathcal{T}_{r,x})g(r)\,\mathrm{d}r\right]$$
$$= \int_0^1 \frac{\mathrm{d}a}{a^{1/\gamma}(1-a)^{1/\gamma}}\mathbf{N}^{(1)}\left[f\circ R_\gamma\left(\mathcal{T}, a^{1-1/\gamma}\right)\right]\mathbf{N}^{(1)}\left[\int_{\mathcal{T}} \mu(\mathrm{d}x) g\left((1-a)^{1-1/\gamma}H(x)\right)\right]. \quad (2.3.22)$$

Another application of Theorem 2.3.2 is the following result giving the moments of the height $H(U)$ of a uniformly distributed leaf $U \in \mathcal{T}$ (*i.e.* according to μ) under $\mathbf{N}^{(1)}$. In particular, this allows to give a nontrivial upper bound for the size of the ball with radius $\varepsilon > 0$ centered around the root of the normalized stable tree. Let us mention that this result is not new since the distribution of $H(U)$ under $\mathbf{N}^{(1)}$ is known: in the Brownian case $\gamma = 2$, H is distributed as $\sqrt{2}e$ where e is the Brownian excursion so $\sqrt{2}H(U)$ has Rayleigh distribution; in the case $\gamma \in (1,2)$, $H(U)$ is distributed as a multiple of the local time at 0 of the Bessel bridge of dimension $2/\gamma$, see [84, Corollary 10].

Lemma 2.3.9. *Let $\mathcal{T}$ be the normalized stable tree with branching mechanism $\psi(\lambda) = \lambda^\gamma$ where $\gamma \in (1,2]$. For every $p \in (-\infty, 2)$, we have*

$$\mathbf{N}^{(1)}\left[\int_{\mathcal{T}} H(x)^{-p}\mu(\mathrm{d}x)\right] = \frac{(\gamma-1)\gamma^{p-1}\Gamma(1-1/\gamma)\Gamma(2-p)}{\Gamma(1-(p-1)(1-1/\gamma))} < \infty. \quad (2.3.23)$$

Proof. Using Bismut's decomposition (2.3.12), we have for every $\lambda > 0$

$$\mathbf{N}\left[\sigma \mathrm{e}^{-\lambda\sigma}\int_{\mathcal{T}} H(x)^{-p}\mu(\mathrm{d}x)\right] = \int_0^\infty t^{-p}\,\mathbb{E}\left[\tau_t \mathrm{e}^{-\lambda\tau_t}\right]\mathrm{d}t = \varphi'(\lambda)\int_0^\infty t^{1-p}\mathrm{e}^{-t\varphi(\lambda)}\,\mathrm{d}t.$$

On the other hand, disintegrating with respect to σ and using Lemma 2.3.5-(ii), we have

$$\begin{aligned}
\mathbf{N}\left[\sigma \mathrm{e}^{-\lambda\sigma}\int_{\mathcal{T}} H(x)^{-p}\mu(\mathrm{d}x)\right] &= \frac{1}{\gamma\Gamma(1-1/\gamma)}\int_0^\infty a\mathrm{e}^{-\lambda a}\mathbf{N}^{(a)}\left[\int_{\mathcal{T}} H(x)^{-p}\mu(\mathrm{d}x)\right]\frac{\mathrm{d}a}{a^{1+1/\gamma}} \\
&= \frac{1}{\gamma\Gamma(1-1/\gamma)}\int_0^\infty \mathrm{e}^{-\lambda a}\frac{\mathrm{d}a}{a^{(p-1)(1-1/\gamma)}}\mathbf{N}^{(1)}\left[\int_{\mathcal{T}} H(x)^{-p}\mu(\mathrm{d}x)\right] \\
&= \frac{\Gamma(1-(p-1)(1-1/\gamma))}{\gamma\Gamma(1-1/\gamma)\lambda^{1-(p-1)(1-1/\gamma)}}\mathbf{N}^{(1)}\left[\int_{\mathcal{T}} H(x)^{-p}\mu(\mathrm{d}x)\right].
\end{aligned}$$

Using (2.3.10), it follows that

$$\begin{aligned}
\mathbf{N}^{(1)}\left[\int_{\mathcal{T}} H(x)^{-p}\mu(\mathrm{d}x)\right] &= \frac{\gamma\Gamma(1-1/\gamma)\lambda^{1-(p-1)(1-1/\gamma)}\varphi'(\lambda)}{\Gamma(1-(p-1)(1-1/\gamma))}\int_0^\infty t^{1-p}\mathrm{e}^{-t\varphi(\lambda)}\,\mathrm{d}t \\
&= \frac{(\gamma-1)\gamma^{p-1}\Gamma(1-1/\gamma)\Gamma(2-p)}{\Gamma(1-(p-1)(1-1/\gamma))}.
\end{aligned}$$

□

Remark 2.3.10. Conditionally on $\mathscr{T}$, let $U \in \mathscr{T}$ be a uniformly distributed leaf. Then we can rewrite (2.3.23) as follows:

$$\frac{1}{c_\gamma}\mathbf{N}^{(1)}\left[\frac{1}{H(U)}\big(\gamma H(U)\big)^p\right] = \frac{\Gamma(p+1)}{\Gamma(p(1-1/\gamma)+1)}, \quad \forall p > -1, \tag{2.3.24}$$

where $c_\gamma = (\gamma-1)\Gamma(1-1/\gamma)$. This implies that, under the probability measure $c_\gamma^{-1}\mathbf{N}^{(1)}[H(U)^{-1}\bullet]$, the random variable $\gamma H(U)$ has Mittag-Leffler distribution with index $1-1/\gamma$, see [137, Eq. (0.42)].

2.4 Zooming in at the root of the stable tree

In this section, we study the shape of the stable tree in a small neighborhood of its root. The main result, Theorem 2.4.2, states that after zooming in and rescaling, one sees a branch on which trees are grafted according to a Poisson point process on $\mathbb{T}$ with intensity $\mathbf{N}_{\mathrm{B}}$ given by

$$\mathbf{N}_{\mathrm{B}}[\mathrm{d}\mathscr{T}] = \begin{cases} 2\,\mathbf{N}[\mathrm{d}\mathscr{T}] & \text{if } \gamma = 2, \\ \int_0^\infty r\pi(\mathrm{d}r)\,\mathbb{P}_r(\mathrm{d}\mathscr{T}) & \text{if } \gamma \in (1,2), \end{cases} \tag{2.4.1}$$

where we recall from Section 2.3 that π is given by (2.3.2) and $\mathbb{P}_r$ is the distribution of the random tree $\mathscr{T}$ obtained by gluing together at their roots a family of trees distributed according to a Poisson point measure with intensity $r\,\mathbf{N}$.

We start with the following result giving the scaling property of the stable tree under $\mathbf{N}_{\mathrm{B}}$.

Lemma 2.4.1. *The following identity holds for every $a > 0$*

$$R_\gamma(\mathscr{T}, a) \quad \textit{under } \mathbf{N}_{\mathrm{B}} \quad \overset{(d)}{=} \quad \mathscr{T} \quad \textit{under } a\mathbf{N}_{\mathrm{B}}\,. \tag{2.4.2}$$

Proof. The case $\gamma = 2$ reduces to the scaling property (2.3.6) so we only need to prove the case $\gamma \in (1,2)$. Thanks to (2.3.6), we deduce that $R_\gamma(\mathscr{T}, a)$ under $\mathbb{P}_r$ has distribution $\mathbb{P}_{a^{1/(\gamma-1)}r}$. It follows from (2.3.2) that under $\mathbf{N}_{\mathrm{B}}$, $R_\gamma(\mathscr{T}, a)$ has distribution

$$\int_0^\infty r\pi(\mathrm{d}r)\,\mathbb{P}_{a^{1/(\gamma-1)}r}(\mathrm{d}\mathscr{T}) = a\int_0^\infty s\pi(\mathrm{d}s)\,\mathbb{P}_s(\mathrm{d}\mathscr{T}) = a\mathbf{N}_{\mathrm{B}}[\mathrm{d}\mathscr{T}].$$

□

Let $(T, \emptyset, d, \mu)$ be a compact real tree and let $x \in T$. Recall from Section 2.3 that T_i, $i \in I_x$ are the trees grafted on the branch $[\![\emptyset, x]\!]$, each one at height h_i. Fix $\mathfrak{f}\colon (0,\infty) \to (0,\infty)$ and define for every $\varepsilon > 0$ a point measure on $[0,\infty)^2 \times \mathbb{T}$ by

$$\mathcal{N}_\varepsilon^{\mathfrak{f}}(x) = \sum_{h_i \le \mathfrak{f}(\varepsilon) H(x)} \delta_{(\varepsilon^{-1} h_i, \varepsilon^{-\gamma/(\gamma-1)}\sigma_i, \mathsf{norm}_\gamma(T_i))}. \tag{2.4.3}$$

We are now in a position to give the main result of this section.

Theorem 2.4.2. *Let $\mathcal{T}$ be the normalized stable tree with branching mechanism $\psi(\lambda) = \lambda^\gamma$ where $\gamma \in (1,2]$. Conditionally on $\mathcal{T}$, let U be a $\mathcal{T}$-valued random variable with distribution μ under $\mathbf{N}^{(1)}$. Let $(\mathsf{T}'_s, s \ge 0)$ be a Poisson point process with intensity $\mathbf{N}_{\mathrm{B}}$, independent of $(\mathcal{T}, H(U))$. Let $\Phi\colon [0,\infty)^2 \times \mathbb{T} \to [0,\infty)$ be a measurable function such that there exists $C > 0$ such that for every $h \ge 0$ and $T \in \mathbb{T}$, we have*

$$|\Phi(h,b,T) - \Phi(h,a,T)| \le C|b-a|. \tag{2.4.4}$$

(i) If $\lim_{\varepsilon\to 0} \varepsilon^{-1/2}\mathfrak{f}(\varepsilon) = 0$ and $\lim_{\varepsilon\to 0}\varepsilon^{-1}\mathfrak{f}(\varepsilon) = \infty$, then we have the following convergence in distribution

$$\left(\mathcal{T}, H(U), \langle \mathcal{N}_\varepsilon^{\mathfrak{f}}(U), \Phi\rangle\right) \xrightarrow[\varepsilon\to 0]{(d)} \left(\mathcal{T}, H(U), \sum_{s\ge 0} \Phi\left(s, \mu(\mathsf{T}'_s), \mathsf{norm}_\gamma\left(\mathsf{T}'_s\right)\right)\right) \tag{2.4.5}$$

in the space $\mathbb{T} \times [0,\infty) \times [0,\infty]$.

(ii) If $\mathfrak{f}(\varepsilon) = \varepsilon$, then we have the following convergence in distribution

$$\left(\mathcal{T}, H(U), \langle \mathcal{N}_\varepsilon^{\mathfrak{f}}(U), \Phi\rangle\right) \xrightarrow[\varepsilon\to 0]{(d)} \left(\mathcal{T}, H(U), \sum_{s\le H(U)} \Phi\left(s, \mu(\mathsf{T}'_s), \mathsf{norm}_\gamma\left(\mathsf{T}'_s\right)\right)\right) \tag{2.4.6}$$

in the space $\mathbb{T} \times [0,\infty) \times [0,\infty]$.

Proof. We only prove (i), the proof of (ii) being similar. Let $f\colon \mathbb{T} \to \mathbb{R}$ and $g\colon [0,\infty) \to \mathbb{R}$ be Lipschitz-continuous and bounded and assume that $\Phi\colon [0,\infty)^2 \times \mathbb{T} \to [0,\infty)$ is measurable and satisfies (2.4.4). We shall consider the following modification of the measure $\mathcal{N}_\varepsilon^{\mathfrak{f}}(U)$:

$$\hat{\mathcal{N}}_\varepsilon^{\mathfrak{f}}(U) := \sum_{h_i \le \mathfrak{f}(\varepsilon) H(U)} \delta_{(\varepsilon^{-1} h_i / H(U), \varepsilon^{-\gamma/(\gamma-1)}\sigma_i, \mathsf{norm}_\gamma(T_i))}.$$

Step 1. Set

$$\begin{aligned}
F(\varepsilon) &:= \mathbf{N}^{(1)}\left[f(\mathcal{T}) g(H(U)) \exp\left\{-\left\langle \hat{\mathcal{N}}_\varepsilon^{\mathfrak{f}}(U), \Phi\right\rangle\right\}\right] \\
&= \mathbf{N}^{(1)}\Bigg[\int_{\mathcal{T}} \mu(\mathrm{d}x) f(\mathcal{T}) g(H(x)) \\
&\qquad \times \exp\left\{-\sum_{h_i \le \mathfrak{f}(\varepsilon) H(x)} \Phi\left(\varepsilon^{-1} h_i / H(x), \varepsilon^{-\gamma/(\gamma-1)}\sigma_i, \mathsf{norm}_\gamma(\mathcal{T}_i)\right)\right\}\Bigg].
\end{aligned}$$

Using Lemma 2.3.5-(i) and Theorem 2.3.2, we have

$$\begin{aligned}\frac{F(\varepsilon)}{\Gamma(1-1/\gamma)} &= \mathbf{N}\Bigg[\frac{1}{\sigma}\mathbf{1}_{\{\sigma>1\}}\int_{\mathscr{T}}\mu(\mathrm{d}x)f\circ\mathrm{norm}_\gamma(\mathscr{T})\,g\left(\sigma^{-1+1/\gamma}H(x)\right)\\ &\qquad\times\exp\Bigg\{-\sum_{h_i\le\mathfrak{f}(\varepsilon)H(x)}\Phi\left(\varepsilon^{-1}h_i/H(x),\varepsilon^{-\gamma/(\gamma-1)}\sigma^{-1}\sigma_i,\mathrm{norm}_\gamma(\mathscr{T}_i)\right)\Bigg\}\Bigg]\\ &= \int_0^\infty \mathrm{d}t\,\mathbb{E}\Bigg[\frac{1}{\tau_t}\mathbf{1}_{\{\tau_t>1\}}f\circ\mathrm{norm}_\gamma\left(\mathsf{T}_t^{\downarrow}\right)g\left(\tau_t^{-1+1/\gamma}t\right)\\ &\qquad\times\exp\Bigg\{-\sum_{s\le\mathfrak{f}(\varepsilon)t}\Phi\left(\varepsilon^{-1}s/t,\varepsilon^{-\gamma/(\gamma-1)}\tau_t^{-1}\mu(\mathsf{T}_s),\mathrm{norm}_\gamma(\mathsf{T}_s)\right)\Bigg\}\Bigg].\end{aligned} \tag{2.4.7}$$

Step 2. The proof of the following lemma is postponed to Section 2.7.1. To simplify notation, we introduce $\mathfrak{g}(\varepsilon)=1-\mathfrak{f}(\varepsilon)$.

Lemma 2.4.3. *Assume that* $\lim_{\varepsilon\to0}\varepsilon^{-1/2}\mathfrak{f}(\varepsilon)=0$. *Let* $f\colon\mathbb{T}\to\mathbb{R}$ *and* $g\colon[0,\infty)\to\mathbb{R}$ *be Lipschitz-continuous and bounded and assume that* $\Phi\colon[0,\infty)^2\times\mathbb{T}\to[0,\infty)$ *is measurable and satisfies* (2.4.4). *We have*

$$\begin{aligned}\lim_{\varepsilon\to0}\Gamma(1-1/\gamma)^{-1}F(\varepsilon)-\int_0^\infty\mathrm{d}t\,\mathbb{E}\Bigg[&\frac{1}{\tau_{\mathfrak{g}(\varepsilon)t}}\mathbf{1}_{\{\tau_{\mathfrak{g}(\varepsilon)t}>1\}}f\circ\mathrm{norm}_\gamma\left(\mathsf{T}^{\downarrow}_{\mathfrak{g}(\varepsilon)t}\right)g\left(\tau_{\mathfrak{g}(\varepsilon)t}^{-1+1/\gamma}t\right)\\ &\times\exp\Bigg\{-\sum_{s\le\mathfrak{f}(\varepsilon)t}\Phi\left(\varepsilon^{-1}s/t,\varepsilon^{-\gamma/(\gamma-1)}\tau^{-1}_{\mathfrak{g}(\varepsilon)t}\mu(\mathsf{T}_s),\mathrm{norm}_\gamma(\mathsf{T}_s)\right)\Bigg\}\Bigg]=0.\end{aligned}$$

Since $(\mathsf{T}_s,0\le s\le t)$ is a Poisson point process, it follows from the definition of $\mathsf{T}^{\downarrow}_{\mathfrak{g}(\varepsilon)t}$ that $(\mathsf{T}_s,0\le s\le\mathfrak{f}(\varepsilon)t)$ is independent of $\mathsf{T}^{\downarrow}_{\mathfrak{g}(\varepsilon)t}$. Thus, denoting by $(\mathsf{T}'_s,s\ge0)$ a Poisson point process with intensity $\mathbf{N}_{\mathrm{B}}$ which is independent of $\mathsf{T}^{\downarrow}_{\mathfrak{g}(\varepsilon)t}$, recalling that $\tau_{\mathfrak{g}(\varepsilon)t}$ is a measurable function of $\mathsf{T}^{\downarrow}_{\mathfrak{g}(\varepsilon)t}$ and making the change of variable $u=\mathfrak{g}(\varepsilon)t$, we have

$$\lim_{\varepsilon\to0}\left|\Gamma(1-1/\gamma)^{-1}F(\varepsilon)-\mathfrak{g}(\varepsilon)^{-1}\int_0^\infty\mathrm{d}u\,\mathbb{E}\left[Y_\varepsilon(u)\right]\right|=0, \tag{2.4.8}$$

where

$$\begin{aligned}Y_\varepsilon(u)=&\frac{1}{\tau_u}\mathbf{1}_{\{\tau_u>1\}}f\circ\mathrm{norm}_\gamma\left(\mathsf{T}_u^{\downarrow}\right)g\left(\mathfrak{g}(\varepsilon)^{-1}\tau_u^{-1+1/\gamma}u\right)\\ &\times\mathbb{E}\left[\exp\Bigg\{-\sum_{s\le\mathfrak{f}(\varepsilon)\mathfrak{g}(\varepsilon)^{-1}u}\Phi\left(\varepsilon^{-1}\mathfrak{g}(\varepsilon)s/u,\varepsilon^{-\gamma/(\gamma-1)}\tau_u^{-1}\mu(\mathsf{T}'_s),\mathrm{norm}_\gamma(\mathsf{T}'_s)\right)\Bigg\}\,\Bigg|\,\mathsf{T}_u^{\downarrow}\right].\end{aligned} \tag{2.4.9}$$

Step 3. For fixed $\lambda>0$, we have

$$\mathbb{E}\left[\exp\Bigg\{-\sum_{s\le\mathfrak{f}(\varepsilon)\mathfrak{g}(\varepsilon)^{-1}u}\Phi\left(\varepsilon^{-1}\mathfrak{g}(\varepsilon)s/u,\varepsilon^{-\gamma/(\gamma-1)}\lambda^{-1}\mu(\mathsf{T}'_s),\mathrm{norm}_\gamma(\mathsf{T}'_s)\right)\Bigg\}\right]$$

$$= \exp\left\{-\int_0^{\mathfrak{f}(\varepsilon)\mathfrak{g}(\varepsilon)^{-1}u} \mathrm{d}s\,\mathbf{N}_{\mathrm{B}}\left[1-\mathrm{e}^{-\Phi(\varepsilon^{-1}\mathfrak{g}(\varepsilon)s/u,\varepsilon^{-\gamma/(\gamma-1)}\lambda^{-1}\sigma,\mathrm{norm}_\gamma(\mathscr{T}))}\right]\right\}$$
$$= \exp\left\{-\mathfrak{g}(\varepsilon)^{-1}\int_0^{\varepsilon^{-1}\mathfrak{f}(\varepsilon)\lambda^{-1+1/\gamma}u} \mathrm{d}r\,\mathbf{N}_{\mathrm{B}}\left[1-\mathrm{e}^{-\Phi(\lambda^{1-1/\gamma}r/u,\sigma,\mathrm{norm}_\gamma(\mathscr{T}))}\right]\right\},$$

where we made the change of variable $r=\varepsilon^{-1}\mathfrak{g}(\varepsilon)\lambda^{-1+1/\gamma}s$ and used Lemma 2.4.1 with $a=\varepsilon\lambda^{1-1/\gamma}$. (Notice that $\mathrm{norm}_\gamma(\mathscr{T})$ has the same distribution under $a\mathbf{N}_{\mathrm{B}}$ for every $a>0$). Thus, we deduce that a.s. for every $u>0$

$$\lim_{\varepsilon\to 0}\mathbb{E}\left[\exp\left\{-\sum_{s\le \mathfrak{f}(\varepsilon)\mathfrak{g}(\varepsilon)^{-1}u}\Phi\left(\varepsilon^{-1}\mathfrak{g}(\varepsilon)s/u,\varepsilon^{-\gamma/(\gamma-1)}\tau_u^{-1}\mu(\mathsf{T}'_s),\mathrm{norm}_\gamma\left(\mathsf{T}'_s\right)\right)\right\}\middle|\mathsf{T}^{\downarrow}_u\right]$$
$$= \lim_{\varepsilon\to 0}\exp\left\{-\mathfrak{g}(\varepsilon)^{-1}\int_0^{\varepsilon^{-1}\mathfrak{f}(\varepsilon)\lambda^{-1+1/\gamma}u}\mathrm{d}r\,\mathbf{N}_{\mathrm{B}}\left[1-\mathrm{e}^{-\Phi(\lambda^{1-1/\gamma}r/u,\sigma,\mathrm{norm}_\gamma(\mathscr{T}))}\right]_{|\lambda=\tau_u}\right\}$$
$$= \exp\left\{-\int_0^{\infty}\mathrm{d}r\,\mathbf{N}_{\mathrm{B}}\left[1-\mathrm{e}^{-\Phi(\lambda^{1-1/\gamma}r/u,\sigma,\mathrm{norm}_\gamma(\mathscr{T}))}\right]_{|\lambda=\tau_u}\right\}$$
$$= \mathbb{E}\left[\exp\left\{-\sum_{s\ge 0}\Phi\left(\tau_u^{1-1/\gamma}s/u,\mu(\mathsf{T}'_s),\mathrm{norm}_\gamma\left(\mathsf{T}'_s\right)\right)\right\}\middle|\mathsf{T}^{\downarrow}_u\right].$$

Step 4. We deduce that a.s. for every $u>0$

$$\lim_{\varepsilon\to 0}Y_\varepsilon(u)=\frac{1}{\tau_u}\mathbf{1}_{\{\tau_u>1\}}f\circ\mathrm{norm}_\gamma\left(\mathsf{T}^{\downarrow}_u\right)g\left(\tau_u^{-1+1/\gamma}u\right)$$
$$\times\,\mathbb{E}\left[\exp\left\{-\sum_{s\ge 0}\Phi\left(\tau_u^{1-1/\gamma}s/u,\mu(\mathsf{T}'_s),\mathrm{norm}_\gamma\left(\mathsf{T}'_s\right)\right)\right\}\middle|\mathsf{T}^{\downarrow}_u\right].\qquad(2.4.10)$$

Since $|Y_\varepsilon(u)|\le\|f\|_\infty\|g\|_\infty\tau_u^{-1}\mathbf{1}_{\{\tau_u>1\}}$ where the right-hand side is integrable with respect to $\mathbf{1}_{(0,\infty)}(u)\,\mathrm{d}u\otimes\mathbb{P}$ thanks to (2.3.13), it follows by dominated convergence that

$$\lim_{\varepsilon\to 0}\int_0^\infty\mathrm{d}u\,\mathbb{E}\left[Y_\varepsilon(u)\right]=\int_0^\infty\mathrm{d}u\,\mathbb{E}\left[\frac{1}{\tau_u}\mathbf{1}_{\{\tau_u>1\}}f\circ\mathrm{norm}_\gamma\left(\mathsf{T}^{\downarrow}_u\right)g\left(\tau_u^{-1+1/\gamma}u\right)\right.$$
$$\left.\times\exp\left\{-\sum_{s\ge 0}\Phi\left(\tau_u^{1-1/\gamma}s/u,\mu(\mathsf{T}'_s),\mathrm{norm}_\gamma\left(\mathsf{T}'_s\right)\right)\right\}\right].\qquad(2.4.11)$$

Step 5. Using Theorem 2.3.2 and Lemma 2.3.5-(i) again, we get that

$$\lim_{\varepsilon\to 0}F(\varepsilon)=\Gamma(1-1/\gamma)\int_0^\infty\mathrm{d}u\,\mathbb{E}\left[\frac{1}{\tau_u}\mathbf{1}_{\{\tau_u>1\}}f\circ\mathrm{norm}_\gamma\left(\mathsf{T}^{\downarrow}_u\right)g\left(\tau_u^{-1+1/\gamma}u\right)\right.$$
$$\left.\times\exp\left\{-\sum_{s\ge 0}\Phi\left(\tau_u^{1-1/\gamma}s/u,\mu(\mathsf{T}'_s),\mathrm{norm}_\gamma\left(\mathsf{T}'_s\right)\right)\right\}\right]$$
$$=\Gamma(1-1/\gamma)\,\mathbf{N}\left[\frac{1}{\sigma}\mathbf{1}_{\{\sigma>1\}}\int_{\mathscr{T}}\mu(\mathrm{d}x)f\circ\mathrm{norm}_\gamma(\mathscr{T})\,g\left(\sigma^{-1+1/\gamma}H(x)\right)\right.$$
$$\left.\times\exp\left\{-\sum_{s\ge 0}\Phi\left(\sigma^{1-1/\gamma}s/H(x),\mu(\mathsf{T}'_s),\mathrm{norm}_\gamma\left(\mathsf{T}'_s\right)\right)\right\}\right]$$

$$= \mathbf{N}^{(1)}\left[\int_{\mathcal{T}} \mu(\mathrm{d}x) f(\mathcal{T}) g(H(x)) \exp\left\{-\sum_{s\geq 0} \Phi\left(s/H(x), \mu(\mathsf{T}'_s), \mathrm{norm}_\gamma\left(\mathsf{T}'_s\right)\right)\right\}\right]$$
$$= \mathbf{N}^{(1)}\left[f(\mathcal{T}) g(H(U)) \exp\left\{-\sum_{s\geq 0} \Phi\left(s/H(U), \mu(\mathsf{T}'_s), \mathrm{norm}_\gamma\left(\mathsf{T}'_s\right)\right)\right\}\right],$$

where, with a slight abuse of notation, we denote by $(\mathsf{T}'_s, s \geq 0)$ a Poisson point process with intensity $\mathbf{N}_{\mathrm{B}}$ under $\mathbf{N}^{(1)}$, independent of $(\mathcal{T}, H(U))$. Since $H(U)$ and $(\mathsf{T}'_s, s \geq 0)$ are independent, this concludes the proof. □

As a consequence of Theorem 2.4.2, the next result gives the asymptotic behavior of the total mass of the subtrees grafted near the root of the stable tree.

Corollary 2.4.4. *Let $\mathcal{T}$ be the normalized stable tree with branching mechanism $\psi(\lambda) = \lambda^\gamma$ where $\gamma \in (1,2]$. Conditionally on $\mathcal{T}$, let U be $\mathcal{T}$-valued random variable with distribution μ under $\mathbf{N}^{(1)}$. Assume that $\lim_{\varepsilon\to 0} \varepsilon^{-1/2}\mathfrak{f}(\varepsilon) = 0$ and $\lim_{\varepsilon\to 0} \varepsilon^{-1}\mathfrak{f}(\varepsilon) = \infty$. Define a process S^ε by*

$$S^\varepsilon_t := \sum_{h_i \leq \varepsilon t \wedge \mathfrak{f}(\varepsilon) H(U)} \varepsilon^{-\gamma/(\gamma-1)} \sigma_i, \quad t \geq 0.$$

Then we have the following convergence in distribution

$$\left(\mathcal{T}, H(U), \left(S^\varepsilon_t, t \geq 0\right)\right) \xrightarrow[\varepsilon\to 0]{(d)} (\mathcal{T}, H(U), (S_t, t \geq 0)) \tag{2.4.12}$$

in the space $\mathbb{T} \times \mathbb{R} \times D[0,\infty)$, where S is a stable subordinator with Laplace exponent φ given by (2.3.10), independent of $(\mathcal{T}, H(U))$.

Proof. We adapt the arguments of [140, Chapter VII, Section 7.2], see also Theorem 3.1 and Corollary 3.4 in [150]. Since the process S has no fixed points of discontinuity, it is enough to show that the convergence (2.4.12) holds in $\mathbb{T} \times \mathbb{R} \times D[0, r]$ for every $r > 0$.

Fix $r > 0$ and let $\delta > 0$. Define

$$S^{\varepsilon,\delta}_t := \sum_{h_i \leq \varepsilon t \wedge \mathfrak{f}(\varepsilon) H(U)} \varepsilon^{-\gamma/(\gamma-1)} \sigma_i \mathbf{1}_{\{\varepsilon^{-\gamma/(\gamma-1)}\sigma_i > \delta\}}, \quad t \geq 0.$$

Recall that for a metric space X, we denote by $\mathcal{M}_p(X)$ the space of point measures on X equipped with the topology of vague convergence. It is known (see [140, p. 215]) that the restriction mapping

$$m \mapsto m_{|[0,\infty)\times(\delta,\infty)}$$

is a.s. continuous from $\mathcal{M}_p([0,\infty)^2)$ to $\mathcal{M}_p([0,\infty) \times (\delta,\infty))$ with respect to the distribution of the Poisson random measure $\sum_{s\geq 0} \delta_{(s,\mu(\mathsf{T}'_s))}$. Furthermore, the summation mapping

$$m \mapsto \left(\int_{[0,t]\times(\delta,\infty)} x\, m(\mathrm{d}s, \mathrm{d}x), 0 \leq t \leq r\right)$$

is a.s. continuous from $\mathscr{M}_p([0,\infty)\times(\delta,\infty))$ to $D[0,r]$ with respect to the same distribution. We deduce from Theorem 2.4.2-(i) and the continuous mapping theorem the following convergence in distribution

$$\left(\mathscr{T},H(U),\left(S_t^{\varepsilon,\delta},0\le t\le r\right)\right)\xrightarrow[\varepsilon\to 0]{(d)}\left(\mathscr{T},H(U),\left(\sum_{s\le t}\mu(\mathsf{T}'_s)\mathbf{1}_{\{\mu(\mathsf{T}'_s)>\delta\}},0\le t\le r\right)\right) \tag{2.4.13}$$

in $\mathbb{T}\times\mathbb{R}\times D[0,r]$, where $(\mathsf{T}'_s,s\ge 0)$ is a Poisson point process with intensity $\mathbf{N}_{\mathrm{B}}$, independent of $(\mathscr{T},H(U))$.

Furthermore, since $\sum_{s\le r}\mu(\mathsf{T}'_s)$ is $\mathbf{N}^{(1)}$-a.s. finite, it is clear by the dominated convergence theorem that $\mathbf{N}^{(1)}$-a.s.

$$\lim_{\delta\to 0}\sup_{t\le r}\left|\sum_{s\le t}\mu(\mathsf{T}'_s)-\sum_{s\le t}\mu(\mathsf{T}'_s)\mathbf{1}_{\{\mu(\mathsf{T}'_s)>\delta\}}\right|=\lim_{\delta\to 0}\sum_{s\le r}\mu(\mathsf{T}'_s)\mathbf{1}_{\{\mu(\mathsf{T}'_s)\le\delta\}}=0.$$

Since uniform convergence on $[0,T]$ implies convergence for the Skorokhod $J1$ topology, we deduce that

$$\left(\mathscr{T},H(U),\left(\sum_{s\le t}\mu(\mathsf{T}'_s)\mathbf{1}_{\{\mu(\mathsf{T}'_s)>\delta\}},0\le t\le r\right)\right)\xrightarrow[\delta\to 0]{(d)}\left(\mathscr{T},H(U),(S_t,0\le t\le r)\right), \tag{2.4.14}$$

where $S_t=\sum_{s\le t}\mu(\mathsf{T}'_s)$ is a stable subordinator with Laplace exponent φ, independent of $(\mathscr{T},H(U))$.

Finally, we shall prove that for every $\eta>0$

$$\lim_{\delta\to 0}\limsup_{\varepsilon\to 0}\mathbf{N}^{(1)}\left[\sup_{0\le t\le T}\left|S_t^{\varepsilon}-S_t^{\varepsilon,\delta}\right|\ge\eta\right]=0. \tag{2.4.15}$$

Let $f\colon[0,\infty)\to[0,\infty)$ be Lipschitz-continuous such that $x\mathbf{1}_{[0,\delta]}(x)\le f(x)\le x\mathbf{1}_{[0,2\delta]}(x)$. We have

$$\begin{aligned}\sup_{0\le t\le r}\left|S_t^{\varepsilon}-S_t^{\varepsilon,\delta}\right|&=\sum_{h_i\le\varepsilon r\wedge\mathfrak{f}(\varepsilon)H(U)}\varepsilon^{-\gamma/(\gamma-1)}\sigma_i\mathbf{1}_{\{\varepsilon^{-\gamma/(\gamma-1)}\sigma_i\le\delta\}}\\&\le\sum_{h_i\le\varepsilon r\wedge\mathfrak{f}(\varepsilon)H(U)}f\left(\varepsilon^{-\gamma/(\gamma-1)}\sigma_i\right).\end{aligned}$$

It follows that

$$\begin{aligned}\limsup_{\varepsilon\to 0}\mathbf{N}^{(1)}\left[\sup_{0\le t\le r}\left|S_t^{\varepsilon}-S_t^{\varepsilon,\delta}\right|\ge\eta\right]&\le\limsup_{\varepsilon\to 0}\mathbf{N}^{(1)}\left[\sum_{h_i\le\varepsilon r\wedge\mathfrak{f}(\varepsilon)H(U)}f\left(\varepsilon^{-\gamma/(\gamma-1)}\sigma_i\right)\ge\eta\right]\\&\le\mathbf{N}^{(1)}\left[\sum_{s\le r}f\left(\mu(\mathsf{T}'_s)\right)\ge\eta\right]\\&\le\mathbf{N}^{(1)}\left[\sum_{s\le r}\mu(\mathsf{T}'_s)\mathbf{1}_{\{\mu(\mathsf{T}'_s)\le 2\delta\}}\ge\eta\right],\end{aligned} \tag{2.4.16}$$

where in the second inequality we used the Portmanteau theorem together with the following convergence in distribution

$$\sum_{h_i \le \varepsilon r \wedge \mathfrak{f}(\varepsilon) H(U)} f\left(\varepsilon^{-\gamma/(\gamma-1)}\sigma_i\right) \xrightarrow[\varepsilon\to 0]{(d)} \sum_{s\le r} f\left(\mu(\mathsf{T}'_s)\right),$$

which holds thanks to Theorem 2.4.2-(i) applied with $\Phi(h,a,T) = \mathbf{1}_{\{h\le r\}} f(a)$. But, by the dominated convergence theorem, we have that $\mathbf{N}^{(1)}$-a.s.

$$\lim_{\delta\to 0} \sum_{s\le r} \mu(\mathsf{T}'_s) \mathbf{1}_{\{\mu(\mathsf{T}'_s)\le 2\delta\}} = 0.$$

Together with (2.4.16), this implies (2.4.15).

Putting together (2.4.13)–(2.4.15), it follows from the second converging together theorem, see e.g. [38, Theorem 3.2], that

$$\left(\mathscr{T}, H(U), \left(S^{\varepsilon}_t, 0\le t\le r\right)\right) \xrightarrow[\varepsilon\to 0]{(d)} (\mathscr{T}, H(U), (S_t, 0\le t\le r))$$

in $\mathbb{T}\times\mathbb{R}\times D[0,r]$. This finishes the proof. □

Remark 2.4.5. Let us comment on the connection between Theorem 2.4.2 and the small time asymptotics of the fragmentation at height of the stable tree F^-, see [32, Section 4] for the Brownian case $\gamma = 2$ and [124] for the case $\gamma \in (1,2)$. We briefly recall its definition. Consider the normalized stable tree $\mathscr{T}$ and denote by $\left(\mathscr{T}_j,\ j\in J_t\right)$ the connected components of the set $\{x\in\mathscr{T} : H(x) > t\}$ obtained from $\mathscr{T}$ by removing vertices located at height $\le t$. Then $F^-(t) = (F^-_1(t), F^-_2(t), \dots)$ is defined as the decreasing sequence of masses $\left(\mu(\mathscr{T}_j),\ j\in J_t\right)$. In [80, Section 5.1], Haas obtains the following functional convergence in distribution as a consequence of a more general result

$$\varepsilon^{-\gamma/(\gamma-1)}(1-F^-_1(\varepsilon\cdot), (F^-_2(\varepsilon\cdot), F^-_3(\varepsilon\cdot),\dots)) \xrightarrow[\varepsilon\to 0]{(d)} (S, FI), \tag{2.4.17}$$

where the convergence holds with respect to the Skorokhod $J1$ topology. Here FI is a fragmentation process with immigration and S is a stable subordinator with index $1-1/\gamma$ representing the total mass of immigrants.

At least heuristically, this can be recovered from Theorem 2.4.2. Let $U\in\mathscr{T}$ be a leaf chosen uniformly at random. It is not difficult to see that for $0\le t\le H(U)$, with high probability as $\varepsilon\to 0$, the biggest fragment at time εt is the one containing U. Thus we get $1-F^-_1(\varepsilon t) = \sum_{h_i\le\varepsilon t}\sigma_i$ and

$$(F^-_2(\varepsilon t), F^-_3(\varepsilon t),\dots) = (\mu(\mathscr{T}_i^{\ge\varepsilon t-h_i}),\ h_i\le\varepsilon t)^{\downarrow}$$

is the decreasing rearrangement of the masses of $\mathscr{T}_i^{\ge\varepsilon t-h_i}$ for the subtrees grafted at height $h_i\le\varepsilon t$. Here we denote by $T^{\ge r} = T\setminus T^{<r} = \{x\in T : H(x)\ge r\}$ the set of vertices of T above

height r. To recover (2.4.17), we may prove the joint convergence of

$$\left(\sum_{h_i \le \varepsilon \cdot \wedge \varepsilon H(U)} \varepsilon^{-\gamma/(\gamma-1)} \sigma_i, \sum_{h_i \le \varepsilon H(U)} \delta_{\left(\mathbf{1}_{\{h_i \le \varepsilon t\}} \varepsilon^{-\gamma/(\gamma-1)} \mu(\mathcal{T}_i^{\ge \varepsilon t - h_i}),\, t \ge 0\right)} \right), \tag{2.4.18}$$

then argue that the convergence of the point measure in (2.4.18) implies that of the re-arranged atoms. Notice that we may obtain the convergence of the first coordinate in (2.4.18) using Theorem 2.4.2-(ii), similarly to how we proved Corollary 2.4.4 using Theorem 2.4.2-(i). For the convergence of the second coordinate, the idea is to consider $\Phi(h, a, T) = F\left((\mathbf{1}_{\{h \le t\}} a \mu(T^{\ge a^{-1+1/\gamma}(t-h)}),\, t \ge 0)\right)$, where $F\colon D[0,\infty) \to [0,\infty)$ is Lipschitz-continuous with compact support. However, Φ is not Lipschitz-continuous with respect to a so our result does not apply directly. Similarly, to get the convergence of the dust, notice that

$$\mu(\mathcal{T}^{<\varepsilon t}) = \sum_{h_i \le \varepsilon t} \mu(\mathcal{T}_i^{<\varepsilon t - h_i}).$$

Thus the idea is to apply Theorem 2.4.2-(ii) with $\Phi(h, a, T) = \mathbf{1}_{\{h \le t\}} a \mu(T^{<a^{-1+1/\gamma}(t-h)})$ which again does not satisfy the assumptions.

2.5 Asymptotic behavior of $\mathbf{Z}_{\alpha,\beta}$ in the case $\beta/\alpha^{1-1/\gamma} \to c \in [0,\infty)$

We start by showing that if $U \in \mathcal{T}$ is a leaf chosen uniformly at random, $Z_{\alpha,\beta}(U)$ defined in (2.1.1) converges in disrtibution after proper rescaling.

Proposition 2.5.1. *Assume that $\alpha \to \infty$, $\beta \ge 0$ and $\beta/\alpha^{1-1/\gamma} \to c \in [0,\infty)$. Let $\mathcal{T}$ be the normalized stable tree with branching mechanism $\psi(\lambda) = \lambda^\gamma$ where $\gamma \in (1,2]$. Conditionally on $\mathcal{T}$, let U be a $\mathcal{T}$-valued random variable with disribution μ under $\mathbf{N}^{(1)}$. Then we have the following convergence in distribution*

$$\left(\mathcal{T}, H(U), \alpha^{1-1/\gamma} \mathfrak{h}^{-\beta} Z_{\alpha,\beta}(U)\right) \xrightarrow[\alpha\to\infty]{(d)} \left(\mathcal{T}, H(U), \int_0^\infty \mathrm{e}^{-S_t - ct/\mathfrak{h}}\, \mathrm{d}t\right), \tag{2.5.1}$$

where $(S_t,\, t \ge 0)$ is a stable subordinator with Laplace exponent φ given by (2.3.10), independent of $(\mathcal{T}, H(U))$.

Proof. Set

$$\varepsilon = \varepsilon(\alpha) := \alpha^{(\delta-1)(1-1/\gamma)} \tag{2.5.2}$$

with $\delta \in (0, 1/3)$ so that $\varepsilon \to 0$ as $\alpha \to \infty$. Define

$$I_\alpha := \alpha^{1-1/\gamma} \int_0^{\varepsilon H(U)} \mathrm{e}^{-\alpha(1-\sigma_{r,U}) - \beta r/\mathfrak{h}}\, \mathrm{d}r. \tag{2.5.3}$$

Lemma 2.5.2. *We have the following convergence in $\mathbf{N}^{(1)}$-probability*

$$\lim_{\alpha\to\infty} \left(\alpha^{1-1/\gamma} \mathfrak{h}^{-\beta} Z_{\alpha,\beta}(U) - I_\alpha\right) = 0.$$

The proof is postponed to Section 2.7.2. Using this together with Slutsky's theorem, it is clear that the proof of (2.5.1) reduces to showing the following convergence in distribution

$$(\mathscr{T}, H(U), I_\alpha) \xrightarrow[\alpha\to\infty]{(d)} \left(\mathscr{T}, H(U), \int_0^\infty \mathrm{e}^{-S_t - ct/\mathfrak{h}}\, \mathrm{d}t\right). \tag{2.5.4}$$

Making the change of variable $t = \alpha^{1-1/\gamma} r$, notice that

$$I_\alpha = \int_0^{\alpha^{1-1/\gamma}\varepsilon H(U)} \exp\left\{-\alpha\left(1-\sigma_{\alpha^{-1+1/\gamma}t,U}\right) - \beta\alpha^{-1+1/\gamma} t/\mathfrak{h}\right\} \mathrm{d}t, \tag{2.5.5}$$

Let $A > 0$. Notice that, applying Corollary 2.4.4, we get the following convergence in distribution

$$\left(\mathscr{T}, H(U), \left(\sum_{h_i \le \alpha^{-1+1/\gamma} t \wedge \varepsilon H(U)} \alpha\sigma_i, 0 \le t \le A\right)\right) \xrightarrow[\alpha\to\infty]{(d)} (\mathscr{T}, H(U), (S_t, 0 \le t \le A)), \tag{2.5.6}$$

where S is a subordinator with Laplace exponent φ, independent of $(\mathscr{T}, H(U))$. Moreover, on the event $\Omega_\alpha := \{\alpha^{-1+1/\gamma} A \le \varepsilon H(U)\}$, we have for every $t \in [0, A]$

$$\sum_{h_i \le \alpha^{-1+1/\gamma} t \wedge \varepsilon H(U)} \sigma_i = \sum_{h_i \le \alpha^{-1+1/\gamma} t} \sigma_i = 1 - \sigma_{\alpha^{-1+1/\gamma}t,U}. \tag{2.5.7}$$

Since $\alpha^{1-1/\gamma}\varepsilon \to \infty$, it is clear that $\lim_{\alpha\to\infty} \mathbf{N}^{(1)}[\Omega_\alpha] = 1$. Thus, it follows from (2.5.6) and (2.5.7) that

$$\left(\mathscr{T}, H(U), \left(\alpha\left(1-\sigma_{\alpha^{-1+1/\gamma}t,U}\right), 0 \le t \le A\right)\right) \xrightarrow[\alpha\to\infty]{(d)} (\mathscr{T}, H(U), (S_t, 0 \le t \le A)).$$

Now a simple application of the continuous mapping theorem gives

$$\begin{aligned}\left(\mathscr{T}, H(U), \int_0^A \exp\left\{-\alpha\left(1-\sigma_{\alpha^{-1+1/\gamma}t,U}\right) - \beta\alpha^{-1+1/\gamma} t/\mathfrak{h}\right\} \mathrm{d}t\right)& \\ \xrightarrow[\alpha\to\infty]{(d)} \left(\mathscr{T}, H(U), \int_0^A \mathrm{e}^{-S_t - ct/\mathfrak{h}}\, \mathrm{d}t\right).&\end{aligned} \tag{2.5.8}$$

On the other hand, applying (2.3.22) with $f(T) = \mathrm{e}^{-\alpha(1-\mu(T))}$ and $g(r) = \mathbf{1}_{\{r \ge \alpha^{-1+1/\gamma} A\}}$, we get

$$\begin{aligned}
\mathbf{N}^{(1)}&\left[\int_A^{\alpha^{1-1/\gamma}\varepsilon H(U)} \exp\left\{-\alpha\left(1-\sigma_{\alpha^{-1+1/\gamma}t,U}\right) - \beta\alpha^{-1+1/\gamma} t/\mathfrak{h}\right\} \mathrm{d}t\right] \\
&\le \alpha^{1-1/\gamma} \mathbf{N}^{(1)}\left[\int_{\alpha^{-1+1/\gamma}A}^{H(U)} \exp\left\{-\alpha\left(1-\sigma_{r,U}\right)\right\} \mathrm{d}r\right] \\
&= \frac{\alpha^{1-1/\gamma}}{\gamma\Gamma(1-1/\gamma)} \int_0^1 x^{-1/\gamma}(1-x)^{-1/\gamma} \mathrm{e}^{-\alpha x} \mathbf{N}^{(1)}\left[(\alpha x)^{1-1/\gamma} H(U) \ge A\right] \mathrm{d}x \\
&= \frac{1}{\gamma\Gamma(1-1/\gamma)} \int_0^\alpha y^{-1/\gamma}\left(1-\frac{y}{\alpha}\right)^{-1/\gamma} \mathrm{e}^{-y} \mathbf{N}^{(1)}\left[y^{1-1/\gamma} H(U) \ge A\right] \mathrm{d}y.
\end{aligned}$$

By the dominated convergence theorem, we have

$$\lim_{\alpha\to\infty}\int_0^{\alpha/2} y^{-1/\gamma}\left(1-\frac{y}{\alpha}\right)^{-1/\gamma} \mathrm{e}^{-y}\mathbf{N}^{(1)}\left[y^{1-1/\gamma}H(U)\geq A\right]\mathrm{d}y \\ = \int_0^{\infty} y^{-1/\gamma}\mathrm{e}^{-y}\mathbf{N}^{(1)}\left[y^{1-1/\gamma}H(U)\geq A\right]\mathrm{d}y.$$

Moreover, we have

$$\int_{\alpha/2}^{\alpha} y^{-1/\gamma}\left(1-\frac{y}{\alpha}\right)^{-1/\gamma} \mathrm{e}^{-y}\mathbf{N}^{(1)}\left[y^{1-1/\gamma}H(U)\geq A\right]\mathrm{d}y \\ \leq \mathrm{e}^{-\alpha/2}\int_{\alpha/2}^{\alpha} y^{-1/\gamma}\left(1-\frac{y}{\alpha}\right)^{-1/\gamma}\mathrm{d}y \\ = \alpha^{1-1/\gamma}\mathrm{e}^{-\alpha/2}\int_{1/2}^{1} z^{-1/\gamma}(1-z)^{-1/\gamma}\,\mathrm{d}z,$$

where the last term converges to 0 as $\alpha\to\infty$. We deduce that

$$\limsup_{\alpha\to\infty}\mathbf{N}^{(1)}\left[\int_A^{\alpha^{1-1/\gamma}\varepsilon H(U)}\exp\left\{-\alpha\left(1-\sigma_{\alpha^{-1+1/\gamma}t,U}\right)-\beta\alpha^{-1+1/\gamma}t/\mathfrak{h}\right\}\mathrm{d}t\right] \\ \leq \frac{1}{\gamma\Gamma(1-1/\gamma)}\int_0^{\infty} y^{-1/\gamma}\mathrm{e}^{-y}\mathbf{N}^{(1)}\left[y^{1-1/\gamma}H(U)\geq A\right]\mathrm{d}y,$$

and, thanks to the dominated convergence theorem,

$$\lim_{A\to\infty}\limsup_{\alpha\to\infty}\mathbf{N}^{(1)}\left[\int_A^{\alpha^{1-1/\gamma}\varepsilon H(U)}\exp\left\{-\alpha\left(1-\sigma_{\alpha^{-1+1/\gamma}t,U}\right)-\beta\alpha^{-1+1/\gamma}t/\mathfrak{h}\right\}\mathrm{d}t\right]=0. \tag{2.5.9}$$

Combining (2.5.8) and (2.5.9) and applying [38, Theorem 3.2], (2.5.4) readily follows. This finishes the proof. □

The next lemma, whose proof is postponed to Section 2.7.3, states that taking a leaf uniformly at random or taking the average over all leaves yields the same limiting behavior for $Z_{\alpha,\beta}(x)$. Recall from (2.1.1) the definition of $\mathbf{Z}_{\alpha,\beta}$.

Lemma 2.5.3. *Under the assumptions of Theorem 2.5.1, we have the convergence in $\mathbf{N}^{(1)}$-probability*

$$\lim_{\alpha\to\infty}\alpha^{1-1/\gamma}\mathfrak{h}^{-\beta}\left(Z_{\alpha,\beta}(U)-\mathbf{Z}_{\alpha,\beta}\right)=0. \tag{2.5.10}$$

Combining Proposition 2.5.1 and Lemma 2.5.3, we get the following result using Slutsky's theorem.

Theorem 2.5.4. *Assume that $\alpha\to\infty$, $\beta\geq 0$ and $\beta/\alpha^{1-1/\gamma}\to c\in[0,\infty)$. Let $\mathcal{T}$ be the stable tree with branching mechanism $\psi(\lambda)=\lambda^\gamma$ where $\gamma\in(1,2]$. Conditionally on $\mathcal{T}$, let U be a $\mathcal{T}$-valued random variable with distribution μ under $\mathbf{N}^{(1)}$. Then we have the following convergence in* *distribution*

$$\left(\mathcal{T}, H(U), \alpha^{1-1/\gamma}\mathfrak{h}^{-\beta} Z_{\alpha,\beta}(U), \alpha^{1-1/\gamma}\mathfrak{h}^{-\beta}\mathbf{Z}_{\alpha,\beta}\right) \xrightarrow[\alpha\to\infty]{(d)} \left(\mathcal{T}, H(U), \int_0^\infty \mathrm{e}^{-S_t - ct/\mathfrak{h}}\,\mathrm{d}t, \int_0^\infty \mathrm{e}^{-S_t-ct/\mathfrak{h}}\,\mathrm{d}t\right), \quad (2.5.11)$$

where S is a stable subordinator with Laplace exponent φ given by (2.3.10)*, independent of* $(\mathcal{T}, H(U))$.

2.6 Asymptotic behavior of $\mathbf{Z}_{\alpha,\beta}$ in the case $\beta/\alpha^{1-1/\gamma} \to \infty$

We treat the case $\beta/\alpha^{1-1/\gamma} \to \infty$. Intuitively, this assumption guarantees that $\mathfrak{h}_{r,x}^{\beta}$ dominates $\sigma_{r,x}^{\alpha}$, thus we get a different asymptotic behavior and there is no longer a subordinator in the limit.

Theorem 2.6.1. *Assume that $\beta \to \infty$, $\alpha \geq 0$ and $\alpha^{1-1/\gamma}/\beta \to 0$. Let $\mathcal{T}$ be the stable tree with branching mechanism $\psi(\lambda) = \lambda^\gamma$ where $\gamma \in (1,2]$. Then we have the following convergence in $\mathbf{N}^{(1)}$-probability*

$$\lim_{\beta\to\infty} \beta\mathfrak{h}^{-\beta}\mathbf{Z}_{\alpha,\beta} = \mathfrak{h}. \quad (2.6.1)$$

Furthermore, if $\alpha^{1-1/\gamma}/\beta^\rho \to 0$ for some $\rho \in (0,1)$, then the convergence holds $\mathbf{N}^{(1)}$-almost surely.

Proof. We start by assuming that $\alpha \to \infty$ and $\alpha^{1-1/\gamma}/\beta \to 0$ (the case α bounded from above is covered by the second part of the theorem). Setting $\varepsilon = (\alpha^{1-1/\gamma}\beta)^{-1/2}$, it is straightforward to check that $\varepsilon \to 0$, $\beta\varepsilon \to \infty$ and $\alpha^{1-1/\gamma}\varepsilon \to 0$. Write

$$\beta\mathfrak{h}^{-\beta}\mathbf{Z}_{\alpha,\beta} = E_\beta + \sum_{i=1}^{4} F_\beta^i \quad (2.6.2)$$

where

$$F_\beta^1 = \beta\int_{\mathcal{T}} \mathbf{1}_{\{H(x)<2\varepsilon\}}\,\mu(\mathrm{d}x)\int_0^{H(x)} \sigma_{r,x}^{\alpha}\left(\frac{\mathfrak{h}_{r,x}}{\mathfrak{h}}\right)^{\beta}\mathrm{d}r,$$

$$F_\beta^2 = \beta\int_{\mathcal{T}} \mathbf{1}_{\{H(x)\geq 2\varepsilon\}}\,\mu(\mathrm{d}x)\int_\varepsilon^{H(x)} \sigma_{r,x}^{\alpha}\left(\frac{\mathfrak{h}_{r,x}}{\mathfrak{h}}\right)^{\beta}\mathrm{d}r,$$

$$F_\beta^3 = \beta\int_{\mathcal{T}} \mathbf{1}_{\{H(x)\geq 2\varepsilon\}}\,\mu(\mathrm{d}x)\int_0^{\varepsilon} \sigma_{r,x}^{\alpha}\left[\left(\frac{\mathfrak{h}_{r,x}}{\mathfrak{h}}\right)^{\beta} - \left(1-\frac{r}{\mathfrak{h}}\right)^{\beta}\right]\mathrm{d}r,$$

$$F_\beta^4 = \beta\int_{\mathcal{T}} \mathbf{1}_{\{H(x)\geq 2\varepsilon\}}\,\mu(\mathrm{d}x)\int_0^{\varepsilon} \sigma_{r,x}^{\alpha}\left[\left(1-\frac{r}{\mathfrak{h}}\right)^{\beta} - \mathrm{e}^{-\beta r/\mathfrak{h}}\right]\mathrm{d}r,$$

$$E_\beta = \beta\int_{\mathcal{T}} \mathbf{1}_{\{H(x)\geq 2\varepsilon\}}\,\mu(\mathrm{d}x)\int_0^{\varepsilon} \sigma_{r,x}^{\alpha}\mathrm{e}^{-\beta r/\mathfrak{h}}\,\mathrm{d}r.$$

We shall prove that $\lim_{\beta\to\infty} F_\beta^i = 0$ in $\mathbf{N}^{(1)}$-probability for every $i \in \{1,2,3,4\}$.

Let $p \in (1,2)$. Using that $\sigma_{r,x} \le 1$ and $\mathfrak{h}_{r,x} \le \mathfrak{h}$ and applying the Markov inequality, it is clear that

$$F^1_\beta \le 2\beta\varepsilon \int_{\mathscr{T}} \mathbf{1}_{\{H(x)<2\varepsilon\}}\, \mu(\mathrm{d}x) \le 2^{1+p}\beta\varepsilon^{1+p} \int_{\mathscr{T}} H(x)^{-p}\, \mu(\mathrm{d}x).$$

Since the last integral has a finite first moment by Lemma 2.3.9 and $\beta\varepsilon^{1+p} \to 0$, we deduce that $\mathbf{N}^{(1)}$-a.s. $\lim_{\beta\to\infty} F^1_\beta = 0$.

Next, using (2.2.5), we get

$$\begin{aligned} F^2_\beta &= \beta \int_{\mathscr{T}} \mathbf{1}_{\{H(x)\ge 2\varepsilon\}}\, \mu(\mathrm{d}x) \int_\varepsilon^{H(x)} \sigma^\alpha_{r,x} \left(\frac{\mathfrak{h}_{r,x}}{\mathfrak{h}}\right)^\beta \mathrm{d}r \\ &\le \beta \left(1 - \frac{\varepsilon}{\mathfrak{h}}\right)^\beta \int_{\mathscr{T}} \mu(\mathrm{d}x) \int_0^{H(x)} \sigma^\alpha_{r,x}\, \mathrm{d}r. \end{aligned} \tag{2.6.3}$$

By [7, Corollary 6.6], we have

$$\mathbf{N}^{(1)}\left[\int_{\mathscr{T}} \mu(\mathrm{d}x) \int_0^{H(x)} \sigma^\alpha_{r,x}\, \mathrm{d}r\right] = \frac{1}{|\Gamma(-1/\gamma)|} \mathrm{B}\big(\alpha + 1 - 1/\gamma, 1 - 1/\gamma\big),$$

where B is the beta function. Using that $\mathrm{B}(x, 1-1/\gamma) \sim \Gamma(1-1/\gamma)x^{-1+1/\gamma}$ as $x \to \infty$, we deduce that

$$\sup_{\alpha\ge 0} \mathbf{N}^{(1)}\left[\alpha^{1-1/\gamma} \int_{\mathscr{T}} \mu(\mathrm{d}x) \int_0^{H(x)} \sigma^\alpha_{r,x}\, \mathrm{d}r\right] < \infty. \tag{2.6.4}$$

On the other hand, let $\theta > 1$. Since the function $x \mapsto x^{1+\theta}\mathrm{e}^{-x}$ is bounded on $[0,\infty)$, it follows that

$$\frac{\beta}{\alpha^{1-1/\gamma}}\left(1 - \frac{\varepsilon}{\mathfrak{h}}\right)^\beta \le \frac{\beta}{\alpha^{1-1/\gamma}} \mathrm{e}^{-\beta\varepsilon/\mathfrak{h}} \le C \frac{\mathfrak{h}^{1+\theta}}{\beta^\theta \varepsilon^{1+\theta} \alpha^{1-1/\gamma}} \tag{2.6.5}$$

for some constant $C > 0$. Notice that $\beta^\theta\varepsilon^{1+\theta}\alpha^{1-1/\gamma} \to \infty$ since $\theta > 1$. Thus the right-hand side of (2.6.5) goes to 0 almost surely. Now putting together (2.6.3), (2.6.4) and (2.6.5), we deduce that $\lim_{\beta\to\infty} F^2_\beta = 0$ in $\mathbf{N}^{(1)}$-probability.

Let $x \in \mathscr{T}$. Recall from (2.2.5) and (2.2.6) that $\mathfrak{h}_{r,x} \le \mathfrak{h} - r$ for every $r \in [0, H(x)]$ and that the equality holds for $r \in [0, H(x \wedge x^*)]$. Therefore, we get

$$\begin{aligned} |F^3_\beta| &= \beta \int_{\mathscr{T}} \mathbf{1}_{\{H(x)\ge 2\varepsilon,\, H(x\wedge x^*)<\varepsilon\}}\, \mu(\mathrm{d}x) \int_{H(x\wedge x^*)}^{\varepsilon} \sigma^\alpha_{r,x} \left[\left(1 - \frac{r}{\mathfrak{h}}\right)^\beta - \left(\frac{\mathfrak{h}_{r,x}}{\mathfrak{h}}\right)^\beta\right] \mathrm{d}r \\ &\le \beta \int_{\mathscr{T}} \mathbf{1}_{\{H(x)\ge 2\varepsilon,\, H(x\wedge x^*)<\varepsilon\}}\, \mu(\mathrm{d}x) \int_{H(x\wedge x^*)}^{\varepsilon} \left(1 - \frac{r}{\mathfrak{h}}\right)^\beta \mathrm{d}r \\ &\le \beta \int_{\mathscr{T}} \mathbf{1}_{\{H(x)\ge 2\varepsilon,\, H(x\wedge x^*)<\varepsilon\}}\, \mu(\mathrm{d}x) \int_{H(x\wedge x^*)}^{\varepsilon} \mathrm{e}^{-\beta r/\mathfrak{h}}\, \mathrm{d}r \\ &\le \mathfrak{h} \int_{\mathscr{T}} \mathrm{e}^{-\beta H(x\wedge x^*)/\mathfrak{h}}\, \mu(\mathrm{d}x). \end{aligned}$$

Since $H(x \wedge x^*) > 0$ for μ-a.e. $x \in \mathscr{T}$, a simple application of the dominated convergence theorem gives that $\mathbf{N}^{(1)}$-a.s. $\lim_{\beta\to\infty} F^3_\beta = 0$.

Furthermore, using the inequality $|\mathrm{e}^b - \mathrm{e}^a| \le |b-a|\mathrm{e}^b$ for $a\le b$ together with the fact that $j\colon y\mapsto -(y+\log(1-y))/y^2$ is increasing on $[0,1)$, we get for $r\in[0,\varepsilon]$

$$\left|\mathrm{e}^{-\beta r/\mathfrak{h}} - \left(1-\frac{r}{\mathfrak{h}}\right)^{\beta}\right| \le \beta\left|\frac{r}{\mathfrak{h}}+\log\left(1-\frac{r}{\mathfrak{h}}\right)\right|\mathrm{e}^{-\beta r/\mathfrak{h}} \le \beta\left(\frac{r}{\mathfrak{h}}\right)^2\mathrm{e}^{-\beta r/\mathfrak{h}} j\left(\frac{\varepsilon}{\mathfrak{h}}\right).$$

Therefore, we deduce that

$$|F_\beta^4| \le j\left(\frac{\varepsilon}{\mathfrak{h}}\right)\int_{\mathcal{T}}\mathbf{1}_{\{H(x)\ge 2\varepsilon\}}\,\mu(\mathrm{d}x)\int_0^\varepsilon\left(\frac{\beta r}{\mathfrak{h}}\right)^2\mathrm{e}^{-\beta r/\mathfrak{h}}\,\mathrm{d}r \le Cj\left(\frac{\varepsilon}{\mathfrak{h}}\right)\varepsilon,$$

where we used that $y\mapsto y^2\mathrm{e}^{-y}$ is bounded on $[0,\infty)$ by some constant $C<\infty$ for the second inequality. Since $\lim_{y\to 0} j(y)=1/2$, we get $\mathbf{N}^{(1)}$-a.s. $\lim_{\beta\to\infty}F_\beta^4=0$. We deduce the following convergence in $\mathbf{N}^{(1)}$-probability

$$\lim_{\beta\to\infty}\sum_{i=1}^4 F_\beta^i = 0. \tag{2.6.6}$$

Notice that

$$E_\beta \le \beta\int_{\mathcal{T}}\mathbf{1}_{\{H(x)\ge 2\varepsilon\}}\,\mu(\mathrm{d}x)\int_0^\varepsilon \mathrm{e}^{-\beta r/\mathfrak{h}}\,\mathrm{d}r = \mathfrak{h}\left(1-\mathrm{e}^{-\beta\varepsilon/\mathfrak{h}}\right)\int_{\mathcal{T}}\mathbf{1}_{\{H(x)\ge 2\varepsilon\}}\,\mu(\mathrm{d}x)\le\mathfrak{h}. \tag{2.6.7}$$

On the other hand, using that $\sigma_{r,x}\ge\sigma_{\varepsilon,x}$ for every $x\in\mathcal{T}$ such that $H(x)\ge 2\varepsilon$ and every $r\in[0,\varepsilon]$, we get

$$E_\beta \ge \mathfrak{h}\left(1-\mathrm{e}^{-\beta\varepsilon/\mathfrak{h}}\right)\int_{\mathcal{T}}\mathbf{1}_{\{H(x)\ge 2\varepsilon\}}\sigma_{\varepsilon,x}^\alpha\,\mu(\mathrm{d}x). \tag{2.6.8}$$

We now shall prove the following convergence in $\mathbf{N}^{(1)}$-probability

$$\lim_{\beta\to\infty}\int_{\mathcal{T}}\mathbf{1}_{\{H(x)\ge 2\varepsilon\}}\sigma_{\varepsilon,x}^\alpha\,\mu(\mathrm{d}x)=1. \tag{2.6.9}$$

Using Lemma 2.3.5-(i) and Bismut's decomposition (2.3.12), we have

$$\begin{aligned}\mathbf{N}^{(1)}&\left[\int_{\mathcal{T}}\mathbf{1}_{\{H(x)\ge 2\varepsilon\}}\sigma_{\varepsilon,x}^\alpha\,\mu(\mathrm{d}x)\right]\\ &= \Gamma(1-1/\gamma)\,\mathbf{N}\left[\frac{1}{\sigma}\mathbf{1}_{\{\sigma>1\}}\int_{\mathcal{T}}\mathbf{1}_{\{\sigma^{-1+1/\gamma}H(x)\ge 2\varepsilon\}}\left(\frac{\sigma_{\sigma^{1-1/\gamma}\varepsilon,x}}{\sigma}\right)^\alpha\mu(\mathrm{d}x)\right]\\ &= \Gamma(1-1/\gamma)\int_0^\infty\mathrm{d}t\,\mathbb{E}\left[\frac{1}{S_t}\mathbf{1}_{\left\{S_t>1,\,t\ge 2\varepsilon S_t^{1-1/\gamma}\right\}}\left(1-\frac{S_{\varepsilon S_t^{1-1/\gamma}}}{S_t}\right)^\alpha\right].\end{aligned} \tag{2.6.10}$$

Recall that S is a stable subordinator with index $1-1/\gamma$. Thus the process T defined by

$$T_r := \frac{1}{\alpha}S_{\alpha^{1-1/\gamma}r},\quad \forall r\ge 0$$

is distributed as S. Applying this, we get that

$$\alpha S\left(\varepsilon S_t^{1-1/\gamma}\right) \overset{(d)}{=} \alpha T\left(\varepsilon T_t{}^{1-1/\gamma}\right) = S\left(\varepsilon S^{1-1/\gamma}_{\alpha^{1-1/\gamma}t}\right). \tag{2.6.11}$$

Now notice that

$$\varepsilon S^{1-1/\gamma}_{\alpha^{1-1/\gamma}t} = \varepsilon \alpha^{1-1/\gamma} T_t^{1-1/\gamma} \overset{(d)}{=} \varepsilon \alpha^{1-1/\gamma} S_t^{1-1/\gamma}.$$

Since $\varepsilon\alpha^{1-1/\gamma} \to 0$, this clearly implies that $\varepsilon S^{1-1/\gamma}_{\alpha^{1-1/\gamma}t} \to 0$ in probability. As S is a.s. continuous at 0, we deduce that $S\left(\varepsilon S^{1-1/\gamma}_{\alpha^{1-1/\gamma}t}\right) \to 0$ in probability. Thus, it follows from (2.6.11) that $\alpha S\left(\varepsilon S_t^{1-1/\gamma}\right) \to 0$ in probability for every $t > 0$ and

$$\alpha \log\left(1 - \frac{S_{\varepsilon S_t^{1-1/\gamma}}}{S_t}\right) \sim -\alpha \frac{S\left(\varepsilon S_t^{1-1/\gamma}\right)}{S_t} \xrightarrow{\mathbb{P}} 0.$$

In particular, this implies the following convergence in probability for every $t > 0$

$$\frac{1}{S_t}\mathbf{1}_{\left\{S_t>1,\, t\geq 2\varepsilon S_t^{1-1/\gamma}\right\}}\left(1 - \frac{S_{\varepsilon S_t^{1-1/\gamma}}}{S_t}\right)^{\alpha} \to \frac{1}{S_t}\mathbf{1}_{\{S_t>1\}}.$$

Since we have the inequality

$$\frac{1}{S_t}\mathbf{1}_{\left\{S_t>1,\, t\geq 2\varepsilon S_t^{1-1/\gamma}\right\}}\left(1 - \frac{S_{\varepsilon S_t^{1-1/\gamma}}}{S_t}\right)^{\alpha} \leq \frac{1}{S_t}\mathbf{1}_{\{S_t>1\}}$$

where the right-hand side is integrable with respect to $\mathbf{1}_{(0,\infty)}(t)\,\mathrm{d}t \otimes \mathbb{P}$ thanks to (2.3.13), the dominated convergence theorem yields

$$\int_0^\infty \mathrm{d}t\,\mathbb{E}\left[\frac{1}{S_t}\mathbf{1}_{\left\{S_t>1,\, t\geq 2\varepsilon S_t^{1-1/\gamma}\right\}}\left(1 - \frac{S_{\varepsilon S_t^{1-1/\gamma}}}{S_t}\right)^{\alpha}\right] \to \int_0^\infty \mathrm{d}t\,\mathbb{E}\left[\frac{1}{S_t}\mathbf{1}_{\{S_t>1\}}\right] = \frac{1}{\Gamma(1-1/\gamma)}.$$

Together with (2.6.10) and the fact that

$$\int_{\mathcal{T}} \mathbf{1}_{\{H(x)\geq 2\varepsilon\}}\sigma^{\alpha}_{\varepsilon,x}\,\mu(\mathrm{d}x) \leq 1,$$

this proves (2.6.9).

Finally, since $\beta\varepsilon \to \infty$, it is clear that $\mathfrak{h}(1-\mathrm{e}^{-\beta\varepsilon/\mathfrak{h}}) \to \mathfrak{h}$ almost surely. In conjunction with (2.6.9), this gives the following convergence in $\mathbf{N}^{(1)}$-probability

$$\mathfrak{h}\left(1-\mathrm{e}^{-\beta\varepsilon/\mathfrak{h}}\right)\int_{\mathcal{T}} \mathbf{1}_{\{H(x)\geq 2\varepsilon\}}\sigma^{\alpha}_{\varepsilon,x}\,\mu(\mathrm{d}x) \to \mathfrak{h}.$$

Thus, using this together (2.6.7) and (2.6.8) yields $\lim_{\beta\to\infty} E_\beta = \mathfrak{h}$ in $\mathbf{N}^{(1)}$-probability. It follows from (2.6.2) and (2.6.6) that $\lim_{\beta\to\infty}\beta\mathfrak{h}^{-\beta}\mathbf{Z}_{\alpha,\beta} = \mathfrak{h}$ in $\mathbf{N}^{(1)}$-probability. This proves the first part of the theorem.

Next, we treat the case $\alpha^{1-1/\gamma}/\beta^{\rho}\to 0$ for some $\rho\in(0,1)$. The proof is similar and we only highlight the differences. Notice that there exists $p,q\in(0,1)$ and $\theta\in(0,\gamma/(\gamma-1))$ such that $(1+p)q>1$ and $q\theta>\rho\gamma/(\gamma-1)$. Taking $\varepsilon=\beta^{-q}$, it is straighforward to check that $\varepsilon\to 0$, $\beta\varepsilon\to\infty$, $\beta\varepsilon^{1+p}\to 0$ and $\alpha\varepsilon^{\theta}\to 0$. As in the first part, we have that $\mathbf{N}^{(1)}$-a.s. $\lim_{\beta\to\infty}F^1_\beta+F^3_\beta+F^4_\beta=0$.

Furthermore, using that $\sigma_{r,x}\le 1$, it follows from (2.6.3) that

$$F^2_\beta\le\beta\left(1-\frac{\varepsilon}{\mathfrak{h}}\right)^{\beta}\mathfrak{h}\le\beta\mathrm{e}^{-\beta\varepsilon/\mathfrak{h}}\mathfrak{h}=\beta\mathrm{e}^{-\beta^{1-q}/\mathfrak{h}}.$$

This proves that $\mathbf{N}^{(1)}$-a.s. $\lim_{\beta\to\infty}F^2_\beta=0$.

Now we shall prove that $\mathbf{N}^{(1)}$-a.s. $\mu(\mathrm{d}x)$-a.s.

$$\lim_{\beta\to\infty}\mathbf{1}_{\{H(x)\ge 2\varepsilon\}}\sigma^{\alpha}_{\varepsilon,x}=1. \tag{2.6.12}$$

Using the same computation as in (2.6.10), we have the following identity in distribution

$$\begin{aligned}&\left(\mathbf{1}_{\{H(x)\ge 2\varepsilon\}}\sigma^{\alpha}_{\varepsilon,x},\,\varepsilon>0\right)\quad\text{under }\mathbf{N}^{(1)}\\ &\overset{(d)}{=}\left(\mathbf{1}_{\left\{t\ge 2\varepsilon S_t^{1-1/\gamma}\right\}}\left(1-\frac{S\left(\varepsilon S_t^{1-1/\gamma}\right)}{S_t}\right)^{\alpha},\,\varepsilon>0\right)\quad\text{under }\int_0^\infty \mathrm{d}t\,\mathbb{E}\left[\frac{1}{S_t}\mathbf{1}_{\{S_t>1\}}\bullet\right].\end{aligned}\tag{2.6.13}$$

Since $\theta<\gamma/(\gamma-1)$, [29, Chapter III, Theorem 9] guarantees that $\mathbb{P}$-a.s. $\limsup_{r\to 0}r^{-\theta}S_r=0$. By composition, it follows that $\mathbb{P}$-a.s. for every $t>0$, $\lim_{\varepsilon\to 0}\varepsilon^{-\theta}S\left(\varepsilon S_t^{1-1/\gamma}\right)=0$. Thus we deduce that

$$\alpha\log\left(1-\frac{S\left(\varepsilon S_t^{1-1/\gamma}\right)}{S_t}\right)\sim-\alpha\frac{S\left(\varepsilon S_t^{1-1/\gamma}\right)}{S_t}=-\alpha\varepsilon^{\theta}\frac{\varepsilon^{-\theta}S\left(\varepsilon S_t^{1-1/\gamma}\right)}{S_t}\to 0$$

since $\alpha\varepsilon^{\theta}\to 0$. This proves that the process in the right-hand side of (2.6.13) goes to 1 $\mathbb{P}$-a.s. as $\varepsilon\to 0$, thus (2.6.12) follows.

Thanks to (2.6.12), since $\sigma_{\varepsilon,x}\le 1$, a simple application of the dominated convergence theorem gives that $\mathbf{N}^{(1)}$-a.s.

$$\lim_{\beta\to\infty}\int_{\mathcal{T}}\mathbf{1}_{\{H(x)\ge 2\varepsilon\}}\sigma^{\alpha}_{\varepsilon,x}\,\mu(\mathrm{d}x)=1.$$

This, together with the estimates (2.6.7) and (2.6.8) yields the $\mathbf{N}^{(1)}$-a.s. convergence $\lim_{\beta\to\infty}E_\beta=\mathfrak{h}$ which concludes the proof of the second part of the theorem. □

2.7 Technical lemmas

2.7.1 Proof of Lemma 2.4.3

Recall that $\mathfrak{g}(\varepsilon) = 1 - \mathfrak{f}(\varepsilon)$. Using the expression of $F(\varepsilon)$ from (2.4.7), we write

$$\begin{aligned}
&\Gamma(1-1/\gamma)^{-1}F(\varepsilon) - \int_0^\infty \mathrm{d}t\,\mathbb{E}\left[\frac{1}{\tau_{\mathfrak{g}(\varepsilon)t}}\mathbf{1}_{\{\tau_{\mathfrak{g}(\varepsilon)t}>1\}} f\circ\mathsf{norm}_\gamma\left(\mathsf{T}^{\downarrow}_{\mathfrak{g}(\varepsilon)t}\right) g\left(\tau_{\mathfrak{g}(\varepsilon)t}^{-1+1/\gamma}t\right)\right.\\
&\left.\times\exp\left\{-\sum_{s\le\mathfrak{f}(\varepsilon)t}\Phi\left(\varepsilon^{-1}s/t,\varepsilon^{-\gamma/(\gamma-1)}\tau_{\mathfrak{g}(\varepsilon)t}^{-1}\mu(\mathsf{T}_s),\mathsf{norm}_\gamma(\mathsf{T}_s)\right)\right\}\right] = \sum_{i=1}^4\int_0^\infty \mathrm{d}t\,\mathbb{E}\left[N^i_\varepsilon(t)\right], \quad (2.7.1)
\end{aligned}$$

where

$$\begin{aligned}
N^1_\varepsilon(t) &= \frac{1}{\tau_t}\mathbf{1}_{\{\tau_t>1\}}\left\{f\circ\mathsf{norm}_\gamma\left(\mathsf{T}^{\downarrow}_t\right) - f\circ\mathsf{norm}_\gamma\left(\mathsf{T}^{\downarrow}_{\mathfrak{g}(\varepsilon)t}\right)\right\} g\left(\tau_t^{-1+1/\gamma}t\right)\\
&\qquad\times\exp\left\{-\sum_{s\le\mathfrak{f}(\varepsilon)t}\Phi\left(\varepsilon^{-1}s/t,\varepsilon^{-\gamma/(\gamma-1)}\tau_t^{-1}\mu(\mathsf{T}_s),\mathsf{norm}_\gamma(\mathsf{T}_s)\right)\right\},\\
N^2_\varepsilon(t) &= \frac{1}{\tau_t}\mathbf{1}_{\{\tau_t>1\}} f\circ\mathsf{norm}_\gamma\left(\mathsf{T}^{\downarrow}_{\mathfrak{g}(\varepsilon)t}\right)\left\{g\left(\tau_t^{-1+1/\gamma}t\right) - g\left(\tau_{\mathfrak{g}(\varepsilon)t}^{-1+1/\gamma}t\right)\right\}\\
&\qquad\times\exp\left\{-\sum_{s\le\mathfrak{f}(\varepsilon)t}\Phi\left(\varepsilon^{-1}s/t,\varepsilon^{-\gamma/(\gamma-1)}\tau_t^{-1}\mu(\mathsf{T}_s),\mathsf{norm}_\gamma(\mathsf{T}_s)\right)\right\},\\
N^3_\varepsilon(t) &= \frac{1}{\tau_t}\mathbf{1}_{\{\tau_t>1\}} f\circ\mathsf{norm}_\gamma\left(\mathsf{T}^{\downarrow}_{\mathfrak{g}(\varepsilon)t}\right) g\left(\tau_{\mathfrak{g}(\varepsilon)t}^{-1+1/\gamma}t\right)\\
&\qquad\times\left[\exp\left\{-\sum_{s\le\mathfrak{f}(\varepsilon)t}\Phi\left(\varepsilon^{-1}s/t,\varepsilon^{-\gamma/(\gamma-1)}\tau_t^{-1}\mu(\mathsf{T}_s),\mathsf{norm}_\gamma(\mathsf{T}_s)\right)\right\}\right.\\
&\qquad\quad\left.-\exp\left\{-\sum_{s\le\mathfrak{f}(\varepsilon)t}\Phi\left(\varepsilon^{-1}s/t,\varepsilon^{-\gamma/(\gamma-1)}\tau_{\mathfrak{g}(\varepsilon)t}^{-1}\mu(\mathsf{T}_s),\mathsf{norm}_\gamma(\mathsf{T}_s)\right)\right\}\right],\\
N^4_\varepsilon(t) &= \left\{\frac{1}{\tau_t}\mathbf{1}_{\{\tau_t>1\}} - \frac{1}{\tau_{\mathfrak{g}(\varepsilon)t}}\mathbf{1}_{\{\tau_{\mathfrak{g}(\varepsilon)t}>1\}}\right\} f\circ\mathsf{norm}_\gamma\left(\mathsf{T}^{\downarrow}_{\mathfrak{g}(\varepsilon)t}\right) g\left(\tau_{\mathfrak{g}(\varepsilon)t}^{-1+1/\gamma}t\right)\\
&\qquad\times\exp\left\{-\sum_{s\le\mathfrak{f}(\varepsilon)t}\Phi\left(\varepsilon^{-1}s/t,\varepsilon^{-\gamma/(\gamma-1)}\tau_{\mathfrak{g}(\varepsilon)t}^{-1}\mu(\mathsf{T}_s),\mathsf{norm}_\gamma(\mathsf{T}_s)\right)\right\}.
\end{aligned}$$

Recall from (2.1.3) the definition of norm_γ and notice that since the total mass of $\mathsf{T}^{\downarrow}_t$ is τ_t, we have $\mathsf{norm}_\gamma(\mathsf{T}^{\downarrow}_t) = R_\gamma(\mathsf{T}^{\downarrow}_t, \tau_t^{-1+1/\gamma})$. It follows that

$$\begin{aligned}
\left|N^1_\varepsilon(t)\right| &\le \|f\|_{\mathrm{L}}\|g\|_\infty\frac{1}{\tau_t}\mathbf{1}_{\{\tau_t>1\}} d_{\mathrm{GHP}}\left(\mathsf{norm}_\gamma\left(\mathsf{T}^{\downarrow}_t\right),\mathsf{norm}_\gamma\left(\mathsf{T}^{\downarrow}_{\mathfrak{g}(\varepsilon)t}\right)\right)\\
&\le \|f\|_{\mathrm{L}}\|g\|_\infty\mathbf{1}_{\{\tau_t>1\}}\left[d_{\mathrm{GHP}}\left(R_\gamma\left(\mathsf{T}^{\downarrow}_t,\tau_t^{-1+1/\gamma}\right),R_\gamma\left(\mathsf{T}^{\downarrow}_{\mathfrak{g}(\varepsilon)t},\tau_t^{-1+1/\gamma}\right)\right)\right.\\
&\qquad\left.+\,d_{\mathrm{GHP}}\left(R_\gamma\left(\mathsf{T}^{\downarrow}_{\mathfrak{g}(\varepsilon)t},\tau_t^{-1+1/\gamma}\right),R_\gamma\left(\mathsf{T}^{\downarrow}_{\mathfrak{g}(\varepsilon)t},\tau_{\mathfrak{g}(\varepsilon)t}^{-1+1/\gamma}\right)\right)\right],
\end{aligned}$$

where $\|f\|_{\mathrm{L}}$ denotes the Lipschitz constant of f. Notice that, by construction, the tree $\mathsf{T}^{\downarrow}_t$ is obtained from $\mathsf{T}^{\downarrow}_{\mathfrak{g}(\varepsilon)t}$ by adding to the root a branch $[0,\mathfrak{f}(\varepsilon)t)$ onto which we graft T_s at height $0 \le s < \mathfrak{f}(\varepsilon)t$. It is clear that the added part has mass $\sum_{s<\mathfrak{f}(\varepsilon)t}\mu(\mathsf{T}_s) = S_{f(\varepsilon)t-}$ and height at most $\max_{s<\mathfrak{f}(\varepsilon)t}\mathfrak{h}(\mathsf{T}_s)+\mathfrak{f}(\varepsilon)t$. Thus, by definition (2.1.2) of the mapping R_γ, we deduce that

$$d_{\mathrm{GHP}}\left(R_\gamma\left(\mathsf{T}^{\downarrow}_t,\tau_t^{-1+1/\gamma}\right),R_\gamma\left(\mathsf{T}^{\downarrow}_{\mathfrak{g}(\varepsilon)t},\tau_t^{-1+1/\gamma}\right)\right) \le \tau_t^{-1}S_{\mathfrak{f}(\varepsilon)t-}+\tau_t^{-1+1/\gamma}\left(\max_{s<\mathfrak{f}(\varepsilon)t}\mathfrak{h}(\mathsf{T}_s)+\mathfrak{f}(\varepsilon)t\right). \quad (2.7.2)$$

Moreover, using Lemma 2.2.2 and again the definition of R_γ, we get

$$d_{\mathrm{GHP}}\left(R_\gamma\left(\mathsf{T}^{\downarrow}_{\mathfrak{g}(\varepsilon)t},\tau_t^{-1+1/\gamma}\right),R_\gamma\left(\mathsf{T}^{\downarrow}_{\mathfrak{g}(\varepsilon)t},\tau_{\mathfrak{g}(\varepsilon)t}^{-1+1/\gamma}\right)\right) \le 2\left(\tau_{\mathfrak{g}(\varepsilon)t}^{-1+1/\gamma}-\tau_t^{-1+1/\gamma}\right)\mathfrak{h}\left(\mathsf{T}^{\downarrow}_{\mathfrak{g}(\varepsilon)t}\right)+\left(\tau_{\mathfrak{g}(\varepsilon)t}^{-1}-\tau_t^{-1}\right)\mu\left(\mathsf{T}^{\downarrow}_{\mathfrak{g}(\varepsilon)t}\right). \quad (2.7.3)$$

From (2.7.2) and (2.7.3), we deduce that

$$\left|N^1_\varepsilon(t)\right| \le \|f\|_{\mathrm{L}}\|g\|_\infty\Big[S_{\mathfrak{f}(\varepsilon)t-}+\max_{s<\mathfrak{f}(\varepsilon)t}\mathfrak{h}(\mathsf{T}_s)+\mathfrak{f}(\varepsilon)t +2\left(\tau_{\mathfrak{g}(\varepsilon)t}^{-1+1/\gamma}-\tau_t^{-1+1/\gamma}\right)\mathfrak{h}\left(\mathsf{T}^{\downarrow}_{\mathfrak{g}(\varepsilon)t}\right)+\left(\tau_{\mathfrak{g}(\varepsilon)t}^{-1}-\tau_t^{-1}\right)\mu\left(\mathsf{T}^{\downarrow}_{\mathfrak{g}(\varepsilon)t}\right)\Big].$$

Therefore it follows that for every $t>0$ $\mathbb{P}$-a.s.

$$\lim_{\varepsilon\to0}N^1_\varepsilon(t)=0. \quad (2.7.4)$$

Furthermore, it is clear that

$$\left|N^2_\varepsilon(t)\right| \le \|f\|_\infty\|g\|_{\mathrm{L}}\,t\left|\tau_t^{-1+1/\gamma}-\tau_{\mathfrak{g}(\varepsilon)t}^{-1+1/\gamma}\right|.$$

Thus, we have for every $t>0$ $\mathbb{P}$-a.s.

$$\lim_{\varepsilon\to0}N^2_\varepsilon(t)=0. \quad (2.7.5)$$

Since

$$\left|N^1_\varepsilon(t)+N^2_\varepsilon(t)\right| \le 4\|f\|_\infty\|g\|_\infty\frac{1}{\tau_t}\mathbf{1}_{\{\tau_t>1\}}$$

where the right-hand side is integrable with respect to $\mathbf{1}_{(0,\infty)}(t)\,\mathrm{d}t\otimes\mathbb{P}$ thanks to (2.3.13), it follows from (2.7.4) and (2.7.5) that

$$\lim_{\varepsilon\to0}\int_0^\infty \mathrm{d}t\,\mathbb{E}\left[N^1_\varepsilon(t)+N^2_\varepsilon(t)\right]=0. \quad (2.7.6)$$

Using the inequality $|e^b - e^a| \le 1 \wedge |b-a|$ for $a \le b \le 0$, we have

$$\begin{aligned}
\left|N_\varepsilon^3(t)\right| &\le \|f\|_\infty \|g\|_\infty \frac{1}{\tau_t}\mathbf{1}_{\{\tau_t>1\}}\Bigg(1\wedge \sum_{s\le \mathfrak{f}(\varepsilon)t}\Big|\Phi\big(\varepsilon^{-1}s/t, \varepsilon^{-\gamma/(\gamma-1)}\tau_t^{-1}\mu(\mathsf{T}_s), \mathsf{norm}_\gamma(\mathsf{T}_s)\big)\\
&\qquad\qquad -\Phi\Big(\varepsilon^{-1}s/t, \varepsilon^{-\gamma/(\gamma-1)}\tau_{\mathfrak{g}(\varepsilon)t}^{-1}\mu(\mathsf{T}_s), \mathsf{norm}_\gamma(\mathsf{T}_s)\Big)\Big|\Bigg)\\
&\le \|f\|_\infty \|g\|_\infty \frac{1}{\tau_t}\mathbf{1}_{\{\tau_t>1\}}\left(1\wedge C\varepsilon^{-\gamma/(\gamma-1)}\left|\tau_t^{-1}-\tau_{\mathfrak{g}(\varepsilon)t}^{-1}\right|\sum_{s\le\mathfrak{f}(\varepsilon)t}\mu(\mathsf{T}_s)\right)\\
&= \|f\|_\infty \|g\|_\infty \frac{1}{\tau_t}\mathbf{1}_{\{\tau_t>1\}}\left(1\wedge C\varepsilon^{-\gamma/(\gamma-1)}\frac{\left(\tau_t-\tau_{\mathfrak{g}(\varepsilon)t}\right)^2}{\tau_t\tau_{\mathfrak{g}(\varepsilon)t}}\right).
\end{aligned} \tag{2.7.7}$$

Since τ is a stable subordinator with index $1-1/\gamma$, we get that

$$\varepsilon^{-\gamma/(\gamma-1)}\left(\tau_t-\tau_{\mathfrak{g}(\varepsilon)t}\right)^2 \overset{(d)}{=} \varepsilon^{-\gamma/(\gamma-1)}\tau_{\mathfrak{f}(\varepsilon)t}^2 \overset{(d)}{=} \left(\varepsilon^{-1}\mathfrak{f}(\varepsilon)^2\right)^{\gamma/(\gamma-1)}\tau_t^2 \xrightarrow[\varepsilon\to0]{(d)} 0$$

as $\varepsilon^{-1}\mathfrak{f}(\varepsilon)^2 \to 0$. We deduce the following convergence in $\mathbb{P}$-probability

$$\lim_{\varepsilon\to0}\frac{1}{\tau_t}\mathbf{1}_{\{\tau_t>1\}}\left(1\wedge C\varepsilon^{-\gamma/(\gamma-1)}\frac{\left(\tau_t-\tau_{\mathfrak{g}(\varepsilon)t}\right)^2}{\tau_t\tau_{\mathfrak{g}(\varepsilon)t}}\right)=0.$$

Thanks to (2.3.13), it follows from the dominated convergence theorem that

$$\lim_{\varepsilon\to0}\int_0^\infty \mathrm{d}t\,\mathbb{E}\left[\frac{1}{\tau_t}\mathbf{1}_{\{\tau_t>1\}}\left(1\wedge C\varepsilon^{-\gamma/(\gamma-1)}\frac{\left(\tau_t-\tau_{\mathfrak{g}(\varepsilon)t}\right)^2}{\tau_t\tau_{\mathfrak{g}(\varepsilon)t}}\right)\right]=0.$$

Together with (2.7.7), this gives

$$\lim_{\varepsilon\to0}\int_0^\infty \mathrm{d}t\,\mathbb{E}\left[N_\varepsilon^3(t)\right]=0. \tag{2.7.8}$$

Finally, notice that

$$\left|\int_0^\infty \mathrm{d}t\,\mathbb{E}\left[N_\varepsilon^4(t)\right]\right| \le \|f\|_\infty\|g\|_\infty\int_0^\infty \mathrm{d}t\,\mathbb{E}\left[\frac{1}{\tau_t}\mathbf{1}_{\{\tau_{\mathfrak{g}(\varepsilon)t}\le1<\tau_t\}}+\frac{\tau_t-\tau_{\mathfrak{g}(\varepsilon)t}}{\tau_t\tau_{\mathfrak{g}(\varepsilon)t}}\mathbf{1}_{\{\tau_{\mathfrak{g}(\varepsilon)t}>1\}}\right]. \tag{2.7.9}$$

Thanks to (2.3.13) and the dominated convergence theorem, it is clear that

$$\lim_{\varepsilon\to0}\int_0^\infty \mathrm{d}t\,\mathbb{E}\left[\frac{1}{\tau_t}\mathbf{1}_{\{\tau_{\mathfrak{g}(\varepsilon)t}\le1<\tau_t\}}\right]=0 \tag{2.7.10}$$

as the process τ is a.s. continuous at t. On the other hand, using the inequality

$$\frac{\tau_t-\tau_{\mathfrak{g}(\varepsilon)t}}{\tau_t\tau_{\mathfrak{g}(\varepsilon)t}}\mathbf{1}_{\{\tau_{\mathfrak{g}(\varepsilon)t}>1\}} \le \left(\frac{\tau_t-\tau_{\mathfrak{g}(\varepsilon)t}}{\tau_t}\right)^{1-q}\frac{\left(\tau_t-\tau_{\mathfrak{g}(\varepsilon)t}\right)^q}{\tau_{\mathfrak{g}(\varepsilon)t}^{1+q}}\mathbf{1}_{\{\tau_{\mathfrak{g}(\varepsilon)t}>1\}}$$

$$\leq \frac{\left(\tau_t - \tau_{\mathfrak{g}(\varepsilon)t}\right)^q}{\tau_{\mathfrak{g}(\varepsilon)t}^{1+q}} \mathbf{1}_{\{\tau_{\mathfrak{g}(\varepsilon)t}>1\}}$$

where $q \in (0, 1-1/\gamma)$, we get that

$$\begin{aligned}
\int_0^\infty \mathrm{d}t \mathbb{E}\left[\frac{\tau_t - \tau_{\mathfrak{g}(\varepsilon)t}}{\tau_t \tau_{\mathfrak{g}(\varepsilon)t}} \mathbf{1}_{\{\tau_{\mathfrak{g}(\varepsilon)t}>1\}}\right] &\leq \int_0^\infty \mathrm{d}t \mathbb{E}\left[\frac{\left(\tau_t - \tau_{\mathfrak{g}(\varepsilon)t}\right)^q}{\tau_{\mathfrak{g}(\varepsilon)t}^{1+q}} \mathbf{1}_{\{\tau_{\mathfrak{g}(\varepsilon)t}>1\}}\right] \\
&= \int_0^\infty \mathrm{d}t \mathbb{E}\left[\tau_{\mathfrak{f}(\varepsilon)t}^q\right] \mathbb{E}\left[\frac{1}{\tau_{\mathfrak{g}(\varepsilon)t}^{1+q}} \mathbf{1}_{\{\tau_{\mathfrak{g}(\varepsilon)t}>1\}}\right] \\
&= \frac{\mathfrak{f}(\varepsilon)^{q\gamma/(\gamma-1)}}{\mathfrak{g}(\varepsilon)^{1+q\gamma/(\gamma-1)}} \mathbb{E}\left[\tau_1^q\right] \int_0^\infty \mathrm{d}r\, r^{q\gamma/(\gamma-1)} \mathbb{E}\left[\frac{1}{\tau_r^{1+q}} \mathbf{1}_{\{\tau_r>1\}}\right] \\
&= \frac{\mathfrak{f}(\varepsilon)^{q\gamma/(\gamma-1)}}{\mathfrak{g}(\varepsilon)^{1+q\gamma/(\gamma-1)}} \mathbb{E}\left[\tau_1^q\right] \mathbb{E}\left[\frac{1}{\tau_1^{1+q}} \int_{\tau_1^{-1+1/\gamma}}^\infty \mathrm{d}r\, r^{-\gamma/(\gamma-1)}\right] \\
&= \frac{\mathfrak{f}(\varepsilon)^{q\gamma/(\gamma-1)}}{\mathfrak{g}(\varepsilon)^{1+q\gamma/(\gamma-1)}} \mathbb{E}\left[\tau_1^q\right] \mathbb{E}\left[\tau_1^{-1-q+1/\gamma}\right], \qquad (2.7.11)
\end{aligned}$$

where we used that $\tau_t - \tau_{\mathfrak{g}(\varepsilon)t}$ is independent of $\tau_{\mathfrak{g}(\varepsilon)t}$ and is distributed as $\tau_{\mathfrak{f}(\varepsilon)t}$ for the first equality and that $\tau_t \overset{(d)}{=} t^{\gamma/(\gamma-1)}\tau_1$ for the second. Thanks to (2.3.11), we have $\mathbb{E}\left[\tau_1^q\right] < \infty$ and $\mathbb{E}\left[\tau_1^{-1+1/\gamma-q}\right] < \infty$. Thus, it follows from (2.7.11) that

$$\lim_{\varepsilon\to 0} \int_0^\infty \mathrm{d}t \mathbb{E}\left[\frac{\tau_t - \tau_{\mathfrak{g}(\varepsilon)t}}{\tau_t \tau_{\mathfrak{g}(\varepsilon)t}} \mathbf{1}_{\{\tau_{\mathfrak{g}(\varepsilon)t}>1\}}\right] = 0. \qquad (2.7.12)$$

Combining (2.7.9), (2.7.10) and (2.7.12), we deduce that

$$\lim_{\varepsilon\to 0} \int_0^\infty \mathrm{d}t \mathbb{E}\left[\left|N_\varepsilon^4(t)\right|\right] = 0. \qquad (2.7.13)$$

It follows from (2.7.1), (2.7.6), (2.7.8) and (2.7.13) that

$$\begin{aligned}
\lim_{\varepsilon\to 0} \Gamma(1-1/\gamma)^{-1} F(\varepsilon) - \int_0^\infty \mathrm{d}t \mathbb{E}\Bigg[\frac{1}{\tau_{\mathfrak{g}(\varepsilon)t}} \mathbf{1}_{\{\tau_{\mathfrak{g}(\varepsilon)t}>1\}} f \circ R\left(\mathsf{T}^{\downarrow}_{\mathfrak{g}(\varepsilon)t}, \tau_{\mathfrak{g}(\varepsilon)t}^{-1}\right) g\left(\tau_{\mathfrak{g}(\varepsilon)t}^{-1+1/\gamma} t\right) \\
\times \exp\left\{-\sum_{s\leq \mathfrak{f}(\varepsilon)t} \Phi\left(\varepsilon^{-1}s/t, \varepsilon^{-\gamma/(\gamma-1)}\tau_{\mathfrak{g}(\varepsilon)t}^{-1}\mu(\mathsf{T}_s), R\left(\mathsf{T}_s, \mu(\mathsf{T}_s)^{-1}\right)\right)\right\}\Bigg] = 0.
\end{aligned}$$

2.7.2 Proof of Lemma 2.5.2

Recall from (2.5.3) the definition of I_α. Write $\alpha^{1-1/\gamma}\mathfrak{h}^{-\beta} Z_{\alpha,\beta}(U) - I_\alpha = \sum_{i=1}^4 J_\alpha^i$ where

$$J_\alpha^1 = \alpha^{1-1/\gamma}\mathfrak{h}^{-\beta} \int_{\varepsilon H(U)}^{H(U)} \sigma_{r,U}^\alpha \mathfrak{h}_{r,U}^\beta \,\mathrm{d}r,$$

$$J_\alpha^2 = \alpha^{1-1/\gamma} \int_0^{\varepsilon H(U)} \sigma_{r,U}^\alpha \left\{ \left(\frac{\mathfrak{h}_{r,U}}{\mathfrak{h}} \right)^\beta - \left(1 - \frac{r}{\mathfrak{h}} \right)^\beta \right\} \mathrm{d}r,$$
$$J_\alpha^3 = \alpha^{1-1/\gamma} \int_0^{\varepsilon H(U)} \sigma_{r,U}^\alpha \left\{ \left(1 - \frac{r}{\mathfrak{h}} \right)^\beta - \mathrm{e}^{-\beta r/\mathfrak{h}} \right\} \mathrm{d}r,$$
$$J_\alpha^4 = \alpha^{1-1/\gamma} \int_0^{\varepsilon H(U)} \left\{ \sigma_{r,U}^\alpha - \mathrm{e}^{-\alpha(1-\sigma_{r,U})} \right\} \mathrm{e}^{-\beta r/\mathfrak{h}} \,\mathrm{d}r.$$

We shall prove that for every $1 \le i \le 4$, $\lim_{\alpha\to\infty} J_\alpha^i = 0$ in $\mathbf{N}^{(1)}$-probability.

We start by showing that $\mathbf{N}^{(1)}$-a.s. $\mu(\mathrm{d}x)$-a.s.

$$\lim_{\alpha\to\infty} \alpha^{1-1/\gamma} \int_{\varepsilon H(x)}^{H(x)} \sigma_{r,x}^\alpha \,\mathrm{d}r = 0. \tag{2.7.14}$$

Recall from (2.3.9) the definition of S. Using Lemma 2.3.5-(i) and Bismut's decomposition (2.3.12), we have

$$\begin{aligned}
&\Gamma(1-1/\gamma)^{-1} \mathbf{N}^{(1)} \left[\mu\left(x \in \mathscr{T} :\ \limsup_{\alpha\to\infty} \alpha^{1-1/\gamma} \int_{\varepsilon H(x)}^{H(x)} \sigma_{r,x}^\alpha \,\mathrm{d}r > 0 \right) \right] \\
&\quad = \mathbf{N}\left[\frac{1}{\sigma} \mathbf{1}_{\{\sigma>1\}} \mu\left(x \in \mathscr{T} :\ \limsup_{\alpha\to\infty} \left(\frac{\alpha}{\sigma}\right)^{1-1/\gamma} \int_{\varepsilon H(x)}^{H(x)} \left(\frac{\sigma_{r,x}}{\sigma}\right)^\alpha \mathrm{d}r > 0 \right) \right] \\
&\quad = \int_0^\infty \mathrm{d}t\, \mathbb{E}\left[\frac{1}{\tau_t} \mathbf{1}_{\{\tau_t>1\}} ; \limsup_{\alpha\to\infty} \left(\frac{\alpha}{\tau_t}\right)^{1-1/\gamma} \int_{\varepsilon t}^t \left(1 - \frac{S_r}{\tau_t}\right)^\alpha \mathrm{d}r > 0 \right].
\end{aligned} \tag{2.7.15}$$

Let $t > 0$. It is clear that

$$\int_{\varepsilon t}^t \left(1 - \frac{S_r}{\tau_t}\right)^\alpha \mathrm{d}r \le \int_{\varepsilon t}^t \mathrm{e}^{-\alpha S_r/\tau_t} \,\mathrm{d}r \le t\mathrm{e}^{-\alpha S_{\varepsilon t}/\tau_t}. \tag{2.7.16}$$

According to [29, Chapter III, Theorem 11], we have that $\mathbb{P}$-a.s.

$$\liminf_{\varepsilon\to 0} \frac{S_{\varepsilon t}}{h(\varepsilon t)} = \gamma - 1 > 0,$$

where $h(r) = r^{\gamma/(\gamma-1)} \log\left(|\log r|\right)^{-1/(\gamma-1)}$. As a consequence, there exist a positive random variable $\rho = \rho(\omega)$ and a constant $c > 0$ such that $\mathbb{P}$-a.s. $S_{\varepsilon t} \ge c h(\varepsilon t)$ for every $\varepsilon \in (0, \rho)$. We deduce that for every $t > 0$, $\mathbb{P}$-a.s.

$$\begin{aligned}
\limsup_{\alpha\to\infty} \alpha^{1-1/\gamma} \mathrm{e}^{-\alpha S_{\varepsilon t}/\tau_t} &\le \limsup_{\alpha\to\infty} \alpha^{1-1/\gamma} \mathrm{e}^{-c\alpha h(\varepsilon t)/\tau_t} \\
&= \limsup_{\alpha\to\infty} \alpha^{1-1/\gamma} \mathrm{e}^{-c t^{\gamma/(\gamma-1)} \alpha^\delta \log(|\log(\varepsilon t)|)^{-1}/\tau_t} = 0,
\end{aligned}$$

where in the second to last equality we used (2.5.2). In conjunction with (2.7.15) and (2.7.16), this yields (2.7.14).

Let $\eta > 0$. Using that $\mathfrak{h}_{r,U} \le \mathfrak{h}$, we have

$$\begin{aligned}\limsup_{\alpha\to\infty} \mathbf{N}^{(1)}\left[J^1_\alpha > \eta\right] &\le \limsup_{\alpha\to\infty} \mathbf{N}^{(1)}\left[\alpha^{1-1/\gamma}\int_{\varepsilon H(U)}^{H(U)} \sigma^\alpha_{r,U}\,\mathrm{d}r > \eta\right] \\ &= \limsup_{\alpha\to\infty} \mathbf{N}^{(1)}\left[\mu\left(x\in\mathscr{T}:\ \alpha^{1-1/\gamma}\int_{\varepsilon H(x)}^{H(x)} \sigma^\alpha_{r,x}\,\mathrm{d}r > \eta\right)\right],\end{aligned}$$

where the last term vanishes thanks to (2.7.14) and the dominated convergence theorem. This gives that $\lim_{\alpha\to\infty} J^1_\alpha = 0$ in $\mathbf{N}^{(1)}$-probability.

Under $\mathbf{N}^{(1)}$, let x^* be the unique leaf realizing the total height, that is the unique $x\in\mathscr{T}$ such that $H(x) = \mathfrak{h}$. Then $\mathbf{N}^{(1)}$-a.s. we have $H(U\wedge x^*) > 0$ and, thanks to (2.2.6), $\mathfrak{h}_{r,U} = \mathfrak{h} - r$ for every $r\in[0,\varepsilon H(U)]$ if $\varepsilon > 0$ is small enough (more precisely for $\varepsilon \le H(U\wedge x^*)/H(U)$). In particular, this implies that $\mathbf{N}^{(1)}$-a.s. $\lim_{\alpha\to\infty} J^2_\alpha = 0$.

Next, we have

$$\begin{aligned}|J^3_\alpha| &\le \alpha^{1-1/\gamma}\int_0^{\varepsilon H(U)} \sigma^\alpha_{r,U}\left|\left(1-\frac{r}{\mathfrak{h}}\right)^\beta - \mathrm{e}^{-\beta r/\mathfrak{h}}\right|\mathrm{d}r \\ &\le \alpha^{1-1/\gamma}\beta\int_0^{\varepsilon H(U)} \sigma^\alpha_{r,U}\left|\log\left(1-\frac{r}{\mathfrak{h}}\right)+\frac{r}{\mathfrak{h}}\right|\mathrm{e}^{-\beta r/\mathfrak{h}}\,\mathrm{d}r \\ &\le \alpha^{1-1/\gamma}\beta j\left(\frac{\varepsilon H(U)}{\mathfrak{h}}\right)\int_0^{\varepsilon H(U)} \sigma^\alpha_{r,U}\frac{r^2}{\mathfrak{h}^2}\mathrm{e}^{-\beta r/\mathfrak{h}}\,\mathrm{d}r \\ &\le C H(U)\, j(\varepsilon)\,\varepsilon^3\alpha^{2(1-1/\gamma)},\end{aligned}$$

where we used that $|\mathrm{e}^b - \mathrm{e}^a| \le |b-a|\mathrm{e}^b$ for $a\le b$ for the second inequality, that the function $j\colon y\mapsto -(y+\log(1-y))/y^2$ is increasing on $[0,1)$ for the third and the fact that $H(U)\le\mathfrak{h}$ and $\beta/\alpha^{1-1/\gamma}$ is bounded by some constant $C>0$ for the last. Using (2.5.2), notice that $\varepsilon^3\alpha^{2(1-1/\gamma)} = \alpha^{(3\delta-1)(1-1/\gamma)} \to 0$ as $\delta < 1/3$. Since $\lim_{y\to 0} j(y) = 1/2$, we deduce that $\mathbf{N}^{(1)}$-a.s. $\lim_{\alpha\to\infty} J^3_\alpha = 0$.

Finally, we have

$$\begin{aligned}|J^4_\alpha| &\le \alpha^{2-1/\gamma}\int_0^{\varepsilon H(U)} \left|\log\left(\sigma_{r,U}\right)+1-\sigma_{r,U}\right|\mathrm{e}^{-\alpha(1-\sigma_{r,U})}\,\mathrm{d}r \\ &\le j\left(1-\sigma_{\varepsilon H(U),U}\right)\alpha^{2-1/\gamma}\int_0^{\varepsilon H(U)}\left(1-\sigma_{r,U}\right)^2\mathrm{e}^{-\alpha(1-\sigma_{r,U})}\,\mathrm{d}r \\ &\le C H(U)\, j\left(1-\sigma_{\varepsilon H(U),U}\right)\alpha^{-1/\gamma}\varepsilon,\end{aligned}$$

where we used that $|\mathrm{e}^b - \mathrm{e}^a| \le |b-a|\mathrm{e}^b$ for $a\le b$ for the first inequality, that the function $j\colon x\mapsto -(x+\log(1-x))/x^2$ is increasing on $[0,1)$ for the second and that the function $x\mapsto x^2\mathrm{e}^{-x}$ is bounded on $[0,\infty)$ for the last. Since $\lim_{x\to 0} j(x) = 1/2$, $\lim_{\varepsilon\to 0}\sigma_{\varepsilon H(U),U} = 1$ and $\alpha^{-1/\gamma}\varepsilon\to 0$, we deduce that $\mathbf{N}^{(1)}$-a.s. $\lim_{\alpha\to\infty} J^4_\alpha = 0$.

2.7.3 Proof of Lemma 2.5.3

It is enough to show that for every Lipschitz-continuous and bounded function $f\colon [0,\infty)\to\mathbb{R}$

$$\lim_{\alpha\to\infty}\mathbf{N}^{(1)}\left[\int_{\mathcal{T}}\mu(\mathrm{d}x)f\left(\alpha^{1-1/\gamma}\mathfrak{h}^{-\beta}\left(Z_{\alpha,\beta}(x)-\mathbf{Z}_{\alpha,\beta}\right)\right)\right]=f(0).$$

Let $\varepsilon=\alpha^{(\delta-1)(1-1/\gamma)}$ with $\delta\in(0,1/2)$. For every $x\in\mathcal{T}$ such that $H(x)\geq\varepsilon$, set

$$Z^{\varepsilon}_{\alpha,\beta}(x)=\int_0^{\varepsilon}\sigma^{\alpha}_{r,x}\mathfrak{h}^{\beta}_{r,x}\,\mathrm{d}r\quad\text{and}\quad\mathbf{Z}^{\varepsilon}_{\alpha,\beta}=\int_{\mathcal{T}}\mathbf{1}_{\{H(x)\geq\varepsilon\}}Z^{\varepsilon}_{\alpha,\beta}(x)\,\mu(\mathrm{d}x).$$

Let $x^*\in\mathcal{T}$ be the unique leaf realizing the height, that is $H(x^*)=\mathfrak{h}$. Using that $\mathfrak{h}\geq H(x\wedge x^*)$ and that $Z^{\varepsilon}_{\alpha,\beta}(x)=Z^{\varepsilon}_{\alpha,\beta}(x^*)$ if $\varepsilon\leq H(x\wedge x^*)$, write

$$\int_{\mathcal{T}}\mu(\mathrm{d}x)f\left(\alpha^{1-1/\gamma}\mathfrak{h}^{-\beta}\left(Z_{\alpha,\beta}(x)-\mathbf{Z}_{\alpha,\beta}\right)\right)=\sum_{i=1}^{4}A^i_\alpha+B_\alpha,$$

where

$$\begin{aligned}
A^1_\alpha&=\int_{\mathcal{T}}\mu(\mathrm{d}x)\mathbf{1}_{\{H(x\wedge x^*)<\varepsilon\}}f\left(\alpha^{1-1/\gamma}\mathfrak{h}^{-\beta}\left(Z_{\alpha,\beta}(x)-\mathbf{Z}_{\alpha,\beta}\right)\right),\\
A^2_\alpha&=\int_{\mathcal{T}}\mu(\mathrm{d}x)\mathbf{1}_{\{H(x\wedge x^*)\geq\varepsilon\}}\left\{f\left(\alpha^{1-1/\gamma}\mathfrak{h}^{-\beta}\left(Z_{\alpha,\beta}(x)-\mathbf{Z}_{\alpha,\beta}\right)\right)\right.\\
&\qquad\qquad\left.-f\left(\alpha^{1-1/\gamma}\mathfrak{h}^{-\beta}\left(Z^{\varepsilon}_{\alpha,\beta}(x)-\mathbf{Z}_{\alpha,\beta}\right)\right)\right\},\\
A^3_\alpha&=\int_{\mathcal{T}}\mu(\mathrm{d}x)\mathbf{1}_{\{H(x\wedge x^*)\geq\varepsilon\}}\left\{f\left(\alpha^{1-1/\gamma}\mathfrak{h}^{-\beta}\left(Z^{\varepsilon}_{\alpha,\beta}(x)-\mathbf{Z}_{\alpha,\beta}\right)\right)\right.\\
&\qquad\qquad\left.-f\left(\alpha^{1-1/\gamma}\mathfrak{h}^{-\beta}\left(Z^{\varepsilon}_{\alpha,\beta}(x)-\mathbf{Z}^{\varepsilon}_{\alpha,\beta}\right)\right)\right\},\\
A^4_\alpha&=-\mu\left(\left\{x\in\mathcal{T}:\ H(x\wedge x^*)<\varepsilon\right\}\right)f\left(\mathbf{1}_{\{\mathfrak{h}\geq\varepsilon\}}\alpha^{1-1/\gamma}\mathfrak{h}^{-\beta}\left(Z^{\varepsilon}_{\alpha,\beta}(x^*)-\mathbf{Z}^{\varepsilon}_{\alpha,\beta}\right)\right),\\
B_\alpha&=f\left(\mathbf{1}_{\{\mathfrak{h}\geq\varepsilon\}}\alpha^{1-1/\gamma}\mathfrak{h}^{-\beta}\left(Z^{\varepsilon}_{\alpha,\beta}(x^*)-\mathbf{Z}^{\varepsilon}_{\alpha,\beta}\right)\right).
\end{aligned}$$

Thanks to the dominated convergence theorem, we have

$$\lim_{\alpha\to\infty}\mathbf{N}^{(1)}[|A^1_\alpha+A^4_\alpha|]\leq 2\,\|f\|_\infty\lim_{\alpha\to\infty}\mathbf{N}^{(1)}\left[\int_{\mathcal{T}}\mu(\mathrm{d}x)\mathbf{1}_{\{H(x\wedge x^*)<\varepsilon\}}\right]=0.\tag{2.7.17}$$

Next, notice that

$$\begin{aligned}
\mathbf{N}^{(1)}[|A^2_\alpha|]&\leq\|f\|_{\mathrm{L}}\mathbf{N}^{(1)}\left[\alpha^{1-1/\gamma}\mathfrak{h}^{-\beta}\int_{\mathcal{T}}\mu(\mathrm{d}x)\mathbf{1}_{\{H(x\wedge x^*)\geq\varepsilon)\}}\left(Z_{\alpha,\beta}(x)-Z^{\varepsilon}_{\alpha,\beta}(x)\right)\right]\\
&\leq\|f\|_{\mathrm{L}}\mathbf{N}^{(1)}\left[\alpha^{1-1/\gamma}\int_{\mathcal{T}}\mu(\mathrm{d}x)\mathbf{1}_{\{H(x)\geq\varepsilon\}}\int_{\varepsilon}^{H(x)}\sigma^{\alpha}_{r,x}\,\mathrm{d}r\right],
\end{aligned}\tag{2.7.18}$$

where we used that $H(x\wedge x^*)\le H(x)$ and $\mathfrak{h}_{r,x}\le \mathfrak{h}$ for the second inequality. Now similarly to (2.7.14), we have $\mathbf{N}^{(1)}$-a.s. $\mu(\mathrm{d}x)$-a.s.

$$\lim_{\alpha\to\infty} \alpha^{1-1/\gamma}\mathbf{1}_{\{H(x)\ge\varepsilon\}}\int_\varepsilon^{H(x)} \sigma^\alpha_{r,x}\,\mathrm{d}r = 0. \tag{2.7.19}$$

Furthermore, applying Corollary 2.3.7, we have

$$\sup_{\alpha\ge 0}\alpha^{2-2/\gamma}\mathbf{N}^{(1)}\left[\int_{\mathscr{T}}\mu(\mathrm{d}x)\left(\mathbf{1}_{\{H(x)\ge\varepsilon\}}\int_\varepsilon^{H(x)}\sigma^\alpha_{r,x}\,\mathrm{d}r\right)^2\right] \\ \le \sup_{\alpha\ge 0}\alpha^{2-2/\gamma}\mathbf{N}^{(1)}\left[\int_{\mathscr{T}}\mu(\mathrm{d}x)\left(\int_0^{H(x)}\sigma^\alpha_{r,x}\,\mathrm{d}r\right)^2\right]<\infty.$$

We deduce that the family

$$\left(\alpha^{1-1/\gamma}\mathbf{1}_{\{H(x)\ge\varepsilon\}}\int_\varepsilon^{H(x)}\sigma^\alpha_{r,x}\,\mathrm{d}r :\ \alpha\ge 0\right)$$

is uniformly integrable under the measure $\mathbf{N}^{(1)}[\mathrm{d}\mathscr{T}]\mu(\mathrm{d}x)$. In conjunction with (2.7.19), this gives

$$\lim_{\alpha\to\infty}\mathbf{N}^{(1)}\left[\alpha^{1-1/\gamma}\int_{\mathscr{T}}\mathbf{1}_{\{H(x)\ge\varepsilon\}}\,\mu(\mathrm{d}x)\int_\varepsilon^{H(x)}\sigma^\alpha_{r,x}\,\mathrm{d}r\right]=0, \tag{2.7.20}$$

which, thanks to (2.7.18), implies that

$$\lim_{\alpha\to\infty}\mathbf{N}^{(1)}[|A^2_\alpha|]=0. \tag{2.7.21}$$

We have

$$\begin{aligned}\mathbf{N}^{(1)}[|A^3_\alpha|] &\le \|f\|_{\mathrm{L}}\mathbf{N}^{(1)}\left[\alpha^{1-1/\gamma}\mathfrak{h}^{-\beta}\int_{\mathscr{T}}\mu(\mathrm{d}x)\mathbf{1}_{\{H(x\wedge x^*)\ge\varepsilon\}}\left(\mathbf{Z}_{\alpha,\beta}-\mathbf{Z}^\varepsilon_{\alpha,\beta}\right)\right]\\ &\le \|f\|_{\mathrm{L}}\mathbf{N}^{(1)}\left[\alpha^{1-1/\gamma}\mathfrak{h}^{-\beta}\left(\mathbf{Z}_{\alpha,\beta}-\mathbf{Z}^\varepsilon_{\alpha,\beta}\right)\right]\\ &\le \|f\|_{\mathrm{L}}\mathbf{N}^{(1)}\left[\alpha^{1-1/\gamma}\int_{\mathscr{T}}\mathbf{1}_{\{H(x)\ge\varepsilon\}}\,\mu(\mathrm{d}x)\int_\varepsilon^{H(x)}\sigma^\alpha_{r,x}\,\mathrm{d}r\right]\\ &\quad+\|f\|_{\mathrm{L}}\mathbf{N}^{(1)}\left[\alpha^{1-1/\gamma}\int_{\mathscr{T}}\mathbf{1}_{\{H(x)<\varepsilon\}}\,\mu(\mathrm{d}x)\int_0^{H(x)}\sigma^\alpha_{r,x}\,\mathrm{d}r\right],\end{aligned} \tag{2.7.22}$$

where we used that $\mathfrak{h}_{r,x}\le\mathfrak{h}$ for the last inequality. Let $p\in(1,2)$ and notice that $\varepsilon^{1+p}\alpha^{1-1/\gamma}\to 0$. Using that $\sigma_{r,x}\le 1$ together with the Markov inequality, we get

$$\mathbf{N}^{(1)}\left[\alpha^{1-1/\gamma}\int_{\mathscr{T}}\mathbf{1}_{\{H(x)<\varepsilon\}}\,\mu(\mathrm{d}x)\int_0^{H(x)}\sigma^\alpha_{r,x}\,\mathrm{d}r\right]\le\mathbf{N}^{(1)}\left[\varepsilon\alpha^{1-1/\gamma}\int_{\mathscr{T}}\mathbf{1}_{\{H(x)<\varepsilon\}}\,\mu(\mathrm{d}x)\right]\\ \le\varepsilon^{1+p}\alpha^{1-1/\gamma}\mathbf{N}^{(1)}\left[\int_{\mathscr{T}}H(x)^{-p}\,\mu(\mathrm{d}x)\right].$$

By Lemma 2.3.9, the last term is finite. This, in conjunction with (2.7.20) and (2.7.22), implies that

$$\lim_{\alpha\to\infty} \mathbf{N}^{(1)}[|A_\alpha^3|] = 0. \tag{2.7.23}$$

It remains to show that $\lim_{\alpha\to\infty}\mathbf{N}^{(1)}[B_\alpha] = f(0)$, which is equivalent to the following convergence in $\mathbf{N}^{(1)}$-probability

$$\lim_{\alpha\to\infty} \mathbf{1}_{\{\mathfrak{h}\geq\varepsilon\}}\alpha^{1-1/\gamma}\mathfrak{h}^{-\beta}\left(Z^{\varepsilon}_{\alpha,\beta}(x^*) - \mathbf{Z}^{\varepsilon}_{\alpha,\beta}\right) = 0. \tag{2.7.24}$$

Again using that $Z^{\varepsilon}_{\alpha,\beta}(x) = Z^{\varepsilon}_{\alpha,\beta}(x^*)$ if $\varepsilon \leq H(x\wedge x^*)$, we write

$$\mathbf{1}_{\{\mathfrak{h}\geq\varepsilon\}}\alpha^{1-1/\gamma}\mathfrak{h}^{-\beta}\left(Z^{\varepsilon}_{\alpha,\beta}(x^*) - \mathbf{Z}^{\varepsilon}_{\alpha,\beta}\right) = B^1_\alpha + B^2_\alpha,$$

where

$$\begin{aligned}
B^1_\alpha &= \alpha^{1-1/\gamma}\mathfrak{h}^{-\beta}\left(\mathbf{1}_{\{\mathfrak{h}\geq\varepsilon\}}Z^{\varepsilon}_{\alpha,\beta}(x^*) - \int_{\mathcal{T}}\mu(\mathrm{d}x)\mathbf{1}_{\{H(x\wedge x^*)\geq\varepsilon\}}Z^{\varepsilon}_{\alpha,\beta}(x^*)\right),\\
B^2_\alpha &= \alpha^{1-1/\gamma}\mathfrak{h}^{-\beta}\left(\int_{\mathcal{T}}\mu(\mathrm{d}x)\mathbf{1}_{\{H(x\wedge x^*)\geq\varepsilon\}}Z^{\varepsilon}_{\alpha,\beta}(x) - \mathbf{1}_{\{\mathfrak{h}\geq\varepsilon\}}\mathbf{Z}^{\varepsilon}_{\alpha,\beta}\right).
\end{aligned}$$

Recall that $\varepsilon = \alpha^{(\delta-1)(1-1/\gamma)} \to 0$ as $\alpha\to\infty$. Fix $\eta>0$ and let $\alpha_0>0$ be large enough so that for every $\alpha\geq\alpha_0$

$$\mathbf{N}^{(1)}\left[\int_{\mathcal{T}}\mu(\mathrm{d}x)\mathbf{1}_{\{H(x\wedge x^*)<\varepsilon\}}\right] \leq \eta.$$

Then we have for every $\alpha\geq\alpha_0$ and $C>0$

$$\begin{aligned}
&\mathbf{N}^{(1)}\left[\alpha^{1-1/\gamma}\mathfrak{h}^{-\beta}Z^{\varepsilon}_{\alpha,\beta}(x^*)\mathbf{1}_{\{\mathfrak{h}\geq\varepsilon\}} \geq C\right]\\
&\quad\leq \mathbf{N}^{(1)}\left[\int_{\mathcal{T}}\mu(\mathrm{d}x)\mathbf{1}_{\left\{\alpha^{1-1/\gamma}\mathfrak{h}^{-\beta}Z^{\varepsilon}_{\alpha,\beta}(x)\geq C,\, H(x\wedge x^*)\geq\varepsilon\right\}}\right] + \mathbf{N}^{(1)}\left[\int_{\mathcal{T}}\mu(\mathrm{d}x)\mathbf{1}_{\{H(x\wedge x^*)<\varepsilon\}}\right]\\
&\quad\leq \frac{\alpha^{2-2/\gamma}}{C^2}\mathbf{N}^{(1)}\left[\int_{\mathcal{T}}\mu(\mathrm{d}x)\mathbf{1}_{\{H(x\wedge x^*)\geq\varepsilon\}}\left(\mathfrak{h}^{-\beta}Z^{\varepsilon}_{\alpha,\beta}(x)\right)^2\right] + \eta\\
&\quad\leq \frac{\alpha^{2-2/\gamma}}{C^2}\mathbf{N}^{(1)}\left[\int_{\mathcal{T}}\mu(\mathrm{d}x)\left(\int_0^{H(x)}\sigma^{\alpha}_{r,x}\,\mathrm{d}r\right)^2\right] + \eta\\
&\quad\leq \frac{M}{C^2} + \eta
\end{aligned} \tag{2.7.25}$$

for some constant $M>0$, where we used that $Z^{\varepsilon}_{\alpha,\beta}(x^*) = Z^{\varepsilon}_{\alpha,\beta}(x)$ for every $x\in\mathcal{T}$ such that $H(x\wedge x^*)\geq\varepsilon$ for the first inequality, the Markov inequality for the second and Corollary 2.3.7 for the last. Thus, we get that the family $(\mathbf{1}_{\{\mathfrak{h}\geq\varepsilon\}}\alpha^{1-1/\gamma}\mathfrak{h}^{-\beta}Z^{\varepsilon}_{\alpha,\beta}(x^*)\colon\ \alpha\geq\alpha_0, \beta\geq 0)$ is tight. Since $\mathbf{N}^{(1)}$-a.s.

$$\lim_{\alpha\to\infty}\int_{\mathcal{T}}\mu(\mathrm{d}x)\mathbf{1}_{\{H(x\wedge x^*)<\varepsilon\}} = 0,$$

we deduce the following convergence in $\mathbf{N}^{(1)}$-probability

$$\lim_{\alpha\to\infty} B^1_\alpha = \lim_{\alpha\to\infty} \mathbf{1}_{\{\mathfrak{h}\geq\varepsilon\}} \alpha^{1-1/\gamma} \mathfrak{h}^{-\beta} Z^{\varepsilon}_{\alpha,\beta}(x^*) \int_{\mathcal{T}} \mu(\mathrm{d}x) \mathbf{1}_{\{H(x\wedge x^*)<\varepsilon\}} = 0.$$

Furthermore, we have

$$\begin{aligned}
\mathbf{N}^{(1)}[|B^2_\alpha|] &= \alpha^{1-1/\gamma} \mathbf{N}^{(1)}\left[\int_{\mathcal{T}} \mu(\mathrm{d}x) \mathbf{1}_{\{H(x)\geq\varepsilon,\, H(x\wedge x^*)<\varepsilon\}} \mathfrak{h}^{-\beta} Z^{\varepsilon}_{\alpha,\beta}(x)\right] \\
&\leq \alpha^{1-1/\gamma} \mathbf{N}^{(1)}\left[\int_{\mathcal{T}} \mu(\mathrm{d}x) \left(\mathbf{1}_{\{H(x\wedge x^*)<\varepsilon\}} \int_0^{H(x)} \sigma^{\alpha}_{r,x}\, \mathrm{d}r\right)\right] \\
&\leq \alpha^{1-1/\gamma} \mathbf{N}^{(1)}\left[\int_{\mathcal{T}} \mu(\mathrm{d}x) \left(\int_0^{H(x)} \sigma^{\alpha}_{r,x}\, \mathrm{d}r\right)^2\right]^{1/2} \mathbf{N}^{(1)}\left[\int_{\mathcal{T}} \mu(\mathrm{d}x) \mathbf{1}_{\{H(x\wedge x^*)<\varepsilon\}}\right]^{1/2} \\
&\leq C\mathbf{N}^{(1)}\left[\int_{\mathcal{T}} \mu(\mathrm{d}x) \mathbf{1}_{\{H(x\wedge x^*)<\varepsilon\}}\right]^{1/2}
\end{aligned}$$

for some constant $C > 0$, where we used the Cauchy-Schwarz inequality for the second inequality and Corollary 2.3.7 for the last. It follows from the dominated convergence theorem that $\lim_{\alpha\to\infty} \mathbf{N}^{(1)}[|B^2_\alpha|] = 0$. This finishes the proof of (2.7.24).

3 Conditioning (sub)critical Lévy trees by their maximal degree

This chapter is based on the preprint [6].

We study the maximal degree of (sub)critical Lévy trees which arise as the scaling limits of Bienaymé-Galton-Watson trees. We determine the genealogical structure of large nodes and establish a Poissonian decomposition of the tree along those nodes. Furthermore, we make sense of the distribution of the Lévy tree conditioned to have a fixed maximal degree. In the case where the Lévy measure is diffuse, we show that the maximal degree is realized by a unique node whose height is exponentially distributed and we also prove that the conditioned Lévy tree can be obtained by grafting a Lévy forest on an independent size-biased Lévy tree with a degree constraint at a uniformly chosen leaf. Finally, we show that the Lévy tree conditioned on having large maximal degree converges locally to an immortal tree (which is the continuous analogue of the Kesten tree) in the critical case and to a condensation tree in the subcritical case. Our results are formulated in terms of the exploration process which allows to drop the Grey condition.

3.1 Introduction and main results

Lévy trees are random metric spaces that encode the genealogical structure of continuous-state branching processes (CB processes for short). As such, they arise as the scaling limits of Bienaymé-Galton-Watson trees. Lévy trees were introduced by Le Gall and Le Jan [119] and Duquesne and Le Gall [57] in order to generalize Aldous' Brownian tree [16]. They also appear as scaling limits of various models of trees and graphs, see e.g. Haas and Miermont [82], and are naturally related to fragmentation processes, see Miermont [124, 125], Haas and Miermont [81], Abraham and Delmas [1].

In the present paper, we study the maximal degree of a general Lévy tree. More precisely, we first establish a Poissonian decomposition of the Lévy tree along large nodes. Then, we make sense of the distribution of the Lévy tree conditioned to have a fixed maximal degree. In the case where the Lévy measure is diffuse, we show that the maximal degree is realized by a unique node, and we describe how to reconstruct the tree by grafting a Lévy forest on an independent size-biased Lévy tree (with a restriction on the maximal degree) at a uniform leaf. Finally, we investigate the asymptotic behavior of the Lévy tree conditioned to have large maximal degree.

These questions arise naturally in the study of random trees and have been thoroughly investigated in the case of Bienaymé-Galton-Watson trees. The first results in this direction were obtained by Jonsson and Stefánsson [98] who showed that a condensation phenomenon appears for a certain class of subcritical Bienaymé-Galton-Watson trees conditioned to have a large size, in the sense that with high probability there exists a unique node with degree proportional to the size. Furthermore, the tree converges locally to a condensation tree consisting of a finite spine with random geometric length onto which independent and identically distributed Bienaymé-Galton-Watson trees are grafted. This was later generalized by Janson [94], with further results by Kortchemski [111], Abraham and Delmas [4], Stufler [146]. On the other hand, He [85] shows that Bienaymé-Galton-Watson trees conditioned on having large maximal degree converge locally to Kesten's tree (which consists of an infinite spine onto which independent and identically distributed Bienaymé-Galton-Watson trees are grafted) in the critical case and to a condensation tree in the subcritical case.

In the continuum setting, Bertoin [34] determined the distribution of the maximal degree of a stable Lévy tree (his result is formulated in terms of Lévy processes). Using the formalism of CB processes, He and Li [88] treated the case of a general branching mechanism (in fact their result is more general as they considered CB processes with immigration). In [87], they also studied the local limit of a CB process conditioned to have large maximal degree (i.e. large maximal jump). In the critical case, they showed that it converges locally to a CB process with immigration. Later , He [86] extended the local convergence result to the whole genealogy: more precisely, he showed that a critical Lévy tree conditioned on having large maximal degree converges locally to an immortal tree (which is the continuous counterpart of Kesten's tree, consisting of an infinite spine onto which trees are grafted according to a Poisson point process). We improve these results by considering the density version of the conditioning instead of the tail version: more explicitly, we study the asymptotic behavior of critical Lévy trees conditioned to have maximal degree equal to (and not larger than) a given value. Density versions are finer than their tail counterparts and are usually more difficult to prove.

The existing litterature in the subcritical case is less developped. He and Li [87] showed that a subcritical CB process conditioned to have large maximal degree converges locally to a CB process with immigration which is killed (i.e. sent to infinity) at an independent exponential time, thus underlining a condensation phenomenon. We improve this result in several directions. Again we consider the density version of the conditioning instead of the tail version. We also extend the convergence result to the whole genealogical structure instead of the population size at a given time: this gives more information and, as an example, allows us to see that only one large node emerges. Finally, we are also able to describe precisely what happens above the condensation node.

For the sake of clarity, we shall formulate our results in terms of Lévy trees in the introduction. This requires an additional assumption on the branching mechanism, namely the Grey condition (see below), in order to have a nice topology on the set of trees. Indeed, this condition ensures that the Lévy tree is a *compact* real tree. However, it is superfluous and will be dropped in the rest of the paper where we will deal with the exploration process instead. Let us mention that a forthcoming work by Duquesne and Winkel [61] should allow us to use the formalism of real trees even for a general branching mechanism not necessarily satisfying the Grey condition.

Before stating our main results, we need to recall some definitions and to set notations.

3.1.1 Real trees

We recall the formalism of real trees, see [69]. A quadruple $(T, d, \emptyset, \mu)$ is called a real tree if (T, d) is a metric space equipped with a distinguished vertex $\emptyset \in T$ called the root and a nonnegative finite measure μ on T and if the following two properties hold for every $x, y \in T$:

(i) (Unique geodesics). There exists a unique isometric map $f_{x,y}\colon [0, d(x, y)] \to T$ such that $f_{x,y}(0) = x$ and $f_{x,y}(d(x, y)) = y$.

(ii) (Loop-free). If φ is a continuous injective map from $[0,1]$ into T such that $\varphi(0) = x$ and $\varphi(1) = y$, then we have $\varphi([0,1]) = f_{x,y}\big([0, d(x, y)]\big)$.

For every vertex $x \in T$, we define its height by $H(x) = d(\emptyset, x)$. The height of the tree is defined by $\mathfrak{h}(T) = \sup_{x\in T} H(x)$. Note that if (T, d) is compact, then $\mathfrak{h}(T) < \infty$.

We will denote by $\mathbb{T}$ the set of (isometry classes of) *compact* real trees. Let us mention that it can be equipped with the Gromov-Hausdorff-Prokhorov distance which makes it a Polish space, see e.g. [13].

We will also need the set $\mathbb{T}^*$ of (isometry classes of) compact real trees that are *marked*, i.e. equipped with a distinguished vertex in addition to the root $\emptyset$. Again, $\mathbb{T}^*$ can be made into a Polish space when equipped with a marked variant of the Gromov-Hausdorff-Prokhorov distance.

3.1.2 Local convergence of real trees

We will make use of the notion of local convergence for *locally compact* real trees which we now recall. For every $h > 0$, define the restriction mapping on the set of (isometry classes of) real trees by:

$$r_h(T, d, \emptyset, \mu) = (T^h, d_{|T^h\times T^h}, \emptyset, \mu_{|T^h}) \quad \text{where } T^h = \{x \in T\colon\ H(x) \le h\}.$$

In other words, $r_h(T)$ is the real tree obtained from T by removing all nodes whose height is larger than h, equipped with the same metric and measure restricted to T^h. Recall that the Hopf-Rinow theorem implies that if T is a locally compact real tree, the closed ball $r_h(T)$ is compact. We say that a sequence T_n of locally compact trees converges locally to a locally compact tree T if for every $h>0$, the sequence $r_h(T_n)$ converges for the Gromov-Hausdorff-Prokhorov distance to $r_h(T)$.

3.1.3 Grafting procedure

Given a real tree $T \in \mathbb{T}$ and a finite or countable family $((x_i, T_i), i \in I)$ of elements of $T \times \mathbb{T}$, we denote by

$$T \circledast_{i \in I} (x_i, T_i)$$

the real tree obtained by grafting T_i on T at the node x_i. For a precise definition, we refer the reader to [3, Section 2.4].

3.1.4 Lévy trees

Let ψ be a branching mechanism given by:

$$\psi(\lambda) = \alpha\lambda + \beta\lambda^2 + \int_{(0,\infty)} \left(e^{-\lambda r} - 1 + \lambda r\right) \pi(dr), \tag{3.1.1}$$

where $\alpha, \beta \geq 0$ and π is a σ-finite measure on $(0,\infty)$ such that $\int_{(0,\infty)} (r \wedge r^2)\, \pi(dr) < \infty$. The branching mechanism ψ is said to be critical (resp. subcritical) if $\alpha = 0$ (resp. $\alpha > 0$). In what follows, we assume that $\pi \neq 0$ as otherwise all branching points of the Lévy tree will be binary. Whenever we are dealing with Lévy trees, we always assume that the Grey condition holds:

$$\int^{\infty} \frac{d\lambda}{\psi(\lambda)} < \infty, \tag{3.1.2}$$

which is equivalent to the compactness of the Lévy tree. In the rest of the paper, this condition will be relaxed to:

$$\beta > 0 \quad \text{or} \quad \int_{(0,1)} r\,\pi(dr) = \infty. \tag{3.1.3}$$

We will consider a Lévy tree $\mathcal{T}$ under its excursion measure which is denoted by $\mathbf{N}^\psi$. Here we briefly recall some results on Lévy trees but we refer the reader to Duquesne and Le Gall [57, 58] for a complete presentation on the subject. One can define a σ-finite measure $\mathbf{N}^\psi$ on the space $\mathbb{T}$, called the excursion measure of the Lévy tree, with the following properties.

(i) **Mass measure.** For $\mathbf{N}^\psi$-almost every $\mathcal{T}$, the mass measure μ is supported on the set of leaves $\mathrm{Lf}(\mathcal{T}) := \{x \in \mathcal{T} : \mathcal{T} \setminus \{x\} \text{ is connected}\}$. Furthermore, the total mass $\sigma := \mu(\mathcal{T})$ satisfies:

$$\mathbf{N}^\psi\left[1 - e^{-\lambda\sigma}\right] = \psi^{-1}(\lambda). \tag{3.1.4}$$

(ii) **Local times.** For $\mathbf{N}^\psi$-almost every $\mathscr{T}$, there exists a process $(L^a,\, a \geq 0)$ with values in the space of finite measures on $\mathscr{T}$ which is càdlàg for the weak topology and such that

$$\mu(\mathrm{d}x) = \int_0^\infty \mathrm{d}a\, L^a(\mathrm{d}x). \tag{3.1.5}$$

For every $a \geq 0$, the measure L^a is supported on $\mathscr{T}(a) := \{x \in \mathscr{T} :\ H(x) = a\}$ the set of nodes at height a. Furthermore, the real-valued process $(L^a_\sigma := \langle L^a, 1\rangle,\, a \geq 0)$ is a ψ-CB process under its canonical measure.

(iii) **Branching property.** For every $a \geq 0$, let $(\mathscr{T}^i,\, i \in I_a)$ be the subtrees of $\mathscr{T}$ originating from level a. Then, under $\mathbf{N}^\psi$ and conditionally on $r_a(\mathscr{T}) := \{x \in \mathscr{T} :\ H(x) \leq a\}$, the measure $\sum_{i\in I_a} \delta_{\mathscr{T}^i}$ is a Poisson point measure with intensity $L^a_\sigma \mathbf{N}^\psi$.

(iv) **Branching points.** For $\mathbf{N}^\psi$-almost every $\mathscr{T}$, the branching points of $\mathscr{T}$ are either binary or of infinite degree. The set of binary branching points is empty if $\beta = 0$ and is a countable dense subset of $\mathscr{T}$ if $\beta > 0$. The set

$$\mathrm{Br}_\infty(\mathscr{T}) := \{x \in \mathscr{T} :\ \mathscr{T} \setminus \{x\} \text{ has infinitely many connected components}\}$$

of infinite branching points is nonempty with $\mathbf{N}^\psi$-positive measure if and only if $\pi \neq 0$. If $\langle \pi, 1\rangle = \infty$, the set $\mathrm{Br}_\infty(\mathscr{T})$ is countable and dense in $\mathscr{T}$ for $\mathbf{N}^\psi$-almost every $\mathscr{T}$. Furthermore, the set $\{H(x),\, x \in \mathrm{Br}_\infty(\mathscr{T})\}$ coincides with the set of discontinuity times of the mapping $a \mapsto L^a$. For every such discontinuity time a, there is a unique $x_a \in \mathrm{Br}_\infty(\mathscr{T}) \cap \mathscr{T}(a)$ and $\Delta_a > 0$ such that

$$L^a = L^{a-} + \Delta_a \delta_{x_a}.$$

For convenience, we define Δ_a for every $a \geq 0$ by setting $\Delta_a = 0$ if $L^a = L^{a-}$. In particular, we have $L^a_\sigma = L^{a-}_\sigma + \Delta_a$, that is Δ_a is exactly the size of the jump of the associated CB process at time a. We will call Δ_a the degree (or the mass) of the node x_a. This is an abuse of language since a node $x_a \in \mathrm{Br}_\infty(\mathscr{T})$ has infinite degree by definition.

3.1.5 Main results

We denote by Δ the maximal degree of the Lévy tree $\mathscr{T}$ under $\mathbf{N}^\psi$:

$$\Delta = \sup_{a\geq 0} \Delta_a. \tag{3.1.6}$$

The first result of this paper gives the joint distribution of the maximal degree Δ and the total mass σ under $\mathbf{N}^\psi$. The distribution of the maximal degree was already obtained by Bertoin [34, Lemma 1] in the stable Lévy case then by He and Li [88] in the general case.

For the sake of notational simplicity, if ν is a measure on $\mathbb{R}$ we will write $\nu(a,b)$ (resp. $\nu[a,b)$) instead of $\nu((a,b))$ (resp. $\nu([a,b))$). We will also write $\nu(a)$ for $\nu(\{a\})$. Denote by $\bar\pi : \mathbb{R}_+ \to (0,\infty]$

the tail of the Lévy measure π:

$$\bar{\pi}(\delta) = \pi(\delta, \infty), \quad \forall \delta \geq 0, \tag{3.1.7}$$

and define the Laplace exponent ψ_δ for every $\delta > 0$ by:

$$\begin{aligned}\psi_\delta(\lambda) &= \left(\alpha + \int_{(\delta,\infty)} r\,\pi(\mathrm{d}r)\right)\lambda + \beta\lambda^2 + \int_{(0,\delta]} \left(\mathrm{e}^{-\lambda r} - 1 + \lambda r\right)\pi(\mathrm{d}r) \\ &= \psi(\lambda) + \int_{(\delta,\infty)} \left(1 - \mathrm{e}^{-\lambda r}\right)\pi(\mathrm{d}r). \end{aligned} \tag{3.1.8}$$

Observe that, in terms of the associated Lévy process, this corresponds to removing all jumps with size larger than δ. If the Lévy measure π is finite, we also define:

$$\psi_0(\lambda) = \left(\alpha + \int_{(0,\infty)} r\,\pi(\mathrm{d}r)\right)\lambda + \beta\lambda^2. \tag{3.1.9}$$

Proposition 3.1.1. *For every $\delta > 0$ and $\lambda \geq 0$, we have:*

$$\mathbf{N}^\psi\left[1 - \mathrm{e}^{-\lambda\sigma}\mathbf{1}_{\{\Delta \leq \delta\}}\right] = \psi_\delta^{-1}(\bar{\pi}(\delta) + \lambda). \tag{3.1.10}$$

Furthermore, if the Lévy measure π is finite, we have:

$$\mathbf{N}^\psi\left[1 - \mathrm{e}^{-\lambda\sigma}\mathbf{1}_{\{\Delta = 0\}}\right] = \psi_0^{-1}(\langle \pi, 1\rangle + \lambda). \tag{3.1.11}$$

The proof is given in Section 3.3.

Remark 3.1.2. Let us make a connection with He and Li [87]. Recall that under $\mathbf{N}^\psi$ the process $(L_\sigma^a, a \geq 0)$ is distributed as a ψ-CB process under its canonical measure and that the maximal degree Δ of the Lévy tree corresponds to the maximal jump of the associated CB process. In particular, taking $\lambda = 0$ in (3.1.10) gives the distribution of the maximal jump of a ψ-CB process, which was already obtained by He and Li, see [87, Corollary 4.2]. In fact, their result is much more general (see [87, Theorem 4.1]) since they consider a CB process with immigration and in this context, they compute the distribution of the *local* maximal jump which in terms of the Lévy tree corresponds to the maximal degree up to a fixed level h. However, they do not give the *joint* distribution of Δ and σ, which in terms of the CB process corresponds to the total mass:

$$\sigma = \int_0^\infty L_\sigma^a\,\mathrm{d}a.$$

Next, we give a decomposition of the Lévy tree along the large nodes. More precisely, we identify the distribution of the pruned Lévy tree obtained by removing all nodes with degree larger than δ (and the subtrees above them). This is again a Lévy tree with branching mechanism ψ_δ under its excursion measure. Furthermore, one can recover the Lévy tree from the pruned one by grafting Lévy forests at uniformly chosen leaves in a Poissonian manner. Before stating the result, we first need to introduce some notations. For every $r > 0$, denote by $\mathbb{P}_r^\psi$

the distribution of the random real tree $\mathscr{T} = \{\emptyset\} \circledast_{i\in I} \mathscr{T}_i$ obtained by gluing together at their root the atoms $(\mathscr{T}_i,\, i \in I)$ of a $\mathbb{T}$-valued Poisson point measure with intensity $r\,\mathbf{N}^{\psi}[\mathrm{d}\mathscr{T}]$. This should be interpreted as the distribution of a Lévy forest with initial degree $r > 0$. Furthermore, for every $\delta > 0$ such that $\bar{\pi}(\delta) > 0$, set:

$$\mathbb{Q}^{\psi}_{\delta}(\mathrm{d}\mathscr{T}) = \frac{1}{\bar{\pi}(\delta)} \int_{(\delta,\infty)} \pi(\mathrm{d}r)\, \mathbb{P}^{\psi}_{r}(\mathrm{d}\mathscr{T})$$

which is the distribution of a Lévy forest with random initial degree with distribution π conditioned on being larger than δ.

Theorem 3.1.3. *Let $\delta \geq 0$ such that $\bar{\pi}(\delta) < \infty$. Under $\mathbf{N}^{\psi_\delta}$ and conditionally on $(\mathscr{T}, \emptyset, d, \mu)$, let $((x_i, \mathscr{T}_i),\, i \in I)$ be the atoms of a Poisson point measure on $\mathscr{T} \times \mathbb{T}$ with intensity $\bar{\pi}(\delta)\, \mu(\mathrm{d}x)\, \mathbb{Q}^{\psi}_{\delta}(\mathrm{d}\widetilde{\mathscr{T}})$. Then, under $\mathbf{N}^{\psi_\delta}$, the random tree $\mathscr{T} \circledast_{i\in I} (x_i, \mathscr{T}_i)$ has distribution $\mathbf{N}^{\psi}$.*

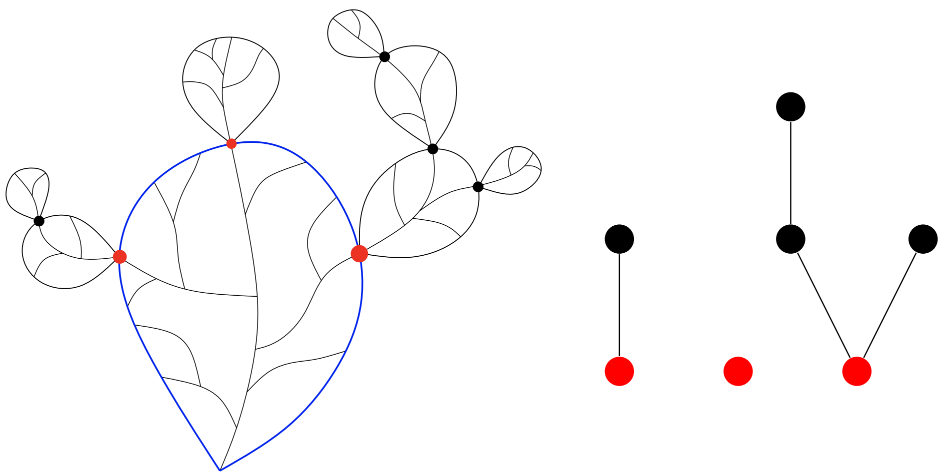

Figure 3.1 – Decomposition of the Lévy tree $\mathscr{T}$ along the nodes with degree larger than δ (left) and the associated discrete forest (right). In blue: the pruned subtree $\mathscr{T}^{\delta}$, in red: the first-generation nodes with degree larger than δ.

See Theorem 3.4.1 for a more precise statement. In particular, the pruned Lévy tree $\mathscr{T}^{\delta}$ which is obtained from $\mathscr{T}$ by removing all nodes with degree larger than δ is again a Lévy tree with branching mechanism ψ_δ. Thanks to this decomposition, we prove in Proposition 3.4.6 that the discrete forest formed by nodes with degree larger than δ is a Bienaymé-Galton-Watson forest and we specify its initial distribution and its offspring distribution, see Figure 3.1.

Remark 3.1.4. Theorem 3.1.3 is a special case of the main result in [8]. In that paper, the authors study a pruning procedure on the Lévy tree defined as follows: they add some marks on the skeleton of the tree according to a Poisson point measure with intensity $\alpha_1 \Lambda$ (where Λ is the length measure on $\mathscr{T}$ which is the equivalent of the Lebesgue measure) and add some other marks on the infinite branching points x_a with probability $p(\Delta_a)$ where p is a nonnegative measurable function satisfying:

$$\int_{(0,\infty)} r\, p(r)\, \pi(\mathrm{d}r) < \infty.$$

Then they show that the subtree $\mathscr{T}^{\alpha_1,p}$ containing the root obtained from $\mathscr{T}$ by removing all the marks is again a Lévy tree and identify its branching mechanism. Furthermore, they determine the distribution of the subtrees above the marks conditionally on $\mathscr{T}^{\alpha_1,p}$. It is obvious that the tree $\mathscr{T}^{\delta}$ coincides with $\mathscr{T}^{\alpha_1,p}$ where $\alpha_1 = 0$ and $p = \mathbf{1}_{(\delta,\infty)}$. Since p satisfies the integrability assumption above (as $\int_{(1,\infty)} r\,\pi(\mathrm{d}r) < \infty$), their result applies and gives the joint distribution of the pruned tree $\mathscr{T}^{\delta}$ and the subtrees originating from the nodes with degree larger than δ. However, the proof is much simpler in our particular setting.

One of our main results is the next theorem giving a decomposition of the Lévy tree at its largest nodes. Under $\mathbf{N}^{\psi}$, denote by M_δ the random variable defined by:

$$M_\delta = \frac{\mathrm{e}^{\mathfrak{g}(\delta)\sigma} - 1}{\mathfrak{g}(\delta)}, \quad \text{where} \quad \mathfrak{g}(\delta) = \pi(\delta)\mathrm{e}^{-\delta \mathbf{N}^{\psi}[\Delta>\delta]}.$$

This should be interpreted as $M_\delta = \sigma$ if $\mathfrak{g}(\delta) = 0$ (i.e. if δ is not an atom of π).

Theorem 3.1.5. *There exists a regular conditional probability $\mathbf{N}^{\psi}[\cdot|\Delta=\delta]$ for $\delta > 0$ such that $\pi[\delta,\infty) > 0$, which is given by, for every measurable and bounded $F\colon \mathbb{T} \to \mathbb{R}$:*

$$\begin{aligned}\mathbf{N}^{\psi}[F(\mathscr{T})|\Delta=\delta] = {} & \frac{1}{\mathbf{N}^{\psi}[M_\delta \mathbf{1}_{\{\Delta<\delta\}}]} \sum_{k=0}^{\infty} \frac{\mathfrak{g}(\delta)^k}{(k+1)!} \\ & \times \mathbf{N}^{\psi}\left[\int \prod_{i=1}^{k+1} \mu(\mathrm{d}x_i)\, \mathbb{P}^{\psi}_{\delta}(\mathrm{d}\mathscr{T}_i|\Delta \le \delta) F(\mathscr{T} \circledast_{i=1}^{k+1} (x_i,\mathscr{T}_i)) \mathbf{1}_{\{\Delta<\delta\}}\right], \end{aligned} \tag{3.1.12}$$

where $\mathbf{N}^{\psi}[M_\delta \mathbf{1}_{\{\Delta<\delta\}}] < \infty$. In particular, if $\delta > 0$ is not an atom of the Lévy measure π, we have:

$$\mathbf{N}^{\psi}[F(\mathscr{T})|\Delta=\delta] = \frac{1}{\mathbf{N}^{\psi}[\sigma \mathbf{1}_{\{\Delta<\delta\}}]} \mathbf{N}^{\psi}\left[\int \mu(\mathrm{d}x)\, \mathbb{P}^{\psi}_{\delta}(\mathrm{d}\widetilde{\mathscr{T}}|\Delta\le\delta) F(\mathscr{T} \circledast (x,\widetilde{\mathscr{T}})) \mathbf{1}_{\{\Delta<\delta\}}\right]. \tag{3.1.13}$$

Furthermore, if $\langle \pi, 1\rangle = \infty$, then $\mathbf{N}^{\psi}$-a.e. $\Delta > 0$, and if $\langle \pi, 1\rangle < \infty$, then we have:

$$\mathbf{N}^{\psi}[F(\mathscr{T})\mathbf{1}_{\{\Delta=0\}}] = \mathbf{N}^{\psi_0}[F(\mathscr{T})\mathrm{e}^{-\langle\pi,1\rangle\sigma}]. \tag{3.1.14}$$

The proof is given in Section 3.5. Some comments are in order.

(i) Recall that the distribution of Δ is given in Proposition 3.1.1. Together with the distribution of $\mathscr{T}$ conditionally on $\Delta=\delta$, we can recover the unconditional distribution of the Lévy tree via the disintegration formula:

$$\mathbf{N}^{\psi}[F(\mathscr{T})] = \mathbf{N}^{\psi}[F(\mathscr{T})\mathbf{1}_{\{\Delta=0\}}] + \int_{(0,\infty)} \mathbf{N}^{\psi}[\Delta \in \mathrm{d}\delta]\, \mathbf{N}^{\psi}[F(\mathscr{T})|\Delta=\delta],$$

where the first term on the right-hand side vanishes if π is infinite.

(ii) Assume that $\delta > 0$ is not an atom of π. Then, conditionally on $\Delta = \delta$, the Lévy tree can be constructed as follows: take $\widetilde{\mathscr{T}}$ with distribution $\mathbf{N}^{\psi}[\sigma \mathbf{1}_{\{\Delta \le \delta\}}]^{-1} \mathbf{N}^{\psi}[\cdot; \sigma \mathbf{1}_{\{\Delta \le \delta\}}]$, choose a leaf uniformly at random in $\widetilde{\mathscr{T}}$ (i.e. according to its normalized mass measure $\tilde{\sigma}^{-1}\tilde{\mu}$) and on this leaf graft an independent Lévy forest with initial degree δ conditioned to have maximal degree $\Delta \le \delta$. In fact, since δ is not an atom, this random forest will have no other nodes with degree δ besides the root. This entails that, conditionally on $\Delta = \delta$, there is a unique node realizing the maximum degree.

(iii) The situation is different when $\delta > 0$ is an atom of π. In that case, conditionally on $\Delta = \delta$, the number of first-generation nodes realizing the maximal degree has a Poisson distribution. More precisely, conditionally on $\Delta = \delta$, the Lévy tree can be constructed as follows: take $\widetilde{\mathscr{T}}$ with distribution $\mathbf{N}^{\psi}[M_\delta \mathbf{1}_{\{\Delta < \delta\}}]^{-1} \mathbf{N}^{\psi}[M_\delta \mathbf{1}_{\{\Delta < \delta\}}\, \mathrm{d}\mathscr{T}]$, and, conditionally on $\widetilde{\mathscr{T}}$, graft a Poisson point measure with intensity $\mathfrak{g}(\delta)\, \tilde{\mu}(\mathrm{d}x)\, \mathbb{P}^{\psi}_{\delta}(\mathrm{d}\mathscr{T} | \Delta \le \delta)$ conditioned on containing at least one point.

As a consequence, we show in Proposition 3.5.11 that if the Lévy measure π is diffuse, then $\mathbf{N}^{\psi}$-a.e. there is a unique node X_Δ with degree Δ. Denote by $H_\Delta = H(X_\Delta)$ its height. Then we give a decomposition of the Lévy tree conditioned on $\Delta = \delta$ and $H_\Delta = h$, see Theorem 3.6.3.

Finally, we turn to the behavior of a Lévy tree conditioned to have a large maximal degree. Other conditionings have been considered in the past. Duquesne [56] (this is also related to Williams' decomposition, see [2]) proved that a (sub)critical Lévy tree conditioned on having a large height converges locally to the immortal tree (which consists of an infinite spine onto which trees are grafted according to a Poisson point process). Later, He [86] proved the same convergence for a critical Lévy tree conditioned on having a large maximal degree $\Delta > \delta$ or a large width. In fact, his result is more general as it allows to condition by any measurable function of the tree satisfying a natural monotonicity property.

Here we treat both the critical and the subcritical cases and we consider the density version $\Delta = \delta$. Similarly to the discrete case, two drastically different types of limiting behavior appear. In the subcritical case, there is a condensation phenomenon where a node with infinite degree emerges at a finite exponentially distributed height. Denote by X_Δ the lowest node with degree Δ and let $\mathscr{F}^+_\Delta$ be the forest above X_Δ, seen as a point measure on $\mathbb{R}_+ \times \mathbb{T}$. To be more precise, the forest $\mathscr{F}^+_\Delta = \sum_{i \in I} \delta_{(\ell_i, \mathscr{T}_i)}$ is obtained by decomposing the path of the exploration process (or the height process) into excursions away from 0, with each excursion arriving at local time ℓ_i and coding a tree $\mathscr{T}_i$. Finally, let $\mathscr{T}^-_\Delta$ be the pruned Lévy tree, that is the Lévy tree $\mathscr{T}$ after removing X_Δ and $\mathscr{F}^+_\Delta$. We refer the reader to Theorem 3.7.5 and Theorem 3.8.2 for a precise statement.

Theorem 3.1.6. *Assume that ψ is subcritical and that the Lévy measure π is unbounded. Let $F\colon \mathbb{T}^* \to \mathbb{R}$ be continuous and bounded, $\Phi\colon \mathbb{R}_+ \times \mathbb{T} \to \mathbb{R}_+$ be continuous with bounded support and let A_δ be equal to any one of the following events: $\{\Delta = \delta\}$, $\{\Delta > \delta\}$, $\{\mathscr{T}$ has exactly one node* *with degree larger than* $\delta\}$ *or* $\{\mathscr{T}$ *has exactly one first-generation node with degree larger than* $\delta\}$. *We have:*

$$\lim_{\delta\to\infty} \mathbf{N}^{\psi}\left[F(\mathscr{T}_{\Delta}^{-}, X_{\Delta})\mathrm{e}^{-\langle \mathscr{F}_{\Delta}^{+},\Phi\rangle}\middle| A_{\delta}\right] = \alpha\mathbf{N}^{\psi}\left[\int_{\mathscr{T}} F(\mathscr{T},x)\,\mu(\mathrm{d}x)\right]\exp\left\{-\int_0^{\infty}\mathrm{d}\ell\,\mathbf{N}^{\psi}\left[1-\mathrm{e}^{-\Phi(\ell,\mathscr{T})}\right]\right\}. \tag{3.1.15}$$

In particular, conditionally on A_δ, *the height* $H(X_\Delta)$ *of* X_Δ *converges to an exponential distribution with mean* $1/\alpha$.

The last result should be interpreted as local convergence in distribution to a "condensation tree" described as follows: start with a size-biased Lévy tree $\widetilde{\mathscr{T}}$ with distribution $\alpha\mathbf{N}^{\psi}[\sigma\mathrm{d}\mathscr{T}]$, choose a leaf uniformly at random in $\widetilde{\mathscr{T}}$ and on this leaf graft an independent Lévy forest with infinite degree (i.e. a Poisson point measure on $\mathbb{R}_+\times\mathbb{T}$ with intensity $\mathrm{d}\ell\,\mathbf{N}^{\psi}[\mathrm{d}\mathscr{T}]$). However, the limiting object is a (random) real tree which is not locally compact and the way to circumvent this difficulty is by considering the subtree above the condensation node as a point measure instead.

In the critical case, it should be no surprise that the density version $\Delta=\delta$ gives rise to the same limiting behavior as the tail version $\Delta>\delta$, namely local convergence to the immortal tree. Intuitively, this means that the condensation node goes to infinity and thus becomes invisible to local convergence. Before stating the result, let us define the immortal tree. Let $\sum_{i\in I}\delta_{(s_i,\mathscr{T}_i)}$ be a Poisson point measure on $\mathbb{R}_+\times\mathbb{T}$ with intensity

$$\mathrm{d}s\left(2\beta\mathbf{N}^{\psi}[\mathrm{d}\mathscr{T}]+\int_0^{\infty} r\,\pi(\mathrm{d}r)\,\mathbb{P}_r^{\psi}(\mathrm{d}\mathscr{T})\right).$$

The immortal Lévy tree $\mathscr{T}_\infty^{\psi}$ with branching mechanism ψ is the real tree obtained by grafting the point measure $\sum_{i\in I}\delta_{(s_i,\mathscr{T}_i)}$ on an infinite branch. More formally, set:

$$\mathscr{T}_\infty^{\psi} = \mathbb{R}_+ \circledast_{i\in I}(s_i,\mathscr{T}_i), \tag{3.1.16}$$

where $\mathbb{R}_+$ is considered as a real tree rooted at 0 and equipped with the Euclidean distance and the zero measure. In particular, thanks to [58, Theorem 4.5], we have the following identity which is simply a restatement of Lemma 3.2 in [56] in terms of trees:

$$\mathbb{E}\left[F(r_h(\mathscr{T}_\infty^{\psi}))\right] = \mathrm{e}^{-\alpha h}\mathbf{N}^{\psi}\left[L_\sigma^h F(r_h(\mathscr{T}))\right], \quad \forall h>0. \tag{3.1.17}$$

Theorem 3.1.7. *Assume that* ψ *is critical and that* π *is unbounded. Either let* $A_\delta=\{\Delta=\delta\}$ *and assume that the additional assumption*

$$\lim_{\delta\to\infty}\frac{\pi(\delta)}{\mathbf{N}^{\psi}[\sigma\mathbf{1}_{\{\Delta<\delta\}}]\bar{\pi}(\delta)\int_{[\delta,\infty)} r\,\pi(\mathrm{d}r)}=0 \tag{3.1.18}$$

holds, or let A_δ *be equal to any of the following events:* $\{\Delta>\delta\}$, $\{\mathscr{T}$ *has exactly one node with degree larger than* $\delta\}$ *or* $\{\mathscr{T}$ *has exactly one first-generation node with degree larger than* $\delta\}$. *Then, conditionally on* A_δ, *the Lévy tree* $\mathscr{T}$ *converges in distribution locally to the immortal*

Lévy tree $\mathscr{T}_\infty^\psi$, i.e. we have:

$$\lim_{\delta\to\infty} \mathbf{N}^\psi\left[F(r_h(\mathscr{T}))|A_\delta\right] = \mathbb{E}\left[F(r_h(\mathscr{T}_\infty^\psi))\right]. \tag{3.1.19}$$

We refer to Theorem 3.7.8 and Theorem 3.8.4 for a precise statement. The assumption (3.1.18) is a technical condition which guarantees a fast decay for the size of the atoms of π. Observe that we have $\lim_{\delta\to\infty} \mathbf{N}^\psi[\sigma\mathbf{1}_{\{\Delta<\delta\}}] = \mathbf{N}^\psi[\sigma]$ which is infinite since ψ is critical. Also notice that (3.1.18) is automatically satisfied if the Lévy measure π is diffuse.

It is worth noting that in the critical case, conditioning by the different events A_δ yields the same limiting behavior even though in general they are not equivalent in $\mathbf{N}^\psi$-measure. In the stable (critical) case $\psi(\lambda) = \lambda^\gamma$ with $\gamma \in (1,2)$, these quantities can be computed explicitly, see Proposition 3.9.2. In that case, we also show in Proposition 3.9.4 that, conditionally on $\Delta > \delta$, the distribution of the Bienaymé-Galton-Watson forest of nodes with degree larger than δ is independent of δ.

The rest of the paper is organized as follows. In Section 3.2, we set notation and we introduce the main object we will be dealing with, namely the exploration process. We compute the distribution of the maximal degree in Section 3.3, then we give a Poissonian decomposition of the exploration process along the large nodes and study their structure in Section 3.4. In Section 3.5 (resp. Section 3.6), we make sense of the exploration process conditioned to have a fixed maximal degree (resp. a fixed maximal degree at a given height). Sections 3.7 and 3.8 deal with the local convergence of the exploration process conditioned to have large maximal degree. Finally, Section 3.9 is devoted to the study of the stable case $\psi(\lambda) = \lambda^\gamma$.

3.2 The exploration process and the Lévy tree

In this section, we will recall the construction of the exploration process introduced in [119] and later developped in [57].

3.2.1 Notation

If E is a Polish space, let $\mathscr{B}_+(E)$ be the set of real-valued and nonnegative measurable functions defined on E endowed with its Borel σ-field. For any measure ν on E and any function $f \in \mathscr{B}_+(E)$, we write $\langle \nu, f\rangle = \int_E f(x)\,\nu(\mathrm{d}x)$. We also denote by $\operatorname{supp}(\nu)$ the closed support of the measure ν in E.

We denote by $\mathscr{M}_f(E)$ the set of finite measures on E endowed with the topology of weak convergence. For every $\nu \in \mathscr{M}_f(\mathbb{R}_+)$, we set:

$$H(\nu) = \sup\operatorname{supp}(\nu), \tag{3.2.1}$$

with the convention that $H(0) = 0$. Moreover, we let

$$\Delta(\nu) = \sup\{\nu(x) : x \geq 0\} \tag{3.2.2}$$

be the largest atomic mass of ν. We say that ν is diffuse if it has no atoms and set $\Delta(\nu) = 0$ by convention.

Denote by

$$\mathscr{D} = D(\mathbb{R}_+, \mathscr{M}_f(\mathbb{R}_+)) \tag{3.2.3}$$

the set of $\mathscr{M}_f(\mathbb{R}_+)$-valued càdlàg functions equipped with the Skorokhod J_1-topology. For a function $\mu = (\mu_t, t \geq 0) \in \mathscr{D}$, let

$$\Delta(\mu) = \sup_{t\geq 0} \Delta(\mu_t) \tag{3.2.4}$$

be the largest atom of the entire path of μ.

3.2.2 The underlying Lévy process and the height process

We consider a (sub)critical branching mechanism of the form

$$\psi(\lambda) = \alpha\lambda + \beta\lambda^2 + \int_{(0,\infty)} (\mathrm{e}^{-\lambda r} - 1 + \lambda r)\,\pi(\mathrm{d}r), \quad \forall \lambda \geq 0, \tag{3.2.5}$$

where $\alpha, \beta \geq 0$ and $\pi \neq 0$ is a σ-finite measure on $(0,\infty)$ satisfying $\int_{(0,\infty)} (r \wedge r^2)\,\pi(\mathrm{d}r) < \infty$. We consider a spectrally positive Lévy process $X = (X_t, t \geq 0)$ with Laplace exponent ψ starting from 0. Namely, we have:

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

We assume that X is of infinite variation a.s. which is equivalent to the following condition:

$$\beta > 0 \quad \text{or} \quad \int_{(0,1)} r\,\pi(\mathrm{d}r) = \infty. \tag{3.2.6}$$

Duquesne and Le Gall [57] proved that there exists a process $H = (H_t, t \geq 0)$ called the ψ-height process such that for every $t \geq 0$, we have the following convergence in probability:

$$H_t = \liminf_{\varepsilon\to 0} \frac{1}{\varepsilon} \int_0^t \mathbf{1}_{\{I_t^s < X_s < I_t^s + \varepsilon\}}\,\mathrm{d}s, \tag{3.2.7}$$

where, for every $0 \leq s \leq t$, $I_t^s = \inf_{s\leq r\leq t} X_r$ is the past infimum. They also proved a Ray-Knight theorem for H which shows that the ψ-height process H describes the genealogy of the ψ-CB process, see [57, Theorem 1.4.1].

3.2.3 The exploration process

Although the height process is not Markov in general, it is a simple function of a measure-valued Markov process, the so-called exploration process that we now introduce. The exploration process $\rho = (\rho_t,\ t \geq 0)$ is the $\mathcal{M}_f(\mathbb{R}_+)$-valued process defined as follows:

$$\rho_t(\mathrm{d}r) = \beta \mathbf{1}_{[0,H_t]}(r)\,\mathrm{d}r + \sum_{\substack{0<s\leq t,\\ X_{s-}<I_t^s}} (I_t^s - X_{s-})\delta_{H_s}(\mathrm{d}r). \tag{3.2.8}$$

In particular, the total mass of ρ_t is $\langle \rho_t, 1\rangle = X_t - I_t$.

We will sometimes refer to $t \geq 0$ as a *node* in reference to the corresponding real tree when it is well defined (see Section 3.2.9). For $s, t \geq 0$, we say that s is an ancestor of t and we write $s \preccurlyeq t$ if $s \leq t$ and $H_s = \inf_{s\leq r\leq t} H_r$. The set

$$\{s \geq 0\colon\ s \preccurlyeq t\} \tag{3.2.9}$$

is called the ancestral line of t. We say that $t \geq 0$ is a first-generation node with property $A \subset \mathcal{M}_f(\mathbb{R}_+)$ if $\rho_t \in A$ and $\rho_s \notin A$ for every (strict) ancestor s of t. For example, we will say that t is a first-generation node with mass larger than $\delta > 0$ if $\Delta(\rho_t) > \delta$ and $\Delta(\rho_s) \leq \delta$ for every $s \preccurlyeq t$ with $s \neq t$. Given $0 \leq t_1 \leq \cdots \leq t_n$, there exists a unique $s \geq 0$ such that $r \preccurlyeq t_i$ for every $1 \leq i \leq n$ if and only if $r \preccurlyeq s$. We write $s = t_1 \wedge \cdots \wedge t_n$ and call it the most recent common ancestor (MRCA for short) of $t_1, \ldots, t_n$.

One can recover the heigth process from the exploration process as follows. Denote by $\Delta X_t = X_t - X_{t-}$ the jump of the process X at time t.

Proposition 3.2.1. *Almost surely for every $t > 0$, we have:*

(i) $H(\rho_t) = H_t$,

(ii) $\rho_t = 0$ *if and only if* $H_t = 0$,

(iii) *if* $\rho_t \neq 0$, *then* $\mathrm{supp}(\rho_t) = [0, H_t]$,

(iv) $\rho_t = \rho_{t-} + \Delta X_t \delta_{H_t}$.

In the definition of the exploration process, since X starts from 0, we have $\rho_0 = 0$. In order to state the Markov property of ρ, we have to define the process ρ starting from any initial measure $\nu \in \mathcal{M}_f(\mathbb{R}_+)$. To that end, for every $a \in [0, \langle \nu, 1\rangle]$, we define the erased measure $k_a\nu$ by:

$$k_a\nu[0,r] = \nu[0,r] \wedge (\langle \nu, 1\rangle - a), \quad \forall r \geq 0.$$

If $a > \langle \nu, 1\rangle$, we set $k_a\nu = 0$. In words, the measure $k_a\nu$ is obtained from ν by erasing a mass a backward starting from $H(\nu)$. For $\mu \in \mathcal{M}_f(\mathbb{R}_+)$ with bounded support, we define the concatenation $[\mu,\nu] \in \mathcal{M}_f(\mathbb{R}_+)$ of the measures μ, ν by:

$$\langle [\mu,\nu], f\rangle = \langle \mu, f\rangle + \langle \nu, f(H(\mu)+\cdot)\rangle, \quad \forall f \in \mathcal{B}_+(\mathbb{R}_+).$$

Finally, we set $\rho_0^\nu = \nu$ and

$$\rho_t^\nu = [k_{-I_t}\nu, \rho_t], \quad \forall t > 0.$$

We say that $\rho^\nu = (\rho_t^\nu,\, t \geq 0)$ is the exploration process started at ν and we write $\mathbb{P}_\nu$ for its distribution.

Proposition 3.2.2. *For any $\nu \in \mathcal{M}_f(\mathbb{R}_+)$, the process $\rho^\nu = (\rho_t^\nu,\, t \geq 0)$ is a càdlàg strong Markov process in $\mathcal{M}_f(\mathbb{R}_+)$.*

3.2.4 The excursion measure of the exploration process

Let us introduce the excursion measure $\mathbf{N}^\psi$. Denote by $I = (I_t,\, t \geq 0)$ the infimum process of X:

$$I_t = \inf_{0\leq s\leq t} X_s. \tag{3.2.10}$$

Standard results (see e.g. [29]) entail that $X - I$ is a strong Markov process with values in $\mathbb{R}_+$ and that the point 0 is regular. Furthermore, $-I$ is a local time at 0 for $X - I$. We denote by $\mathbf{N}^\psi$ the associated excursion measure of the process $X - I$ away from 0. It is not difficult to see from (3.2.7) that H_t (and thus also ρ_t) only depends on the excursion of $X - I$ above 0 which straddles time t. It follows that the excursion measure of ρ away from 0 is the "distribution" of ρ under $\mathbf{N}^\psi$. We still denote it by $\mathbf{N}^\psi$ and we let

$$\sigma = \inf\{t > 0 \colon\ \rho_t = 0\} \tag{3.2.11}$$

be the lifetime of ρ under $\mathbf{N}^\psi$ (this coincides with the lifetime of $X - I$ under $\mathbf{N}^\psi$). In particular, the following holds for every $\lambda > 0$:

$$\mathbf{N}^\psi\left[1 - \mathrm{e}^{-\lambda\sigma}\right] = \psi^{-1}(\lambda) \quad \text{and} \quad \mathbf{N}^\psi\left[\sigma \mathrm{e}^{-\lambda\sigma}\right] = \frac{1}{\psi' \circ \psi^{-1}(\lambda)}, \tag{3.2.12}$$

where ψ^{-1} is the inverse function of ψ. By letting $\lambda \to 0$ we obtain:

$$\mathbf{N}^\psi[\sigma] = \frac{1}{\alpha}, \tag{3.2.13}$$

with the convention that $1/0 = \infty$. Let us recall Bismut's decomposition for the exploration process. Let J_a be the random element in $\mathcal{M}_f(\mathbb{R}_+)$ defined by $J_a(\mathrm{d}r) = \mathbf{1}_{[0,a]}(r)\,\mathrm{d}U_r$, where U is a subordinator with Laplace exponent

$$\varphi(\lambda) = \frac{\psi(\lambda)}{\lambda} - \alpha = \beta\lambda + \int_0^\infty \left(1 - \mathrm{e}^{-\lambda r}\right)\bar{\pi}(r)\,\mathrm{d}r, \tag{3.2.14}$$

where the tail $\bar{\pi}$ of the Lévy measure π is defined in (3.1.7).

Proposition 3.2.3. *For every $F \in \mathscr{B}_+(\mathscr{M}_f(\mathbb{R}_+))$, we have:*

$$\mathbf{N}^\psi\left[\int_0^\sigma F(\rho_t)\,\mathrm{d}t\right] = \int_0^\infty \mathrm{d}a\,\mathrm{e}^{-\alpha a}\,\mathbb{E}[F(J_a)]. \tag{3.2.15}$$

3.2.5 Local times of the height process

Although the height process H is not Markov in general, one can show that its local time process exists under $\mathbb{P}$ or $\mathbf{N}^\psi$. More precisely, for every $a > 0$, there exists a continuous nondecreasing process $(L_s^a,\, s \geq 0)$ which can be characterized via the approximation:

$$\lim_{\varepsilon\to 0}\mathbf{N}^\psi\left[\mathbf{1}_{\{\sup H>h\}}\sup_{0\leq s\leq t}\left|\frac{1}{\varepsilon}\int_0^s \mathbf{1}_{\{a-\varepsilon<H_r\leq a\}}\,\mathrm{d}r - L_s^a\right|\right] = 0, \quad \forall\, t, h \geq 0.$$

Moreover, $\mathbf{N}^\psi$-a.e. the support of the measure $L^a(\mathrm{d}s) := \mathrm{d}_s L_s^a$ is contained in $\{s \geq 0\colon\ H_s = a\}$ and we have the occupation time formula $\int_0^\infty \mathrm{d}a\, L^a(\mathrm{d}s) = \mathbf{1}_{[0,\sigma]}(s)\,\mathrm{d}s$. Furthermore, the process $(L_\sigma^a,\, a \geq 0)$ is a ψ-CB process under its canonical measure.

Let us recall the excursion decomposition of the exploration process above level $h > 0$. Set $\tau_s^h = \inf\{t > 0\colon\ \int_0^t \mathbf{1}_{\{H_r\leq h\}}\,\mathrm{d}r > s\}$ and define the truncated exploration process by:

$$r_h(\rho) = (\rho_{\tau_s^h},\, s \geq 0). \tag{3.2.16}$$

Denote by $\mathscr{E}_h$ the σ-field generated by the process $r_h(\rho)$. Let $(\alpha_i, \beta_i), i \in I_h$ denote the excursion intervals of H above level h. For every $i \in I$, we define the measure-valued process ρ^i by:

$$\langle \rho_s^i, f\rangle = \int_{(a,\infty)} f(r-a)\,\rho_{\alpha_i+s}(\mathrm{d}r) \quad \text{if } 0 < s < \beta_i - \alpha_i$$

and $\rho_s^i = 0$ if $s = 0$ or $s \geq \beta_i - \alpha_i$. Finally, let $\ell_i = L_{\alpha_i}^h$ be the local time at level h at the beginning of the excursion ρ^i.

Proposition 3.2.4. *Under $\mathbf{N}^\psi$, conditionally on $\mathscr{E}_h$, the random measure $\sum_{i\in I}\delta_{(\ell_i,\rho^i)}$ is a Poisson point measure with intensity $\mathbf{1}_{[0,L_\sigma^h]}(\ell)\,\mathrm{d}\ell\,\mathbf{N}^\psi[\mathrm{d}\rho]$.*

3.2.6 The dual process

We shall need the $\mathscr{M}_f(\mathbb{R}_+)$-valued process $\eta = (\eta_t,\, t \geq 0)$ defined by:

$$\eta_t(\mathrm{d}r) = \beta\mathbf{1}_{[0,H_t]}(r)\,\mathrm{d}r + \sum_{\substack{0<s\leq t,\\ X_{s-}<I_t^s}} (X_s - I_t^s)\delta_{H_s}(\mathrm{d}r). \tag{3.2.17}$$

The process η is the dual process of ρ under $\mathbf{N}^\psi$ thanks to the following time-reversal property.

Proposition 3.2.5. *The processes $((\rho_t, \eta_t),\, t \geq 0)$ and $((\eta_{(\sigma-s)-}, \rho_{(\sigma-s)-}),\, t \geq 0)$ have the same distribution under $\mathbf{N}^\psi$.*

3.2.7 Grafting procedure

We now explain how to insert a finite collection of measured-valued processes into a measure-valued process. Let $\mu = (\mu(t), 0 \le t < \sigma)$ be a $\mathscr{M}_f(\mathbb{R}_+)$-valued function with lifetime $\sigma \in (0,\infty]$ such that $\mu(t)$ has bounded support for every $t \in [0,\sigma)$ and let $\sum_{i=1}^N \delta_{(s_i,\mu_i)}$ be a finite point measure on $\mathbb{R}_+ \times \mathscr{D}$ where the s_i are arranged in increasing order and each μ_i has a finite lifetime σ^i. Set $s_0 = \Sigma_0 = 0$ and

$$\Sigma_i = \sum_{k=1}^{i} \sigma^k, \quad \forall i \ge 1.$$

Define a measure-valued process $\tilde{\mu}$ by:

$$\tilde{\mu}(t) = \begin{cases} \mu(t - \Sigma_i) & \text{if } s_{i-1} + \Sigma_{i-1} \le t < (s_i \wedge \sigma) + \Sigma_{i-1}, \\ [\mu(s_i), \mu_i(t - s_i - \Sigma_{i-1})] & \text{if } s_i + \Sigma_{i-1} \le t < s_i + \Sigma_i \text{ and } s_i < \sigma. \end{cases}$$

Observe that the (s_i, μ_i) such that $s_i \ge \sigma$ do not play a role in this construction and that $\tilde{\mu}$ has lifetime

$$\sigma + \sum_{i:\, s_i < \sigma} \sigma^k.$$

We denote this grafting procedure by:

$$\mu \circledast_{i=1}^{N} (s_i, \mu_i) = \tilde{\mu}. \tag{3.2.18}$$

In words, this is the process obtained from μ by inserting the measure-valued process μ_i into μ at time $s_i < \sigma$.

3.2.8 A Poissonian decomposition of the exploration process

Let $\nu \in \mathscr{M}_f(\mathbb{R}_+)$. We write $\mathbb{P}_\nu^{\psi,*}$ for the distribution of the exploration process ρ starting at ν and killed when it first reaches 0. Let us introduce two probability measures on $\mathscr{D}$ that will play a major role in the rest of the paper. For every $r > 0$, we will write $\mathbb{P}_r^\psi$ for $\mathbb{P}_{r\delta_0}^{\psi,*}$. This should be interpreted as the distribution of the exploration processes with initial mass r. Furthermore, for every $\delta > 0$ such that $\bar{\pi}(\delta) > 0$, set:

$$\mathbb{Q}_\delta^\psi(\mathrm{d}\rho) = \frac{1}{\bar{\pi}(\delta)} \int_{(\delta,\infty)} \pi(\mathrm{d}r)\, \mathbb{P}_r^\psi(\mathrm{d}\rho), \tag{3.2.19}$$

which is the distribution of the exploration process starting from a random initial mass with distribution π conditioned on being larger than δ.

We decompose the path of ρ under $\mathbb{P}_\nu^{\psi,*}$ according to excursions of the total mass of ρ above its minimum. Let $(\alpha_i, \beta_i), i \in I$ denote the excursion intervals of the process $X - I$ away from 0 under $\mathbb{P}_\nu^{\psi,*}$. Define the measure-valued process ρ^i by $\rho_{(\alpha_i+s)\wedge\beta_i} = [k_{-I_{\alpha_i}}\nu, \rho_s^i]$.

Lemma 3.2.6. *The random measure* $\sum_{i\in I} \delta_{(-I_{\alpha_i}, \rho^i)}$ *is under* $\mathbb{P}_\nu^{\psi,*}$ *a Poisson point measure with* *intensity* $\mathbf{1}_{[0,\langle\nu,1\rangle)}(u)\,\mathrm{d}u\,\mathbf{N}^{\psi}[\mathrm{d}\rho]$. *In particular, under* $\mathbb{P}^{\psi}_{r}$, *it is a Poisson point measure with intensity* $\mathbf{1}_{[0,r]}(u)\,\mathrm{d}u\,\mathbf{N}^{\psi}[\mathrm{d}\rho]$.

Using this decomposition, we can give another useful interpretation of the measure $\mathbb{P}^{\psi}_{r}$. Let ρ be the exploration process starting from 0 and let $(L^0_s, s\ge 0)$ be its local time process at 0. Then the process $(\tilde\rho^{(r)}_t, t\ge 0)$ defined by:

$$\tilde\rho^{(r)}_t = (r-L^0_t)_+\delta_0 + \rho_t\mathbf{1}_{\{L^0_t\le r\}} \tag{3.2.20}$$

has distribution $\mathbb{P}^{\psi}_{r}$.

In the next lemma, we identify the distribution of the exploration process above level $H(\nu)$ starting from ν. For a measure $\nu\in\mathcal{M}_f(\mathbb{R}_+)$ and a positive real $a>0$, define $\theta_a(\nu)$ as the measure ν shifted by a. More formally, define a measure $\theta_a(\nu)$ on $\mathbb{R}_+$ by setting:

$$\langle\theta_a(\nu), f\rangle = \int_{[a,\infty)} f(x-a)\,\nu(\mathrm{d}x),$$

for every $f\in\mathcal{B}_+(\mathbb{R}_+)$ if $a\le H(\nu)$ and $\theta_a(\nu)=0$ if $a>H(\nu)$.

Lemma 3.2.7. *Let* $\nu\in\mathcal{M}_f(\mathbb{R}_+)$ *such that* $\nu(H(\nu))>0$. *Under* $\mathbb{P}^{\psi,*}_{\nu}$, *the process* $\tilde\rho=(\theta_{H(\nu)}(\rho_t), t\ge 0)$ *stopped at the first time it hits* 0 *has distribution* $\mathbb{P}^{\psi}_{\nu(H(\nu))}$.

Proof. We shall use the Poisson decomposition of Lemma 3.2.6. Using its notations, we have $\rho_{(t+\alpha_i)\wedge\beta_i} = [k_{-I_{\alpha_i}}\nu, \rho^i_t]$ where $\sum_{i\in I}\delta_{(-I_{\alpha_i},\rho^i)}$ is a Poisson point measure with intensity $\mathbf{1}_{[0,\langle\nu,1\rangle)}(u)\,\mathrm{d}u\,\mathbf{N}^{\psi}[\mathrm{d}\rho]$. Thus, the exploration process above level $H(\nu)$ stopped at the first time it hits 0 satisfies:

$$\theta_{H(\nu)}(\rho_{(t+\alpha_i)\wedge\beta_i}) = (\nu(H(\nu))+I_{\alpha_i})\delta_0 + \rho^i_t$$

if $-I_{\alpha_i}\le\mu(H(\nu))$ and it is zero if $-I_{\alpha_i}>\nu(H(\nu))$. Applying Lemma 3.2.6 again, it is easy to see that this is also the Poisson decomposition of ρ under $\mathbb{P}^{\psi}_{\nu(H(\nu))}$. This proves the desired result. □

3.2.9 The Lévy tree

Recall that the Grey condition

$$\int^{\infty}\frac{\mathrm{d}\lambda}{\psi(\lambda)}<\infty \tag{3.2.21}$$

is equivalent to the almost sure extinction of the ψ-CB process in finite time. If it holds, then the height process H admits a continuous version and one can use the coding of real trees by continuous excursions (see e.g. [69]) in order to define the Lévy tree $\mathcal{T}$ as the tree coded by the height process H under its excursion measure $\mathbf{N}^{\psi}$. Then the Grey condition implies that $\mathcal{T}$ is a *compact* real tree. In the rest of the paper we forego this assumption, but we still interpret the results in terms of trees as they are easier to understand.

3.3 Distribution of the maximal degree

Under $\mathbf{N}^\psi$, denote by $\Delta = \Delta(\rho)$ the largest atomic mass of the exploration process. Thanks to [58, Theorem 4.6], if $\langle \pi, 1\rangle < \infty$ then the set of discontinuity times of ρ is $\mathbf{N}^\psi$-a.e. finite (and possibly empty). On the other hand, if $\langle \pi, 1\rangle = \infty$ then it is countable and dense in $[0,\sigma]$. In particular, in that case we have that $\mathbf{N}^\psi$-a.e. $\Delta > 0$. The main result of this section is the following proposition giving the joint distribution of the lifetime σ and the maximal degree Δ under $\mathbf{N}^\psi$. Recall from (3.1.7) and (3.1.8) the definitions of $\bar\pi$ and ψ_δ, and define:

$$\psi_{\delta-}(\lambda) = \lim_{\varepsilon\uparrow 0} \psi_{\delta-\varepsilon}(\lambda) = \left(\alpha + \int_{[\delta,\infty)} r\,\pi(\mathrm{d}r)\right)\lambda + \beta\lambda^2 + \int_{(0,\delta)} \left(\mathrm{e}^{-\lambda r} - 1 + \lambda r\right)\pi(\mathrm{d}r). \tag{3.3.1}$$

Proposition 3.3.1. *For every $\delta > 0$ and $\lambda \geq 0$, we have:*

$$\mathbf{N}^\psi\left[1 - \mathrm{e}^{-\lambda\sigma}\mathbf{1}_{\{\Delta\leq\delta\}}\right] = \psi_\delta^{-1}(\bar\pi(\delta) + \lambda), \tag{3.3.2}$$

$$\mathbf{N}^\psi\left[1 - \mathrm{e}^{-\lambda\sigma}\mathbf{1}_{\{\Delta<\delta\}}\right] = \psi_{\delta-}^{-1}(\pi[\delta,\infty) + \lambda). \tag{3.3.3}$$

In particular, we have:

$$\mathbf{N}^\psi[\Delta > \delta] = \psi_\delta^{-1}(\bar\pi(\delta)), \tag{3.3.4}$$

$$\mathbf{N}^\psi[\Delta \geq \delta] = \psi_{\delta-}^{-1}(\pi[\delta,\infty)). \tag{3.3.5}$$

Furthermore, if $\langle\pi,1\rangle < \infty$, then we have:

$$\mathbf{N}^\psi\left[1 - \mathrm{e}^{-\lambda\sigma}\mathbf{1}_{\{\Delta=0\}}\right] = \psi_0^{-1}(\langle\pi,1\rangle + \lambda). \tag{3.3.6}$$

Proof. We only prove (3.3.2), the proof of (3.3.3) being similar. Fix $\delta > 0$ and let $\lambda, \mu \geq 0$. Let

$$A = \{\nu \in \mathcal{M}_f(\mathbb{R}_+) :\ \nu \text{ has an atom with mass} > \delta\}.$$

We shall compute

$$v(\lambda,\mu) = \mathbf{N}^\psi\left[1 - \mathrm{e}^{-\lambda\sigma - \mu\int_0^\sigma \mathrm{d}t\,\mathbf{1}_{\{\rho_t\in A\}}}\right]. \tag{3.3.7}$$

We have:

$$\begin{aligned} v(\lambda,\mu) &= \mathbf{N}^\psi\left[\int_0^\sigma \mathrm{d}t\,(\lambda + \mu\mathbf{1}_{\{\rho_t\in A\}})\mathrm{e}^{-\lambda(\sigma-t) - \mu\int_t^\sigma \mathrm{d}s\,\mathbf{1}_{\{\rho_s\in A\}}}\right] \\ &= \mathbf{N}^\psi\left[\int_0^\sigma \mathrm{d}t\,(\lambda + \mu\mathbf{1}_{\{\rho_t\in A\}})\mathbb{E}^{\psi,*}_{\rho_t}\left[\mathrm{e}^{-\lambda\sigma - \mu\int_0^\sigma \mathrm{d}s\,\mathbf{1}_{\{\rho_s\in A\}}}\right]\right], \end{aligned}$$

where we applied the Markov property for the last equality. We shall use Lemma 3.2.6 to compute the last expectation.

For a measure $\nu \in \mathcal{M}_f(\mathbb{R}_+)$, denote by $H'(\nu)$ the first atom of ν with mass larger than δ:

$$H'(\nu) = \inf\{x \geq 0 :\ \nu(x) > \delta\},$$

with the convention that $\inf \emptyset = +\infty$.

Suppose that $\nu \in A$. Recall from Section 3.2.8 the excursion decomposition of the exploration process above the minimum of its total mass under $\mathbb{P}^{\psi,*}_{\nu}$. Notice that if $-I_{\alpha_i} < \nu\left([H'(\nu), H(\nu)]\right) - \delta$, then $\rho_{(\alpha_i+s)\wedge\beta_i} \in A$ for every $s \geq 0$. On the other hand, if $-I_{\alpha_i} > \nu\left([H'(\nu), H(\nu)]\right) - \delta$, then $\rho_{(\alpha_i+s)\wedge\beta_i} \in A$ if and only if $\rho^i_s \in A$. It follows that

$$
\begin{aligned}
&\mathbb{E}^{\psi,*}_{\nu}\left[\mathrm{e}^{-\lambda\sigma-\mu\int_0^\sigma \mathrm{d}s\,\mathbf{1}_{\{\rho_s\in A\}}}\right] \\
&\qquad = \mathbb{E}^{\psi,*}_{\nu}\left[\exp\left\{-\sum_{i\in I}\left(\lambda\sigma^i + \mu\int_0^{\beta_i-\alpha_i}\mathrm{d}s\,\mathbf{1}_{\{\rho_{\alpha_i+s}\in A\}}\right)\right\}\right] \\
&\qquad = \exp\left\{-\int_0^{\langle\nu,1\rangle}\mathrm{d}u\,\mathbf{N}^{\psi}\left[1-\mathrm{e}^{-(\lambda+\mu\mathbf{1}_{\{u<\nu[H'(\nu),H(\nu)]-\delta\}})\sigma-\mu\mathbf{1}_{\{u>\nu[H'(\nu),H(\nu)]-\delta\}}\int_0^\sigma \mathrm{d}s\,\mathbf{1}_{\{\rho_s\in A\}}}\right]\right\} \\
&\qquad = \exp\left\{-\left(\nu[H'(\nu),H(\nu)]-\delta\right)\psi^{-1}(\lambda+\mu)-\left(\nu[0,H'(\nu))+\delta\right)v(\lambda,\mu)\right\}.
\end{aligned}
$$

Now suppose that $\nu \notin A$. Then $\mathbb{P}^{\psi,*}_{\nu}$-a.s. we have the equality $\{\rho_{(\alpha_i+s)\wedge\beta_i} \in A\} = \{\rho^i_s \in A\}$. It follows that

$$\mathbb{E}^{\psi,*}_{\nu}\left[\mathrm{e}^{-\lambda\sigma-\mu\int_0^\sigma \mathrm{d}s\,\mathbf{1}_{\{\rho_s\in A\}}}\right] = \exp\left\{-\langle\nu,1\rangle\, v(\lambda,\mu)\right\}.$$

We deduce that $v(\lambda,\mu)$ is equal to

$$
\begin{aligned}
(\lambda+\mu)\,\mathbf{N}^{\psi}&\left[\int_0^\sigma \mathrm{d}t\,\mathbf{1}_{\{\rho_t\in A\}}\exp\left\{-\left(\rho_t[H'_t,H_t]-\delta\right)\psi^{-1}(\lambda+\mu)-\left(\rho_t[0,H'_t)+\delta\right)v(\lambda,\mu)\right\}\right] \\
&+\lambda\,\mathbf{N}^{\psi}\left[\int_0^\sigma \mathrm{d}t\,\mathbf{1}_{\{\rho_t\notin A\}}\exp\left\{-\langle\rho_t,1\rangle\, v(\lambda,\mu)\right\}\right], \qquad (3.3.8)
\end{aligned}
$$

where $H'_t = H'(\rho_t)$.

Thanks to Proposition 3.2.3, for every $\theta, \omega \geq 0$ we have:

$$\mathbf{N}^{\psi}\left[\int_0^\sigma \mathrm{d}t\,\mathbf{1}_{\{\rho_t\in A\}}\exp\left\{-\theta\rho_t[0,H'_t)-\omega\rho_t[H'_t,H_t]\right\}\right] = \int_0^\infty \mathrm{d}a\,\mathrm{e}^{-\alpha a} f(a,\theta,\omega), \qquad (3.3.9)$$

where we set

$$f(a,\theta,\omega) := \mathbb{E}\left[\mathbf{1}_{\{J_a\in A\}}\mathrm{e}^{-\theta J_a[0,H'(J_a))-\omega J_a[H'(J_a),H(J_a)]}\right].$$

Recall that $J_a(\mathrm{d}r) = \mathbf{1}_{[0,a]}(r)\,\mathrm{d}U_r$ where U is a subordinator with Laplace exponent φ defined in (3.2.14). Denote by T be the time of the first jump of U exceeding δ:

$$T := \inf\{r > 0 \colon \Delta U_r > \delta\} \qquad (3.3.10)$$

Then it is clear that $H(J_a) = a$, $H'(J_a) = T$ and $\{J_a \in A\} = \{T \leq a\}$. Thus, we get:

$$
\begin{aligned}
f(a,\theta,\omega) &= \mathbb{E}\left[\mathbf{1}_{\{T\leq a\}}\mathrm{e}^{-\theta U_{T-}-\omega\Delta U_T-\omega(U_a-U_T)}\right] \\
&= \mathbb{E}\left[\mathbf{1}_{\{T\leq a\}}\mathrm{e}^{-\theta U_{T-}-\omega\Delta U_T-\varphi(\omega)(a-T)}\right],
\end{aligned}
$$

where we used the strong Markov property at time T for the last equality.

Set
$$s_\delta = \int_\delta^\infty \bar{\pi}(r)\,\mathrm{d}r \quad \text{and} \quad \varphi_\delta(\lambda) = \beta\lambda + \int_0^\delta (1-\mathrm{e}^{-\lambda r})\bar{\pi}(r)\,\mathrm{d}r. \tag{3.3.11}$$
Using basic results on Poisson point processes, we have that T is exponentially distributed with mean $1/s_\delta$, ΔU_T has distribution $s_\delta^{-1}\mathbf{1}_{[\delta,\infty)}(x)\bar{\pi}(x)\,\mathrm{d}x$ and is independent of T, and the process $(U_r, 0 \le r < T)$ is distributed as $(V_r, 0 \le r < T)$, where V is a subordinator with Laplace exponent φ_δ, independent of $(T, \Delta U_T)$. Therefore, it follows that
$$f(a,\theta,\omega) = \int_0^a \mathrm{d}t\,\mathrm{e}^{-s_\delta t - \varphi_\delta(\theta)t - \varphi(\omega)(a-t)} \int_\delta^\infty \mathrm{d}x\,\bar{\pi}(x)\mathrm{e}^{-\omega x}.$$
We deduce from (3.3.9) that
$$\mathbf{N}^\psi\left[\int_0^\sigma \mathrm{d}t\,\mathbf{1}_{\{\rho_t\in A\}}\exp\left\{-\theta\rho_t([0,H'_t)) - \omega\rho_t([H'_t,H_t])\right\}\right] = \frac{1}{(\alpha+\varphi(\omega))(s_\delta+\alpha+\varphi_\delta(\theta))}\int_\delta^\infty \mathrm{d}x\,\bar{\pi}(x)\mathrm{e}^{-\omega x}. \tag{3.3.12}$$
Similar arguments yield
$$\begin{aligned}\mathbf{N}^\psi\left[\int_0^\sigma \mathrm{d}t\,\mathbf{1}_{\{\rho_t\notin A\}}\exp\left\{-\theta\langle\rho_t,1\rangle\right\}\right] &= \int_0^\infty \mathrm{d}a\,\mathrm{e}^{-\alpha a}\,\mathbb{E}\left[\mathbf{1}_{\{J_a\notin A\}}\mathrm{e}^{-\theta\langle J_a,1\rangle}\right]\\ &= \int_0^\infty \mathrm{d}a\,\mathrm{e}^{-\alpha a}\,\mathbb{E}\left[\mathbf{1}_{\{T>a\}}\mathrm{e}^{-\theta U_a}\right]\\ &= \frac{1}{s_\delta+\alpha+\varphi_\delta(\theta)}.\end{aligned} \tag{3.3.13}$$

It follows from (3.3.8), (3.3.12) and (3.3.13) that
$$v(\lambda,\mu) = \frac{(\lambda+\mu)\mathrm{e}^{\delta(\psi^{-1}(\lambda+\mu)-v(\lambda,\mu))}}{(\alpha+\varphi\circ\psi^{-1}(\lambda+\mu))(s_\delta+\alpha+\varphi_\delta\circ v(\lambda,\mu))}\int_\delta^\infty \mathrm{d}x\,\bar{\pi}(x)\mathrm{e}^{-\psi^{-1}(\lambda+\mu)x} + \frac{\lambda}{s_\delta+\alpha+\varphi_\delta\circ v(\lambda,\mu)}. \tag{3.3.14}$$
From (3.3.7), it is clear by monotone convergence that $v(\lambda,\mu)\uparrow v(\lambda)$ as $\mu\uparrow\infty$, where
$$v(\lambda) := \mathbf{N}^\psi\left[1-\mathrm{e}^{-\lambda\sigma}\mathbf{1}_{\{\Delta\le\delta\}}\right].$$
Furthermore, thanks to a Tauberian theorem, we have as $\mu\to\infty$:
$$\int_\delta^\infty \mathrm{e}^{-\psi^{-1}(\lambda+\mu)x}\bar{\pi}(x)\,\mathrm{d}x \sim \frac{\bar{\pi}(\delta)\mathrm{e}^{-\delta\psi^{-1}(\lambda+\mu)}}{\psi^{-1}(\lambda+\mu)}. \tag{3.3.15}$$

Thus, letting $\mu \to \infty$ in (3.3.14) and using that $\psi^{-1}(x)\left(\alpha + \varphi \circ \psi^{-1}(x)\right) = x$ for every $x > 0$, we get:

$$v(\lambda) = \frac{\bar{\pi}(\delta)\mathrm{e}^{-\delta v(\lambda)} + \lambda}{s_\delta + \alpha + \varphi_\delta \circ v(\lambda)}. \tag{3.3.16}$$

Notice that for every $x > 0$, we have:

$$\begin{aligned} s_\delta + \alpha + \varphi_\delta(x) &= \alpha + \varphi(x) + \int_\delta^\infty \mathrm{e}^{-xr}\bar{\pi}(r)\,\mathrm{d}r \\ &= \frac{1}{x}\left(\psi(x) + \int_{(\delta,\infty)} \left(1-\mathrm{e}^{-xr}\right)\pi(\mathrm{d}r) - \bar{\pi}(\delta)\left(1-\mathrm{e}^{-x\delta}\right)\right) \\ &= \frac{1}{x}\left(\psi_\delta(x) - \bar{\pi}(\delta)\left(1-\mathrm{e}^{-x\delta}\right)\right), \end{aligned}$$

where we used (3.2.14) and Fubini's theorem for the second equality and the definition of ψ_δ for the last. Thus (3.3.16) becomes

$$\psi_\delta \circ v(\lambda) = \bar{\pi}(\delta) + \lambda.$$

This yields (3.3.2). Then (3.3.6) follows by letting $\delta \to 0$. □

As a consequence, the following corollary states that the distribution of Δ under $\mathbf{N}^\psi$ on $(0,\infty)$ and the Lévy measure π have the same support and the same atoms.

Corollary 3.3.2. *The measures $\mathbf{N}^\psi[\Delta \in \cdot]_{|(0,\infty)}$ and π have the same support. Furthermore, for every $\delta > 0$, $\mathbf{N}^\psi[\Delta = \delta] > 0$ if and only if δ is an atom of the Lévy measure π.*

Proof. This is clear from (3.3.4) and (3.3.5). □

Remark 3.3.3. More precisely, if $\delta > 0$ is an atom of π, we have:

$$\mathbf{N}^\psi[\Delta = \delta] = \psi_{\delta-}^{-1}(\pi[\delta,\infty)) - \psi_\delta^{-1}(\bar{\pi}(\delta)). \tag{3.3.17}$$

Furthermore, if $\langle \pi, 1\rangle < \infty$, then we have:

$$\mathbf{N}^\psi[\Delta > 0] = \psi_0^{-1}(\langle \pi, 1\rangle) > 0. \tag{3.3.18}$$

3.4 Degree decomposition of the Lévy tree

In this section, we give a decomposition of the Lévy tree along the large nodes. More precisely, we identify the distribution of the pruned Lévy tree obtained by removing large nodes. Furthermore, we show that the initial Lévy tree can be recovered in distribution from the pruned one by grafting Lévy forests in a Poissonian manner. We apply this decomposition to describe the structure of the discrete tree formed by large nodes.

3.4.1 A Poissonian decomposition of the Lévy tree

The main result of this section is the following Poissonian decomposition along the nodes with mass larger than δ. Recall from (3.2.18) the definition of the grafting procedure $\circledast$.

Theorem 3.4.1. *The following holds:*

(i) Let $\delta \geq 0$ such that $\bar{\pi}(\delta) < \infty$. Under $\mathbf{N}^{\psi_\delta}$, let $\big((s_i, \rho_i),\, i \in I\big)$ be the atoms of a Poisson point measure with intensity $\bar{\pi}(\delta)\,\mathrm{d}s\,\mathbb{Q}_\delta^\psi(\mathrm{d}\tilde{\rho})$, independent of ρ. Then, under $\mathbf{N}^{\psi_\delta}$, the process $\rho \circledast_{i\in I} (s_i, \rho_i)$ has distribution $\mathbf{N}^\psi$.

(ii) Let $\delta > 0$. Under $\mathbf{N}^{\psi_{\delta-}}$, let $\big((s_i, \rho_i),\, i \in I\big)$ be the atoms of a Poisson point measure with intensity

$$\mathrm{d}s \int_{[\delta,\infty)} \pi(\mathrm{d}r)\,\mathbb{P}_r^\psi(\mathrm{d}\tilde{\rho}).$$

Then, under $\mathbf{N}^{\psi_{\delta-}}$, the process $\rho \circledast_{i\in I} (s_i, \rho_i)$ has distribution $\mathbf{N}^\psi$.

Remark 3.4.2. As mentioned in the introduction, the above theorem is a special case of the main result in [8] where the number of marks is finite. This greatly simplifies the proof which is why we choose include it. Observe however that the decomposition in [8] is proved under $\mathbb{P}$ and that an additional argument is needed to show that it still holds under the excursion measures, see the end of the proof below.

Proof. We only prove the first part, the second one being similar. Notice that the statement is trivial if $\bar{\pi}(\delta) = 0$ since in that case we have $\psi_\delta = \psi$ and the intensity of the Poisson point measure is 0. Thus we may assume that $\bar{\pi}(\delta) \in (0,\infty)$. We shall start by proving the identity under $\mathbb{P}$ using a coupling argument. Let $X^\delta = (X_t^\delta,\, t \geq 0)$ be a Lévy process with Laplace exponent ψ_δ and let $e = (e_t,\, t \geq 0)$ be an independent Poisson point process on $\mathbb{R}_+$ with intensity $\mathbf{1}_{\{r>\delta\}}\,\pi(\mathrm{d}r)$. Define the process $X = (X_t,\, t \geq 0)$ by:

$$X_t = X_t^\delta + \sum_{s\leq t} e_s, \quad \forall\, t \geq 0.$$

Then the process X is also a Lévy process with Laplace transform $\psi_\delta(\lambda) + \int_{(\delta,\infty)} (\mathrm{e}^{-\lambda r} - 1)\,\pi(\mathrm{d}r) = \psi(\lambda)$. In words, the process X^δ is obtained from X by removing jumps of size larger than δ.

Denote by ρ (resp. ρ^δ) the exploration process associated with X (resp. X^δ). Let $T_\delta := \inf\{t > 0 \colon \Delta(\rho_t) > \delta\}$ be the first time ρ contains an atom with mass larger than δ. It is clear from the definition that the process ρ jumps exactly when X does, so that $T_\delta = \inf\{t > 0 \colon \Delta X_t > \delta\}$. Therefore, we have that $X_t = X_t^\delta$ for $t < T_\delta$, which implies that $\rho_t = \rho_t^\delta$ for $t < T_\delta$.

Now, from the construction of X, we get that $T_\delta = \inf\{t > 0 \colon e_t > \delta\}$, that is T_δ is the first time that the Poisson point process e enters in (δ,∞). Therefore the random time T_δ is exponentially distributed with mean $1/\bar{\pi}(\delta)$ and the jump $\Delta X_{T_\delta} = e_{T_\delta}$ has distribution $\mathbf{1}_{\{r>\delta\}}\,\pi(\mathrm{d}r)/\bar{\pi}(\delta)$ and is independent of T_δ. Furthermore, the pair $(T_\delta, \Delta X_{T_\delta})$ is independent of X^δ.

Recall from (3.2.9) the definition of the ancestral line of $t \in [0, \sigma]$. Let $\Delta_t = \sup_{s \preccurlyeq t} \Delta X_s = \sup_{s \preccurlyeq t} \Delta(\rho_s)$ be the maximal degree of the ancestral line of t. For every $t \geq 0$, let

$$A(t) := \int_0^t \mathbf{1}_{\{\Delta_s \leq \delta\}} \, \mathrm{d}s \tag{3.4.1}$$

be the Lebesgue measure of the set of individuals prior to t whose lineage does not contain any node with mass larger than δ. Let $C_t := \inf\{s \geq 0 \colon A_s > t\}$ be the right-continuous inverse of A and define the pruned exploration process $\tilde{\rho} = (\tilde{\rho}_t = \rho_{C_t}, t \geq 0)$. In other words, we remove from the tree all the individuals above a node with mass larger than δ and the pruned exploration process $\tilde{\rho}$ codes the remaining tree.

Next, let us consider excursions of ρ above nodes of mass larger than δ. Let $T_\delta^{(1)} = T_\delta$ be the first time ρ contains an atom with mass larger than δ and $L_\delta^{(1)} = L_\delta = \inf\{t > T_\delta \colon H_t < H_{T_\delta}\}$ be the first time that atom is erased. Define recursively the stopping times $T_\delta^{(k)} = \inf\{t > L_\delta^{(k-1)} \colon \Delta(\rho_t) > \delta\}$ the k-th time ρ contains a (first-generation) node with mass larger than δ and $L_\delta^{(k)} = \inf\{t > T_\delta^{(k)} \colon H_t < H_{T_\delta^{(k)}}\}$ the first time that node is erased. Finally, let $\rho^{(k)}$ be the path of the exploration process above level $H_{T_\delta^{(k)}}$ between times $T_\delta^{(k)}$ and $L_\delta^{(k)}$, defined by:

$$\rho_t^{(k)} = \theta_{H_{T_\delta^{(k)}}}(\rho_{t+T_\delta^{(k)}}), \quad \forall 0 \leq t \leq L_\delta^{(k)} - T_\delta^{(k)}.$$

Notice that by construction, we have:

$$\rho = \tilde{\rho} \circledast_{k=1}^{\infty} (A(T_\delta^{(k)}), \rho^{(k)}).$$

Using the strong Markov property under $\mathbb{P}$ at time T_δ and Lemma 3.2.7, we get that, conditionally on $\Delta(\rho_{T_\delta})$ (which is equal to ΔX_{T_δ}), the process $\rho^{(1)}$ has distribution $\mathbb{P}^{\psi}_{\Delta X_{T_\delta}}$.

But the random time T_δ is exponentially distributed with mean $1/\bar{\pi}(\delta)$, the jump ΔX_{T_δ} has distribution $\mathbf{1}_{(\delta,\infty)}(r)\,\pi(\mathrm{d}r)/\bar{\pi}(\delta)$ and they are independent. We deduce that $\rho^{(1)}$ is independent of T_δ and has distribution $\mathbb{Q}^{\psi}_{\delta}$. Furthermore, $(T_\delta, \Delta X_{T_\delta})$ is generated by the Poisson point process e while $\tilde{\rho}$ is generated by X^δ. These being independent, we deduce that $\tilde{\rho}$ is independent of $(T_\delta, \Delta X_{T_\delta})$, and thus of $(A(T_\delta^{(1)}) = T_\delta, \rho^{(1)})$. Iterating this argument and using the strong Markov property, we get that the random measure

$$\sum_{k=1}^{\infty} \delta_{(A(T_\delta^{(k)}), \rho^{(k)})}$$

is a Poisson point measure with intensity $\bar{\pi}(\delta)\,\mathrm{d}s\,\mathbb{Q}^{\psi}_{\delta}(\mathrm{d}\rho)$ and is independent of $\tilde{\rho}$.

It remains to show that $\tilde{\rho}$ is distributed as ρ^δ. Recall that $\tilde{\rho}_t = \rho_{C_t}$. From this, it is clear that the two processes are equal to ρ before time T_δ. Furthermore, at time T_δ we have $\tilde{\rho}_{T_\delta} = \rho_{T_\delta -} = \rho_{L_\delta} = \rho^{\delta}_{T_\delta}$. Now applying the strong Markov property to ρ at L_δ gives that, conditionally on $\tilde{\rho}_{T_\delta}$, the process $(\rho_{t+L_\delta}, t \geq 0)$ has distribution $\mathbb{P}_{\tilde{\rho}_{T_\delta}}$. As a consequence, conditionally on $\tilde{\rho}_{T_\delta}$, the process $(\tilde{\rho}_t = \rho_{t+L_\delta - T_\delta},\ A(T_\delta^{(1)}) \le t < A(T_\delta^{(2)}))$ is distributed as $(\rho_t^\delta,\ S_1 \le t < S_2)$, where $0 \le S_1 \le S_2 \le \ldots$ are the ordered atoms of a Poisson point process on $\mathbb{R}_+$ with intensity $\bar{\pi}(\delta)\,\mathrm{d}s$, independent of ρ^δ. Iterating this argument, we deduce that $\tilde{\rho}$ and ρ^δ have the same distribution. This proves the Poisson decomposition under $\mathbb{P}$. Therefore, the same decomposition holds under the excursion measures up to a normalizing constant: there exists a constant $c > 0$ such that, under $\mathbf{N}^{\psi_\delta}$, the process $\rho \circledast_{i\in I} (s_i, \rho_i)$ has distribution $c\mathbf{N}^\psi$, where the random measure $\sum_{i\in I} \delta_{(s_i,\rho_i)}$ is under $\mathbf{N}^{\psi_\delta}$ a Poisson point measure with intensity $\bar{\pi}(\delta)\,\mathrm{d}s\,\mathbb{Q}_\delta^\psi(\mathrm{d}\tilde{\rho})$. Let $\zeta = \mathrm{Card}\{i \in I\colon\ s_i < \sigma\}$. Then, under $\mathbf{N}^{\psi_\delta}$ and conditionally on ρ, the random variable ζ has Poisson distribution with parameter $\bar{\pi}(\delta)\sigma$. It follows that

$$\mathbf{N}^{\psi_\delta}[\zeta \ge 1] = \mathbf{N}^{\psi_\delta}\left[\mathbf{N}^{\psi_\delta}\left[\zeta \ge 1 \middle| \rho\right]\right] = \mathbf{N}^{\psi_\delta}\left[1 - \mathrm{e}^{-\bar{\pi}(\delta)\sigma}\right] = \psi_\delta^{-1}(\bar{\pi}(\delta)) = \mathbf{N}^\psi[\Delta > \delta],$$

where in the last equality we used Proposition 3.3.1. This gives $c = 1$ and the result readily follows. □

The following corollary is an immediate consequence of the Poissonian decomposition from Theorem 3.4.1.

Corollary 3.4.3. *Let $\delta > 0$ and let $F \in \mathscr{B}_+(\mathscr{D})$. We have:*

$$\mathbf{N}^\psi\left[F(\rho)\mathbf{1}_{\{\Delta\le\delta\}}\right] = \mathbf{N}^{\psi_\delta}\left[F(\rho)\mathrm{e}^{-\bar{\pi}(\delta)\sigma}\right], \tag{3.4.2}$$

$$\mathbf{N}^\psi\left[F(\rho)\mathbf{1}_{\{\Delta<\delta\}}\right] = \mathbf{N}^{\psi_{\delta-}}\left[F(\rho)\mathrm{e}^{-\pi[\delta,\infty)\sigma}\right] \tag{3.4.3}$$

Furthermore, if $\langle\pi, 1\rangle < \infty$, then we have:

$$\mathbf{N}^\psi\left[F(\rho)\mathbf{1}_{\{\Delta=0\}}\right] = \mathbf{N}^{\psi_0}\left[F(\rho)\mathrm{e}^{-\langle\pi,1\rangle\sigma}\right]. \tag{3.4.4}$$

The Poissonian decomposition of Theorem 3.4.1 also holds for forests.

Proposition 3.4.4. *Let $\delta > 0$ such that $\bar{\pi}(\delta) < \infty$ and let $r > 0$. Under $\mathbb{P}_r^{\psi_\delta}$ (resp. $\mathbb{Q}_\delta^{\psi_\delta}$), let $\big((s_i, \rho_i),\ i \in I\big)$ be the atoms of a Poisson point measure with intensity $\bar{\pi}(\delta)\,\mathrm{d}s\,\mathbb{Q}_\delta^\psi(\mathrm{d}\tilde{\rho})$. Then, under $\mathbb{P}_r^{\psi_\delta}$ (resp. $\mathbb{Q}_\delta^{\psi_\delta}$), the process $\rho \circledast_{i\in I} (s_i, \rho_i)$ has distribution $\mathbb{P}_r^\psi$ (resp. $\mathbb{Q}_\delta^\psi$).*

3.4.2 Structure of nodes with mass larger than δ

Here, we give a description of the structure of nodes with mass larger than δ under $\mathbf{N}^\psi$. Let us start by determining the distribution of the height of MRCA (see Section 3.2.3 for the definition) of the set of nodes with mass larger than δ.

Proposition 3.4.5. *Under $\mathbf{N}^\psi$, conditionally on $\Delta > \delta$, the height of the MRCA of the set of nodes with mass larger than δ is exponentially distributed with mean $1/\psi_\delta'(\mathbf{N}^\psi[\Delta > \delta])$.*

Notice that, as $\delta \to \infty$, $\psi_\delta'(\mathbf{N}^\psi[\Delta > \delta])$ converges to α which is positive in the subcritical case and 0 in the critical case (this implies that the height of the MRCA goes to infinity).

Proof. Under $\mathbf{N}^{\psi_\delta}$, denote by $\tau_1 \le \tau_2 \le \ldots$ the jump times of a standard Poisson process with intensity $\bar\pi(\delta)$. Denote by $M = \sup\{i \ge 1 :\ \tau_i \le \sigma\}$ the number of marks which arrive during the lifetime σ and set:

$$J = \begin{cases} \inf\{H_s :\ \tau_1 \le s \le \tau_M\} & \text{if } M \ge 1, \\ \infty & \text{if } M = 0. \end{cases}$$

It is clear from Theorem 3.4.1 that, under $\mathbf{N}^{\psi}$, the height of the MRCA of the set of nodes with mass larger than δ is distributed as J under $\mathbf{N}^{\psi_\delta}$, with the convention that this height is equal to ∞ if there are no such nodes. Thus, we need to determine the distribution of J under $\mathbf{N}^{\psi_\delta}$ and conditionally on $M \ge 1$.

Notice that, on the event $\{M \ge 2\}$, J agrees with the random variable K defined in [57, p.96]. Proposition 3.2.3 therein gives:

$$\mathbf{N}^{\psi_\delta}\left[f(J)\mathbf{1}_{\{M\ge 2\}}\middle| M \ge 1\right] = \left(\psi'_\delta(\mathbf{N}^{\psi}[\Delta > \delta]) - \frac{\bar\pi(\delta)}{\mathbf{N}^{\psi}[\Delta > \delta]}\right)\int_0^\infty f(a)\mathrm{e}^{-a\psi'_\delta(\mathbf{N}^{\psi}[\Delta>\delta])}\,\mathrm{d}a, \qquad (3.4.5)$$

where we used that $\psi_\delta(\mathbf{N}^{\psi}[\Delta > \delta]) = \bar\pi(\delta)$ by (3.3.4).

Next, notice that under $\mathbf{N}^{\psi_\delta}$, conditionally on ρ, M has Poisson distribution with parameter $\bar\pi(\delta)\sigma$. Furthermore, conditionally on ρ and on $M = 1$, τ_1 is uniformly distributed on $[0,\sigma]$. Thus, by conditioning on ρ, we get:

$$\begin{aligned} \mathbf{N}^{\psi_\delta}\left[f(J)\mathbf{1}_{\{M=1\}}\right] &= \bar\pi(\delta)\,\mathbf{N}^{\psi_\delta}\left[\int_0^\sigma f(H_t)\mathrm{e}^{-\bar\pi(\delta)\sigma}\,\mathrm{d}t\right] \\ &= \bar\pi(\delta)\,\mathbf{N}^{\psi_\delta}\left[\int_0^\sigma f(H_t)\mathrm{e}^{-\bar\pi(\delta)t}\mathrm{e}^{-\bar\pi(\delta)(\sigma-t)}\,\mathrm{d}t\right] \\ &= \bar\pi(\delta)\,\mathbf{N}^{\psi_\delta}\left[\int_0^\sigma f(H_t)\mathrm{e}^{-\bar\pi(\delta)t}\,\mathbb{E}^{\psi_\delta,*}_{\rho_t}\left[\mathrm{e}^{-\bar\pi(\delta)\sigma}\right]\mathrm{d}t\right], \end{aligned}$$

where we used the Markov property of the exploration process under $\mathbf{N}^{\psi_\delta}$ for the last equality. Thanks to Lemma 3.2.6, for every $\nu \in \mathcal{M}_f(\mathbb{R}_+)$ we have:

$$\mathbb{E}^{\psi_\delta,*}_{\nu}\left[\mathrm{e}^{-\bar\pi(\delta)\sigma}\right] = \mathrm{e}^{-\psi_\delta^{-1}(\bar\pi(\delta))\langle\nu,1\rangle} = \mathrm{e}^{-\mathbf{N}^{\psi}[\Delta>\delta]\langle\nu,1\rangle},$$

where we used (3.3.4) for the last equality.

Therefore, we get:

$$\begin{aligned} \mathbf{N}^{\psi_\delta}\left[f(J)\mathbf{1}_{\{M=1\}}\right] &= \bar\pi(\delta)\,\mathbf{N}^{\psi_\delta}\left[\int_0^\sigma f(H_t)\mathrm{e}^{-\bar\pi(\delta)t}\mathrm{e}^{-\langle\rho_t,1\rangle\mathbf{N}^{\psi}[\Delta>\delta]}\,\mathrm{d}t\right] \\ &= \bar\pi(\delta)\,\mathbf{N}^{\psi_\delta}\left[\int_0^\sigma f(H_t)\mathrm{e}^{-\bar\pi(\delta)(\sigma-t)}\mathrm{e}^{-\langle\eta_t,1\rangle\mathbf{N}^{\psi}[\Delta>\delta]}\,\mathrm{d}t\right] \\ &= \bar\pi(\delta)\,\mathbf{N}^{\psi_\delta}\left[\int_0^\sigma f(H_t)\mathrm{e}^{-\langle\rho_t+\eta_t,1\rangle\mathbf{N}^{\psi}[\Delta>\delta]}\,\mathrm{d}t\right], \end{aligned}$$

where we used the time-reversal property of the exploration process for the second equality and the Markov property for the last. By [57, Proposition 3.1.3], we deduce that

$$\mathbf{N}^{\psi_\delta}\left[f(J)\mathbf{1}_{\{M=1\}}\right] = \bar{\pi}(\delta)\int_0^\infty f(a)\mathrm{e}^{-\psi_\delta'(\mathbf{N}^\psi[\Delta>\delta])a}\,\mathrm{d}a.$$

Thanks to Theorem 3.4.1, it is clear that $\mathbf{N}^{\psi_\delta}[M\geq 1] = \mathbf{N}^\psi[\Delta>\delta]$. It follows that

$$\mathbf{N}^{\psi_\delta}\left[f(J)\mathbf{1}_{\{M=1\}}\middle| M\geq 1\right] = \frac{\bar{\pi}(\delta)}{\mathbf{N}^\psi[\Delta>\delta]}\int_0^\infty f(a)\mathrm{e}^{-\psi_\delta'(\mathbf{N}^\psi[\Delta>\delta])a}\,\mathrm{d}a.$$

In conjunction with (3.4.5), this yields:

$$\mathbf{N}^{\psi_\delta}\left[f(J)\middle| M\geq 1\right] = \psi_\delta'(\mathbf{N}^\psi[\Delta>\delta])\int_0^\infty f(a)\mathrm{e}^{-\psi_\delta'(\mathbf{N}^\psi[\Delta>\delta])a}\,\mathrm{d}a.$$

This shows that, under $\mathbf{N}^{\psi_\delta}$ and conditionally on $M\geq 1$, J is exponentially distributed with mean $1/\psi_\delta'(\mathbf{N}^\psi[\Delta>\delta])$ and the proof is now complete. □

Let $\mathbf{t}_\delta$ be the (random) discrete forest spanned by nodes with mass larger than δ. More explicitly, $\mathbf{t}_\delta$ starts with Z_0^δ individuals, where Z_0^δ is the number of first-generation nodes of ρ with mass larger than δ (that is nodes of ρ with mass larger than δ having no ancestors with mass larger than δ). Then, each node v of $\mathbf{t}_\delta$ gets ξ_v^δ children, where ξ_v^δ is the number of first-generation descendants with mass larger than δ of the corresponding node in ρ. Finally, denote by W^δ the total population of $\mathbf{t}_\delta$ or equivalently the total number of nodes of ρ with mass larger than δ. We shall identify the distribution of $\mathbf{t}_\delta$. Given two $\mathbb{N}$-valued random variables Z_0 and ξ, we call a (Z_0,ξ)-Bienaymé-Galton-Watson forest a collection of Z_0 independent Bienaymé-Galton-Watson trees with offspring distribution (the law of) ξ.

Under $\mathbf{N}^{\psi_\delta}$ (resp. under $\mathbb{Q}_\delta^{\psi_\delta}$), let $\sum_{i\in I}\delta_{(s_i,\rho^i)}$ be a Poisson point measure with intensity $\bar{\pi}(\delta)\,\mathrm{d}s\ \mathbb{Q}_\delta^\psi(\mathrm{d}\tilde{\rho})$ independent of ρ and let

$$\zeta = \mathrm{Card}\{i\in I:\ s_i<\sigma\} \tag{3.4.6}$$

be the number of points arriving during the lifetime σ. Basic properties of Poisson point measures imply that, under $\mathbf{N}^{\psi_\delta}$ (resp. under $\mathbb{Q}_\delta^{\psi_\delta}$) and conditionally on ρ, the random variable ζ has Poisson distribution with parameter $\bar{\pi}(\delta)\sigma$.

Proposition 3.4.6. *Let $\delta>0$ such that $\bar{\pi}(\delta)>0$. Under $\mathbf{N}^\psi$, the random forest $\mathbf{t}_\delta$ is a (Z_0^δ,ξ^δ)-Bienaymé-Galton-Watson forest, where Z_0^δ is distributed as ζ under $\mathbf{N}^{\psi_\delta}$ and ξ^δ is distributed as ζ under $\mathbb{Q}_\delta^{\psi_\delta}$. Their Laplace transforms are given by, for every $\lambda>0$:*

$$\mathbf{N}^\psi\left[1-\mathrm{e}^{-\lambda Z_0^\delta}\right] = \psi_\delta^{-1}\left((1-\mathrm{e}^{-\lambda})\bar{\pi}(\delta)\right), \tag{3.4.7}$$

$$\mathbf{N}^\psi\left[\mathrm{e}^{-\lambda\xi^\delta}\right] = \frac{1}{\bar{\pi}(\delta)}\int_{(\delta,\infty)}\mathrm{e}^{-r\psi_\delta^{-1}\left((1-\mathrm{e}^{-\lambda})\bar{\pi}(\delta)\right)}\,\pi(\mathrm{d}r). \tag{3.4.8}$$

Proof. That $\mathbf{t}_\delta$ is under $\mathbf{N}^\psi$ a Bienaymé-Galton-Watson forest with the mentioned distribution is an immediate consequence of the Poissonian decompositions given in Thereom 3.4.1 and Proposition 3.4.4. Let us compute the Laplace transforms.

Recall that, under $\mathbf{N}^{\psi_\delta}$ and conditionally on ρ, ζ has Poisson distribution with parameter $\bar{\pi}(\delta)\sigma$. Using this, we have:

$$\mathbf{N}^\psi\left[1-\mathrm{e}^{-\lambda Z_0^\delta}\right]=\mathbf{N}^{\psi_\delta}\left[1-\mathrm{e}^{-\lambda\zeta}\right]=\mathbf{N}^{\psi_\delta}\left[1-\mathrm{e}^{-(1-\mathrm{e}^{-\lambda})\bar{\pi}(\delta)\sigma}\right]=\psi_\delta^{-1}\left((1-\mathrm{e}^{-\lambda})\bar{\pi}(\delta)\right). \tag{3.4.9}$$

This proves (3.4.7). Similarly, since under $\mathbb{Q}_\delta^{\psi_\delta}$ and conditionally on ρ, ζ has Poisson distribution with parameter $\bar{\pi}(\delta)\sigma$, a similar computation yields:

$$\mathbf{N}^\psi\left[\mathrm{e}^{-\lambda\xi^\delta}\right]=\mathbb{Q}_\delta^{\psi_\delta}(\mathrm{e}^{-\lambda\zeta})=\mathbb{Q}_\delta^{\psi_\delta}\left(\mathrm{e}^{-(1-\mathrm{e}^{-\lambda})\bar{\pi}(\delta)\sigma}\right)=\frac{1}{\bar{\pi}(\delta)}\int_{(\delta,\infty)}\pi(\mathrm{d}r)\mathbb{P}_r^{\psi_\delta}\left(\mathrm{e}^{-(1-\mathrm{e}^{-\lambda})\bar{\pi}(\delta)\sigma}\right).$$

But, using the Poisson decomposition of Lemma 3.2.6, we get that:

$$\mathbb{P}_r^{\psi_\delta}(\mathrm{e}^{-x\sigma})=\exp\left\{-r\,\mathbf{N}^{\psi_\delta}\left[1-\mathrm{e}^{-x\sigma}\right]\right\}=\mathrm{e}^{-r\psi_\delta^{-1}(x)},\quad \forall x\geq 0, \tag{3.4.10}$$

and (3.4.8) readily follows. □

We end this section with the following result on the criticality of the random forest $\mathbf{t}_\delta$.

Proposition 3.4.7. *Let $\delta>0$ such that $\bar{\pi}(\delta)>0$. The mean of ξ^δ is given by:*

$$\mathbf{N}^\psi[\xi^\delta]=\frac{\int_{(\delta,\infty)}r\,\pi(\mathrm{d}r)}{\alpha+\int_{(\delta,\infty)}r\,\pi(\mathrm{d}r)}. \tag{3.4.11}$$

In particular, under $\mathbf{N}^\psi$, the Bienaymé-Galton-Watson forest $\mathbf{t}_\delta$ is critical (resp. subcritical) if ψ is critical (resp. subcritical).

Proof. Thanks to Proposition 3.4.6, we have:

$$\mathbf{N}^\psi[\xi^\delta]=\mathbb{Q}_\delta^{\psi_\delta}(\zeta)=\bar{\pi}(\delta)\,\mathbb{Q}_\delta^{\psi_\delta}(\sigma)=\int_{(\delta,\infty)}\pi(\mathrm{d}r)\mathbb{P}_r^{\psi_\delta}(\sigma).$$

But the Poissonian decomposition of $\mathbb{P}_r^{\psi_\delta}$ gives:

$$\mathbb{P}_r^{\psi_\delta}(\sigma)=r\,\mathbf{N}^{\psi_\delta}[\sigma]=\frac{r}{\alpha+\int_{(\delta,\infty)}z\,\pi(\mathrm{d}z)},$$

where we used (3.2.13) for the second equality. This yields (3.4.11). □

3.5 Conditioning on $\Delta = \delta$

The goal of this section is to make sense of the conditional measure $\mathbf{N}^{\psi}[\cdot|\Delta = \delta]$. For every $\delta > 0$, we set:

$$w(\delta) = \mathbf{N}^{\psi}[\sigma \mathbf{1}_{\{\Delta<\delta\}}] \quad \text{and} \quad w_{+}(\delta) = \mathbf{N}^{\psi}[\sigma \mathbf{1}_{\{\Delta\leq\delta\}}]. \tag{3.5.1}$$

Notice that if $\delta > 0$ is not an atom of the Lévy measure π, then we have $w(\delta) = w_{+}(\delta)$ by Lemma 3.3.2. Furthermore, thanks to Corollary 3.4.3, (3.2.12) and (3.3.5), we have:

$$w(\delta) = \mathbf{N}^{\psi_{\delta-}}\left[\sigma \mathrm{e}^{-\pi[\delta,\infty)\sigma}\right] = \frac{1}{\psi'_{\delta-} \circ \psi_{\delta-}^{-1}(\pi[\delta,\infty))} = \frac{1}{\psi'_{\delta-}(\mathbf{N}^{\psi}[\Delta \geq \delta])}. \tag{3.5.2}$$

Similarly, we have:

$$w_{+}(\delta) = \mathbf{N}^{\psi_{\delta}}\left[\sigma \mathrm{e}^{-\bar{\pi}(\delta)\sigma}\right] = \frac{1}{\psi'_{\delta} \circ \psi_{\delta}^{-1}(\bar{\pi}(\delta))} = \frac{1}{\psi'_{\delta}(\mathbf{N}^{\psi}[\Delta > \delta])}. \tag{3.5.3}$$

For $\delta > 0$, denote by $\mathbf{P}^{\psi}_{\delta}$ the probability measure on the space $\mathbb{R}_{+} \times \mathscr{D}$ defined by:

$$\int_{\mathbb{R}_{+}\times\mathscr{D}} F \,\mathrm{d}\mathbf{P}^{\psi}_{\delta} = \frac{1}{w(\delta)} \mathbf{N}^{\psi}\left[\int_0^{\sigma} F(s,\rho)\,\mathrm{d}s\, \mathbf{1}_{\{\Delta<\delta\}}\right], \tag{3.5.4}$$

for every $F \in \mathscr{B}_{+}(\mathbb{R}_{+} \times \mathscr{D})$. Similarly, we set:

$$\int_{\mathbb{R}_{+}\times\mathscr{D}} F \,\mathrm{d}\mathbf{P}^{\psi}_{\delta+} = \frac{1}{w_{+}(\delta)} \mathbf{N}^{\psi}\left[\int_0^{\sigma} F(s,\rho)\,\mathrm{d}s\, \mathbf{1}_{\{\Delta\leq\delta\}}\right]. \tag{3.5.5}$$

Observe that $\mathbf{P}^{\psi}_{\delta+} = \lim_{\varepsilon\to 0+} \mathbf{P}^{\psi}_{\delta+\varepsilon}$ in the sense of weak convergence of measures.

For every $\delta, \varepsilon > 0$, let

$$E_{\delta,\varepsilon} = \{\delta - \varepsilon < \Delta < \delta + \varepsilon,\, Z_0^{\delta-\varepsilon} = 1\} \tag{3.5.6}$$

be the event that the maximal degree is between $\delta - \varepsilon$ and $\delta + \varepsilon$ and there is a unique first-generation node with mass larger than $\delta - \varepsilon$. The next lemma states that, under the assumption that δ is not an atom of the Lévy measure π, the two events $E_{\delta,\varepsilon}$ and $\{\delta - \varepsilon < \Delta < \delta + \varepsilon\}$ are equivalent in $\mathbf{N}^{\psi}$-measure as $\varepsilon \to 0$. Recall that π is a measure on $(0,\infty)$ and as such, its support $\operatorname{supp}(\pi)$ does not contain 0.

Lemma 3.5.1. *Assume that $\delta \in \operatorname{supp}(\pi)$ is not an atom of the Lévy measure π and that $\bar{\pi}(\delta) > 0$. We have $\mathbf{N}^{\psi}[\delta - \varepsilon < \Delta < \delta + \varepsilon] \sim \mathbf{N}^{\psi}[E_{\delta,\varepsilon}]$ as $\varepsilon \to 0$.*

Proof. We start by observing that, thanks to the Poissonian decomposition of $\mathbb{P}^{\psi}_{r}$ given in Lemma 3.2.6, we have:

$$\mathbb{P}^{\psi}_{r}(\Delta < \delta) = \begin{cases} 0 & \text{if } r \leq \delta, \\ \mathrm{e}^{-r\mathbf{N}^{\psi}[\Delta\geq\delta]} & \text{if } r > \delta. \end{cases} \tag{3.5.7}$$

Similarly, we have:

$$\mathbb{P}_r^{\psi}(\Delta \le \delta) = \begin{cases} 0 & \text{if } r < \delta, \\ \mathrm{e}^{-r\mathbf{N}^{\psi}[\Delta > \delta]} & \text{if } r \ge \delta. \end{cases} \tag{3.5.8}$$

We deduce that

$$\begin{aligned} \mathbb{Q}_{\delta-\varepsilon}^{\psi}(\Delta < \delta+\varepsilon) &= \frac{1}{\bar{\pi}(\delta-\varepsilon)} \int_{(\delta-\varepsilon,\delta+\varepsilon)} \pi(\mathrm{d}r)\, \mathbb{P}_r^{\psi}(\Delta < \delta+\varepsilon) \\ &= \frac{1}{\bar{\pi}(\delta-\varepsilon)} \int_{(\delta-\varepsilon,\delta+\varepsilon)} \mathrm{e}^{-r\mathbf{N}^{\psi}[\Delta \ge \delta+\varepsilon]}\, \pi(\mathrm{d}r). \end{aligned} \tag{3.5.9}$$

Since $\pi(\delta) = 0$ and $\bar{\pi}(\delta) > 0$, this implies that

$$\lim_{\varepsilon\to 0} \mathbb{Q}_{\delta-\varepsilon}^{\psi}(\Delta < \delta+\varepsilon) = 0. \tag{3.5.10}$$

Under $\mathbf{N}^{\psi_{\delta-\varepsilon}}$ and conditionally on ρ, let ζ be a Poisson random variable with parameter $\bar{\pi}(\delta-\varepsilon)\sigma$ and let $\big((s_i,\rho_i),\, i \ge 1\big)$ be independent with distribution $\sigma^{-1}\mathbf{1}_{[0,\sigma]}(s)\,\mathrm{d}s\,\mathbb{Q}_{\delta-\varepsilon}^{\psi}(\mathrm{d}\tilde{\rho})$, independent of ζ. Thanks to Theorem 3.4.1, we have:

$$\begin{aligned} \mathbf{N}^{\psi}[E_{\delta,\varepsilon}] &= \mathbf{N}^{\psi_{\delta-\varepsilon}}[\zeta = 1,\, \Delta(\rho_1) < \delta+\varepsilon] \\ &= \mathbf{N}^{\psi_{\delta-\varepsilon}}[\bar{\pi}(\delta-\varepsilon)\sigma \mathrm{e}^{-\bar{\pi}(\delta-\varepsilon)\sigma}]\, \mathbb{Q}_{\delta-\varepsilon}^{\psi}(\Delta < \delta+\varepsilon) \\ &= \bar{\pi}(\delta-\varepsilon)\, w_+(\delta-\varepsilon)\, \mathbb{Q}_{\delta-\varepsilon}^{\psi}(\Delta < \delta+\varepsilon), \end{aligned} \tag{3.5.11}$$

where we used (3.5.3) for the last equality. Similarly, we have:

$$\begin{aligned} \mathbf{N}^{\psi}[\delta-\varepsilon < \Delta < \delta+\varepsilon] &= \mathbf{N}^{\psi_{\delta-\varepsilon}}[\zeta \ge 1;\, \forall i \le \zeta,\, \Delta(\rho_i) < \delta+\varepsilon] \\ &= \mathbf{N}^{\psi_{\delta-\varepsilon}}[\zeta \ge 1; \mathbb{Q}_{\delta-\varepsilon}^{\psi}(\Delta < \delta+\varepsilon)^{\zeta}] \\ &= \mathbf{N}^{\psi_{\delta-\varepsilon}}\left[\mathrm{e}^{-\bar{\pi}(\delta-\varepsilon)\sigma}\left(\mathrm{e}^{\bar{\pi}(\delta-\varepsilon)\sigma\,\mathbb{Q}_{\delta-\varepsilon}^{\psi}(\Delta<\delta+\varepsilon)} - 1\right)\right]. \end{aligned}$$

Therefore, using the inequality $\mathrm{e}^x - 1 - x \le x^2\mathrm{e}^x/2$ and the fact that the function $x \mapsto x\mathrm{e}^{-x}$ is bounded on $\mathbb{R}_+$ by some constant $C > 0$, we deduce that

$$\begin{aligned} 0 &\le \mathbf{N}^{\psi}[\delta-\varepsilon < \Delta < \delta+\varepsilon] - \mathbf{N}^{\psi}[E_{\delta,\varepsilon}] \\ &\le \frac{1}{2}\bar{\pi}(\delta-\varepsilon)^2\,\mathbf{N}^{\psi_{\delta-\varepsilon}}\left[\sigma^2\mathrm{e}^{-\bar{\pi}(\delta-\varepsilon)\sigma\,\mathbb{Q}_{\delta-\varepsilon}^{\psi}(\Delta\ge\delta+\varepsilon)}\right]\mathbb{Q}_{\delta-\varepsilon}^{\psi}(\Delta < \delta+\varepsilon)^2 \\ &\le \frac{C}{2}\bar{\pi}(\delta-\varepsilon)\,\mathbf{N}^{\psi_{\delta-\varepsilon}}[\sigma]\,\frac{\mathbb{Q}_{\delta-\varepsilon}^{\psi}(\Delta < \delta+\varepsilon)^2}{\mathbb{Q}_{\delta-\varepsilon}^{\psi}(\Delta \ge \delta+\varepsilon)} \\ &= \frac{C\bar{\pi}(\delta-\varepsilon)}{2(\alpha + \int_{(\delta-\varepsilon,\infty)} r\,\pi(\mathrm{d}r))}\,\frac{\mathbb{Q}_{\delta-\varepsilon}^{\psi}(\Delta < \delta+\varepsilon)^2}{\mathbb{Q}_{\delta-\varepsilon}^{\psi}(\Delta \ge \delta+\varepsilon)} \\ &\le C_\delta\,\frac{\mathbb{Q}_{\delta-\varepsilon}^{\psi}(\Delta < \delta+\varepsilon)^2}{\mathbb{Q}_{\delta-\varepsilon}^{\psi}(\Delta \ge \delta+\varepsilon)} \end{aligned}$$

for $\varepsilon > 0$ small enough and some constant C_δ which is independent of ε, where we used (3.2.13) for the equality.

Furthermore, it is clear from (3.5.3) that

$$\bar{\pi}(\delta-\varepsilon)\,w_+(\delta-\varepsilon) = \bar{\pi}(\delta-\varepsilon)\,\mathbf{N}^\psi[\sigma\mathbf{1}_{\{\Delta\le\delta-\varepsilon\}}] \ge \bar{\pi}(\delta/2)\,\mathbf{N}^\psi[\sigma\mathbf{1}_{\{\Delta\le\delta/2\}}],$$

for $\varepsilon > 0$ small enough. In particular, it follows from (3.5.11) that there exists a constant $C'_\delta > 0$ such that

$$0 \le \frac{\mathbf{N}^\psi[\delta-\varepsilon<\Delta<\delta+\varepsilon]-\mathbf{N}^\psi[E_{\delta,\varepsilon}]}{\mathbf{N}^\psi[E_{\delta,\varepsilon}]} \le C'_\delta\,\frac{\mathbb{Q}^\psi_{\delta-\varepsilon}(\Delta<\delta+\varepsilon)}{\mathbb{Q}^\psi_{\delta-\varepsilon}(\Delta\ge\delta+\varepsilon)},$$

where the right-hand side goes to 0 as $\varepsilon\to 0$ thanks to (3.5.10). This concludes the proof. □

As a consequence, since $E_{\delta,\varepsilon}\subset\{\delta-\varepsilon<\Delta<\delta+\varepsilon\}$, conditioning on either event is equivalent as $\varepsilon\to 0$. We choose to work with the former as computations will be simpler. We shall next give a description of the exploration process conditioned on $E_{\delta,\varepsilon}$.

Let

$$T_\delta = \inf\{t>0:\ \Delta(\rho_t)>\delta\} \tag{3.5.12}$$

be the first time that the exploration process contains an atom with mass larger than δ and let

$$L_\delta = \inf\{t>T_\delta:\ H(\rho_t)<H(\rho_{T_\delta})\} \tag{3.5.13}$$

be the first time that node is erased. We split the path of the exploration process into two parts: $\rho^{\delta,-}$ is the pruned exploration process (that is the exploration process minus the first node with mass larger than δ):

$$\rho^{\delta,-}_t = \begin{cases}\rho_t & \text{if } t<T_\delta,\\ \rho_{t-T_\delta+L_\delta} & \text{if } t\ge T_\delta,\end{cases} \tag{3.5.14}$$

and $\rho^{\delta,+}$ is the path of the exploration process above the unique first-generation node with mass larger than δ:

$$\rho^{\delta,+}_t = \theta_{H_{T_\delta}}(\rho_{(t+T_\delta)\wedge L_\delta}),\quad \forall t\ge 0. \tag{3.5.15}$$

Notice that $\rho^{\delta,+}_0$ is a multiple of the Dirac measure at 0.

Lemma 3.5.2. *Let $F,G\in\mathscr{B}_+(\mathbb{R}_+\times\mathscr{D})$. For every $\delta,\varepsilon>0$ such that $\bar{\pi}(\delta-\varepsilon)>0$, we have:*

$$\mathbf{N}^\psi\left[F(T_{\delta-\varepsilon},\rho^{\delta-\varepsilon,-})G(\rho^{\delta-\varepsilon,+})\middle|E_{\delta,\varepsilon}\right] = \int_{\mathbb{R}_+\times\mathscr{D}} F\,\mathrm{d}\mathbf{P}^\psi_{(\delta-\varepsilon)+}\times\mathbb{Q}^\psi_{\delta-\varepsilon}(G(\rho)|\Delta<\delta+\varepsilon). \tag{3.5.16}$$

Proof. By Theorem 3.4.1, we have

$$\mathbf{N}^\psi\left[F(T_{\delta-\varepsilon},\rho^{\delta-\varepsilon,-})G(\rho^{\delta-\varepsilon,+})\mathbf{1}_{E_{\delta,\varepsilon}}\right] = \mathbf{N}^{\psi_{\delta-\varepsilon}}\left[F(U,\rho)G(\rho^{\delta-\varepsilon})\mathbf{1}_{\{\zeta=1,\Delta(\rho^{\delta-\varepsilon})<\delta+\varepsilon\}}\right],$$

where, under $\mathbf{N}^{\psi_{\delta-\varepsilon}}$, conditionally on ρ, U is uniformly distributed on $[0,\sigma]$, ζ is a Poisson random variable with parameter $\bar{\pi}(\delta-\varepsilon)\sigma$, $\rho^{\delta-\varepsilon}$ has distribution $\mathbb{Q}^{\psi}_{\delta-\varepsilon}$ and they are independent. We deduce that

$$\begin{aligned}\mathbf{N}^{\psi}&\left[F(T_{\delta-\varepsilon},\rho^{\delta-\varepsilon,-})G(\rho^{\delta-\varepsilon,+})\mathbf{1}_{E_{\delta,\varepsilon}}\right]\\ &= \mathbf{N}^{\psi_{\delta-\varepsilon}}\left[\bar{\pi}(\delta-\varepsilon)\mathrm{e}^{-\bar{\pi}(\delta-\varepsilon)\sigma}\int_0^{\sigma} F(s,\rho)\,\mathrm{d}s\right]\mathbb{Q}^{\psi}_{\delta-\varepsilon}(G(\rho)\mathbf{1}_{\{\Delta<\delta+\varepsilon\}}).\end{aligned}$$

Together with Corollary 3.4.3, (3.5.4) and (3.5.11), this yields the desired result. □

We now turn to the study of the asymptotic behavior of the measures appearing in the right-hand side of (3.5.16). Recall that the total variation distance of two probability measures P,Q on some measurable space $(E,\mathscr{E})$ is given by:

$$d_{\mathrm{TV}}(P,Q) = \sup\{|P(A)-Q(A)|:\ A\in\mathscr{E}\}.$$

Lemma 3.5.3. *Assume that $\delta > 0$ is not an atom of the Lévy measure π. Then, the mapping $r \mapsto \mathbf{P}^{\psi}_{r+}$ is continuous at δ in total variation distance and $\mathbf{P}^{\psi}_{\delta+} = \mathbf{P}^{\psi}_{\delta}$.*

Proof. Thanks to Corollary 3.3.2, we have $\mathbf{N}^{\psi}[\Delta=\delta] = 0$. Then the result readily follows from the definition of the measure $\mathbf{P}^{\psi}_{r+}$. □

Recall that the space $\mathscr{M}_f(\mathbb{R}_+)$ is equipped with the topology of weak convergence which makes it a Polish space, see [40, Section 8.3]. It can be metrized by the so-called bounded Lipschitz distance defined for every $\mu,\nu\in\mathscr{M}_f(\mathbb{R}_+)$ by $d_{\mathrm{BL}}(\mu,\nu) = \sup|\langle\mu,f\rangle-\langle\nu,f\rangle|$, where the supremum is taken over all Lipschitz-continuous and bounded functions $f:\mathbb{R}_+\to\mathbb{R}$ such that

$$\sup_{x\geq 0}|f(x)| + \sup_{x\neq y}\frac{|f(x)-f(y)|}{|x-y|} \leq 1.$$

Recall that $\mathscr{D}$ is the space of càdlàg $\mathscr{M}_f(\mathbb{R}_+)$-valued functions defined on $\mathbb{R}_+$, equipped with the Skorokhod J_1-topology and let d_S be the Skorokhod distance associated with the distance d_{BL} on $\mathscr{M}_f(\mathbb{R}_+)$. Denote by $\mathscr{D}_0$ the subset of $\mathscr{D}$ consisting of excursions:

$$\mathscr{D}_0 := \{\mu = (\mu_t,\, t\geq 0)\in\mathscr{D}:\ \sigma(\mu)<\infty,\ \mu_t\neq 0,\ \forall 0<t<\sigma(\mu) \text{ and } \mu_{\sigma(\mu)-} = 0 \text{ if } \sigma(\mu)>0\}, \tag{3.5.17}$$

where $\sigma(\mu) = \inf\{t>0:\ \mu(t+\cdot)\equiv 0\}$ is the lifetime of μ. Notice that if $\mu\in\mathscr{D}_0$ such that $\sigma(\mu)=0$ then necessarily $\mu\equiv 0$. Observe that the mapping $\mu\mapsto\sigma(\mu)$ is measurable with respect to the Skorokhod topology since $\sigma(\mu) = \inf\{t\in\mathbb{Q}\cap(0,\infty):\ \mu_t = 0\}$ and $\mu\mapsto\mu_t$ is measurable. We equip $\mathscr{D}_0$ with the following distance:

$$d_0(\mu,\nu) = d_S(\mu,\nu) + |\sigma(\mu)-\sigma(\nu)|.$$

Lemma 3.5.4. *Let $\nu \in \mathscr{D}$ and $s > 0$. The mapping $\mu \mapsto \nu \circledast (s, \mu)$ is continuous from $(\mathscr{D}_0, d_0)$ to $(\mathscr{D}, d_S)$.*

Proof. Denote by Λ the set of all continuous functions $\lambda\colon \mathbb{R}_+ \to \mathbb{R}_+$ that are (strictly) increasing, with $\lambda(0) = 0$ and $\lim_{t\to\infty} \lambda(t) = \infty$. Let μ_n be a sequence in $\mathscr{D}_0$ converging to μ with respect to the distance d_0. By definition of the Skorokhod topology (see e.g. Jacod and Shiryaev [92, Chapter VI]), this means that there exists a sequence $\lambda_n \in \Lambda$ of time changes such that

$$\lim_{n\to\infty} |\sigma_n - \sigma| = 0, \quad \lim_{n\to\infty} \sup_{t\in\mathbb{R}_+} |\lambda_n(t) - t| = 0 \quad \text{and} \quad \lim_{n\to\infty} \sup_{t\le N} d_{\mathrm{BL}}\big(\mu_n \circ \lambda_n(t), \mu(t)\big) = 0,$$

for every $N \ge 1$, where we set $\sigma_n = \sigma(\mu_n)$ and $\sigma = \sigma(\mu)$.

Let $\kappa_n = \nu \circledast (s, \mu_n)$ and $\kappa = \nu \circledast (s, \mu)$. Our goal is to show that κ_n converges to κ with respect to the Skorokhod topology. To this end, let $\varepsilon_n > 0$ be a sequence converging to 0 such that $\varepsilon_n > \lambda_n(\sigma) - \sigma_n$ and let $\tilde\lambda_n \in \Lambda$ be a time change such that $\tilde\lambda_n(t) = t$ if $t \le s$, $\tilde\lambda_n(s+t) = s + \lambda_n(t)$ if $t \le \sigma$, $\tilde\lambda_n(s+\sigma+t) = s + \sigma_n + t$ if $t \ge \varepsilon_n$ and $\tilde\lambda_n([s+\sigma, s+\sigma+\varepsilon_n]) = [s + \lambda_n(\sigma), s + \sigma_n + \varepsilon_n]$. Notice that if $t \in [s+\sigma, s+\sigma+\varepsilon_n]$, we have:

$$\big|\tilde\lambda_n(t) - t\big| \le |\lambda_n(\sigma) - \sigma - \varepsilon_n| + |\sigma_n + \varepsilon_n - \sigma| \le |\lambda_n(\sigma) - \sigma| + |\sigma_n - \sigma| + 2\varepsilon_n.$$

It follows that

$$\begin{aligned}\sup_{t\in\mathbb{R}_+} \big|\tilde\lambda_n(t) - t\big| &\le \sup_{s\le t\le s+\sigma} \big|\tilde\lambda_n(t) - t\big| + \sup_{s+\sigma\le t\le s+\sigma+\varepsilon_n} \big|\tilde\lambda_n(t) - t\big| + \sup_{t\ge s+\sigma+\varepsilon_n} \big|\tilde\lambda_n(t) - t\big| \\ &\le \sup_{t\le\sigma} |\lambda_n(t) - t| + |\lambda_n(\sigma) - \sigma| + 2\,|\sigma_n - \sigma| + 2\varepsilon_n,\end{aligned}$$

where the right-hand side goes to 0 as $n \to \infty$.

In order to show that κ_n converges to κ in $\mathscr{D}$, it is enough to check that

$$\lim_{n\to\infty} \sup_{t\le N} d_{\mathrm{BL}}\big(\kappa_n \circ \tilde\lambda_n(t), \kappa(t)\big) = 0, \quad \forall N \ge 1.$$

If $t \le s$, we have $\kappa_n \circ \tilde\lambda_n(t) = \kappa(t) = \nu(t)$. If $t \le \sigma$ and $\lambda_n(t) \le \sigma_n$, we have:

$$\kappa_n \circ \tilde\lambda_n(s+t) = \kappa_n(s + \lambda_n(t)) = \big[\nu(s), \mu_n \circ \lambda_n(t)\big] \quad \text{and} \quad \kappa(s+t) = [\nu(s), \mu(t)].$$

It follows that

$$d_{\mathrm{BL}}\big(\kappa_n \circ \tilde\lambda_n(s+t), \kappa(s+t)\big) \le d_{\mathrm{BL}}(\mu_n \circ \lambda_n(t), \mu(t)).$$

On the other hand, if $t \le \sigma$ and $\lambda_n(t) > \sigma_n$, we have:

$$\kappa_n \circ \tilde\lambda_n(s+t) = \nu(s + \lambda_n(t) - \sigma_n) \quad \text{and} \quad \kappa(s+t) = [\nu(s), \mu(t)].$$

In that case, we get:

$$\begin{aligned} d_{\mathrm{BL}}\left(\kappa_n \circ \tilde{\lambda}_n(s+t), \kappa(s+t)\right) &\leq d_{\mathrm{BL}}\left(\nu(s+\lambda_n(t)-\sigma_n), \nu(s)\right) + d_{\mathrm{BL}}\left(\nu(s), [\nu(s), \mu(t)]\right) \\ &\leq d_{\mathrm{BL}}\left(\nu(s+\lambda_n(t)-\sigma_n), \nu(s)\right) + \langle \mu(t), 1\rangle. \end{aligned}$$

If $t \in [s+\sigma, s+\sigma+\varepsilon_n]$, then $\kappa_n \circ \tilde{\lambda}_n(t)$ is of the form $\nu(u)$ with $u \in [s, s+\varepsilon_n]$ or $[\nu(s), \mu_n(u)]$ with $u \in [\lambda_n(\sigma), \sigma_n]$. We deduce that

$$\begin{aligned} d_{\mathrm{BL}}\left(\kappa_n \circ \tilde{\lambda}_n(t), \kappa(t)\right) &\leq \sup_{s\leq u\leq s+\varepsilon_n} d_{\mathrm{BL}}\left(\nu(u), \nu(t-\sigma)\right) + \sup_{\lambda_n(\sigma)\leq u\leq \sigma_n} d_{\mathrm{BL}}\left([\nu(s), \mu_n(u)], \nu(t-\sigma)\right) \\ &\leq 3 \sup_{s\leq u\leq s+\varepsilon_n} d_{\mathrm{BL}}\left(\nu(u), \nu(s)\right) + \sup_{\lambda_n(\sigma)\leq u\leq \sigma_n} \langle \mu_n(u), 1\rangle. \end{aligned}$$

Finally, if $t \geq \varepsilon_n$, then we have $\kappa_n \circ \tilde{\lambda}_n(s+\sigma+t) = \kappa(s+\sigma+t) = \nu(t)$. We deduce that

$$\begin{aligned} &\sup_{t\leq N} d_{\mathrm{BL}}\left(\kappa_n \circ \tilde{\lambda}_n(t), \kappa(t)\right) \\ &\quad \leq \sup_{t\leq N} d_{\mathrm{BL}}\left(\mu_n \circ \lambda_n(t), \mu(t)\right) + \sup_{s\leq u\leq s+(\lambda_n(\sigma)-\sigma_n)_+} d_{\mathrm{BL}}\left(\nu(u), \nu(s)\right) \\ &\quad\quad + 3 \sup_{s\leq u\leq s+\varepsilon_n} d_{\mathrm{BL}}\left(\nu(u), \nu(s)\right) + \sup_{u\leq\sigma, \lambda_n(u)>\sigma_n} \langle \mu(u), 1\rangle + \sup_{\sigma\leq u\leq N} \langle \mu_n \circ \lambda_n(u), 1\rangle. \end{aligned} \tag{3.5.18}$$

Observe that

$$\sup_{u\leq\sigma, \lambda_n(u)>\sigma_n} \langle \mu(u), 1\rangle = \sup_{\lambda_n^{-1}(\sigma_n)<u\leq\sigma} \langle \mu(u), 1\rangle \to 0,$$

since $\lambda_n^{-1}(\sigma_n) \to \sigma$ and since μ is left-continuous at σ and $\mu(\sigma) = 0$. Furthermore, using that $\mu(u) = 0$ for $u \geq \sigma$, we have:

$$\sup_{\sigma\leq u\leq N} \langle \mu_n \circ \lambda_n(u), 1\rangle \leq \sup_{\sigma\leq u\leq N} d_{\mathrm{BL}}(\mu_n \circ \lambda_n(u), \mu(u)) \to 0.$$

Since ν is right-continuous at s, we deduce that the right-hand side of (3.5.18) converges to 0, which concludes the proof. □

Lemma 3.5.5. *For every $\delta \in \mathrm{supp}(\pi)$, the measure $\mathbb{Q}^{\psi}_{\delta-\varepsilon}(\cdot|\Delta < \delta+\varepsilon)$ converges weakly to $\mathbb{P}^{\psi}_{\delta}(\cdot|\Delta \leq \delta)$ as $\varepsilon \to 0$ on the space $(\mathscr{D}_0, d_0)$.*

Remark 3.5.6. Notice that if $\delta = \inf \mathrm{supp}(\pi)$ is positive, then the measure π is necessarily finite and we have:

$$\mathbb{P}^{\psi}_{\delta}(\Delta \leq \delta) \geq \mathbb{P}^{\psi}_{\delta}(\Delta = 0) = \mathrm{e}^{-\delta \mathbf{N}^{\psi}[\Delta>0]} > 0.$$

This implies that the conditional measure $\mathbb{P}^{\psi}_{\delta}(\mathrm{d}\rho|\Delta \leq \delta)$ is well defined.

Proof. It is enough to show that for every Lipschitz-continuous and bounded function $F\colon \mathscr{D}_0 \to \mathbb{R}$, the following convergence holds:

$$\lim_{\varepsilon\to 0} \mathbb{Q}^{\psi}_{\delta-\varepsilon}\left(F(\rho)\middle|\Delta < \delta+\varepsilon\right) = \mathbb{P}^{\psi}_{\delta}\left(F(\rho)\middle|\Delta \leq \delta\right).$$

Fix such a function F. From the definition of $\mathbb{Q}^{\psi}_{\delta-\varepsilon}$, we have:

$$\mathbb{Q}^{\psi}_{\delta-\varepsilon}\left(F(\rho)\mathbf{1}_{\{\Delta<\delta+\varepsilon\}}\right) = \frac{1}{\bar{\pi}(\delta-\varepsilon)}\int_{(\delta-\varepsilon,\delta+\varepsilon)} \pi(\mathrm{d}r)\,\mathbb{P}^{\psi}_{r}\left(F(\rho)\mathbf{1}_{\{\Delta<\delta+\varepsilon\}}\right).$$

In conjunction with (3.5.9), this gives:

$$\mathbb{Q}^{\psi}_{\delta-\varepsilon}\left(F(\rho)\middle|\Delta<\delta+\varepsilon\right)$$
$$= \frac{1}{\int_{(\delta-\varepsilon,\delta+\varepsilon)} \mathrm{e}^{-r\mathbf{N}^{\psi}[\Delta\geq\delta+\varepsilon]}\,\pi(\mathrm{d}r)}\int_{(\delta-\varepsilon,\delta+\varepsilon)} \pi(\mathrm{d}r)\,\mathbb{P}^{\psi}_{r}\left(F(\rho)\mathbf{1}_{\{\Delta<\delta+\varepsilon\}}\right).$$

Now it is not difficult to show that, as $\varepsilon\to 0$, we have:

$$\int_{(\delta-\varepsilon,\delta+\varepsilon)} \mathrm{e}^{-r\mathbf{N}^{\psi}[\Delta\geq\delta+\varepsilon]}\,\pi(\mathrm{d}r) \sim \pi(\delta-\varepsilon,\delta+\varepsilon)\mathrm{e}^{-\delta\mathbf{N}^{\psi}[\Delta>\delta]}.$$

Thus, as $\varepsilon\to 0$, we have:

$$\mathbb{Q}^{\psi}_{\delta}\left(F(\rho)\middle|\Delta<\delta+\varepsilon\right) \sim \frac{\mathrm{e}^{\delta\mathbf{N}^{\psi}[\Delta>\delta]}}{\pi(\delta-\varepsilon,\delta+\varepsilon)}\int_{(\delta-\varepsilon,\delta+\varepsilon)} \pi(\mathrm{d}r)\,\mathbb{P}^{\psi}_{r}\left(F(\rho)\mathbf{1}_{\{\Delta<\delta+\varepsilon\}}\right).$$

Thanks to (3.5.8), we have $\mathbb{P}^{\psi}_{\delta}(\Delta\leq\delta) = \mathrm{e}^{-\delta\mathbf{N}^{\psi}[\Delta>\delta]}$. Thus, in order to prove the result, it is enough to show that

$$\lim_{\varepsilon\to 0}\frac{1}{\pi(\delta-\varepsilon,\delta+\varepsilon)}\int_{(\delta-\varepsilon,\delta+\varepsilon)} \pi(\mathrm{d}r)\,\mathbb{P}^{\psi}_{r}\left(F(\rho)\mathbf{1}_{\{\Delta<\delta+\varepsilon\}}\right) = \mathbb{P}^{\psi}_{\delta}\left(F(\rho)\mathbf{1}_{\{\Delta\leq\delta\}}\right). \tag{3.5.19}$$

Write:

$$\frac{1}{\pi(\delta-\varepsilon,\delta+\varepsilon)}\int_{(\delta-\varepsilon,\delta+\varepsilon)} \pi(\mathrm{d}r)\,\mathbb{P}^{\psi}_{r}\left(F(\rho)\mathbf{1}_{\{\Delta<\delta+\varepsilon\}}\right) - \mathbb{P}^{\psi}_{\delta}\left(F(\rho)\mathbf{1}_{\{\Delta\leq\delta\}}\right)$$
$$= \frac{1}{\pi(\delta-\varepsilon,\delta+\varepsilon)}\int_{(\delta-\varepsilon,\delta+\varepsilon)} \pi(\mathrm{d}r)\left[\mathbb{P}^{\psi}_{r}\left(F(\rho)\mathbf{1}_{\{\Delta<\delta+\varepsilon\}}\right) - \mathbb{P}^{\psi}_{\delta}\left(F(\rho)\mathbf{1}_{\{\Delta<\delta+\varepsilon\}}\right)\right]$$
$$+\mathbb{P}^{\psi}_{\delta}\left(F(\rho)\mathbf{1}_{\{\delta<\Delta<\delta+\varepsilon\}}\right).$$

By dominated convergence, it is clear that the second term on the right-hand side converges to 0.

For the first term, one can couple the measures $\mathbb{P}^{\psi}_{r}$ and $\mathbb{P}^{\psi}_{\delta}$ in the following way. Let ρ be the exploration process with branching mechanism ψ starting from 0 and let $(L^0_t,\, t\geq 0)$ be its local time process at 0. Then the process $\tilde{\rho}^{(r)}$ defined in (3.2.20) has distribution $\mathbb{P}^{\psi}_{r}$ while $\tilde{\rho}^{(\delta)}$ has distribution $\mathbb{P}^{\psi}_{\delta}$. It follows that

$$\left|\mathbb{P}^{\psi}_{r}\left(F(\rho)\mathbf{1}_{\{\Delta<\delta+\varepsilon\}}\right) - \mathbb{P}^{\psi}_{\delta}\left(F(\rho)\mathbf{1}_{\{\Delta<\delta+\varepsilon\}}\right)\right|$$

$$= \left| \mathbb{E}\left[F(\tilde{\rho}^{(r)}) \mathbf{1}_{\left\{\sup_{L_t^0 \le r} \Delta(\rho_t) < \delta + \varepsilon\right\}} - F(\tilde{\rho}^{(\delta)}) \mathbf{1}_{\left\{\sup_{L_t^0 \le \delta} \Delta(\rho_t) < \delta + \varepsilon\right\}} \right] \right|$$
$$\le C \mathbb{P}\left(\sup_{\delta < L_t^0 \le r} \Delta(\rho_t) \ge \delta + \varepsilon \right) + C \mathbb{E}\left[1 \wedge d_0(\tilde{\rho}^{(r)}, \tilde{\rho}^{(\delta)}) \right]. \tag{3.5.20}$$

Using the Poissonian decomposition from Lemma 3.2.6, we have for $r \in (\delta, \delta + \varepsilon)$:

$$\mathbb{P}\left(\sup_{\delta < L_t^0 \le r} \Delta(\rho_t) \ge \delta + \varepsilon \right) \le \mathbb{P}\left(\sup_{\delta < L_t^0 < \delta + \varepsilon} \Delta(\rho_t) \ge \delta \right) = \mathbb{P}\left(\sup_{\delta < -I_{\alpha_i} < \delta + \varepsilon} \Delta(\rho^i) \ge \delta \right) = 1 - \mathrm{e}^{-\varepsilon \mathbf{N}^{\psi}[\Delta \ge \delta]}. \tag{3.5.21}$$

Next, by definition of d_0 we have that $d_0(\tilde{\rho}^{(r)}, \tilde{\rho}^{(\delta)}) = |\sigma(\tilde{\rho}^{(r)}) - \sigma(\tilde{\rho}^{(\delta)})| + d_S(\tilde{\rho}^{(r)}, \tilde{\rho}^{(\delta)})$. We introduce the right-continuous inverse S of the local time process at 0 given by:

$$S_r = \inf\{t > 0 \colon L_t^0 > r\}, \quad \forall r > 0.$$

It is well known that the process S is a subordinator. Then the process $\tilde{\rho}^{(r)}$ has lifetime S_r. Furthermore, we have:

$$d_S(\tilde{\rho}^{(r)}, \tilde{\rho}^{(\delta)}) \le \sup_{t \ge 0} d_{\mathrm{BL}}(\tilde{\rho}_t^{(r)}, \tilde{\rho}_t^{(\delta)}).$$

For $L_t^0 \le \delta$, the processes $\tilde{\rho}^{(r)}$ and $\tilde{\rho}^{(\delta)}$ differ only by their masses at 0 so that $d_{\mathrm{BL}}(\tilde{\rho}_t^{(r)}, \tilde{\rho}_t^{(\delta)}) \le r - \delta \le \varepsilon$. On the other hand, for $L_t^0 > \delta$, we have $\tilde{\rho}^{(\delta)} = 0$ so that

$$d_{\mathrm{BL}}(\tilde{\rho}_t^{(r)}, \tilde{\rho}_t^{(\delta)}) = \langle \tilde{\rho}_t^{(r)}, 1 \rangle = (r - L_t^0) + \langle \rho_t, 1 \rangle \le \varepsilon + \langle \rho_t, 1 \rangle,$$

where we recall from Section 3.2.3 that $\langle \rho_t, 1 \rangle = X_t - I_t$, where X is the underlying Lévy process and I is its running infimum. It follows that

$$d_0(\tilde{\rho}^{(r)}, \tilde{\rho}^{(\delta)}) \le S_{\delta+\varepsilon} - S_\delta + \varepsilon + \sup_{\delta < L_t^0 < \delta + \varepsilon} (X_t - I_t), \tag{3.5.22}$$

where the right-hand side converges to 0 a.s. as $\varepsilon \to 0$.

Combining (3.5.20)–(3.5.22), we deduce that

$$\left| \mathbb{P}_r^{\psi}\left(F(\rho) \mathbf{1}_{\{\Delta < \delta + \varepsilon\}} \right) - \mathbb{P}_\delta^{\psi}\left(F(\rho) \mathbf{1}_{\{\Delta < \delta + \varepsilon\}} \right) \right| \le C_1(\varepsilon),$$

for every $r \in (\delta, \delta + \varepsilon)$, where $C_1(\varepsilon)$ does not depend on r and goes to 0 as $\varepsilon \to 0$. Similarly, for every $r \in (\delta - \varepsilon, \delta)$, we have:

$$\left| \mathbb{P}_r^{\psi}\left(F(\rho) \mathbf{1}_{\{\Delta < \delta + \varepsilon\}} \right) - \mathbb{P}_\delta^{\psi}\left(F(\rho) \mathbf{1}_{\{\Delta < \delta + \varepsilon\}} \right) \right| \le C_2(\varepsilon).$$

Finally, we deduce that

$$\frac{1}{\pi(\delta - \varepsilon, \delta + \varepsilon)} \int_{(\delta - \varepsilon, \delta + \varepsilon)} \pi(\mathrm{d}r) \left| \mathbb{P}_r^{\psi}\left(F(\rho) \mathbf{1}_{\{\Delta < \delta + \varepsilon\}} \right) - \mathbb{P}_\delta^{\psi}\left(F(\rho) \mathbf{1}_{\{\Delta < \delta + \varepsilon\}} \right) \right| \le C_1(\varepsilon) + C_2(\varepsilon).$$

Letting $\varepsilon \to 0$ proves (3.5.19) and the proof is complete. □

We are now in a position to prove the main result of this section which gives a description of the Lévy tree conditioned on having maximal degree δ. For every atom $\delta > 0$ of π, we set:

$$\mathfrak{g}(\delta) = \pi(\delta)\,\mathbb{P}^{\psi}_{\delta}(\Delta \le \delta) = \pi(\delta)\mathrm{e}^{-\delta \mathbf{N}^{\psi}[\Delta > \delta]}, \tag{3.5.23}$$

where the last equality is due to (3.5.8). Under $\mathbf{N}^{\psi}$, denote by M_δ the random variable defined by:

$$M_\delta = \frac{\mathrm{e}^{\mathfrak{g}(\delta)\sigma} - 1}{\mathfrak{g}(\delta)}.$$

This should be interpreted as $M_\delta = \sigma$ if δ is not an atom of π.

For every atom $\delta > 0$ of the Lévy measure π, we define a probability measure $\mathbf{P}^{\psi,\mathrm{a}}_{\delta}$ on the space $\mathscr{D}$ as follows. Take $\tilde{\rho}$ with distribution $\mathbf{N}^{\psi}[M_\delta \mathbf{1}_{\{\Delta<\delta\}}]^{-1}\mathbf{N}^{\psi}[M_\delta \mathbf{1}_{\{\Delta<\delta\}}\,\mathrm{d}\rho]$, and, conditionally on $\tilde{\rho}$, let $\big((s_i,\rho_i),\, i \in I\big)$ be the atoms of a Poisson point measure with intensity $\mathfrak{g}(\delta)\mathbf{1}_{[0,\sigma]}(s)\,\mathrm{d}s\, \mathbb{P}^{\psi}_{\delta}(\mathrm{d}\hat{\rho}|\Delta \le \delta)$ conditioned on containing at least one point. Then $\mathbf{P}^{\psi,\mathrm{a}}_{\delta}$ is defined as the distribution of the process $\tilde{\rho} \circledast_{i\in I} (s_i,\rho_i)$.

Theorem 3.5.7. *There exists a regular conditional probability* $\mathbf{N}^{\psi}[\cdot|\Delta=\delta]$ *for* $\delta>0$ *such that* $\pi[\delta,\infty) > 0$, *which is given by, for every* $F \in \mathscr{B}_+(\mathscr{D})$:

$$\begin{aligned}\mathbf{N}^{\psi}[F(\rho)|\Delta=\delta] = \frac{1}{\mathbf{N}^{\psi}[M_\delta \mathbf{1}_{\{\Delta<\delta\}}]} \sum_{k=0}^{\infty} \frac{\mathfrak{g}(\delta)^k}{(k+1)!} \\ \times \mathbf{N}^{\psi}\left[\int \prod_{i=1}^{k+1} \mathbf{1}_{[0,\sigma]}(s_i)\,\mathrm{d}s_i\, \mathbb{P}^{\psi}_{\delta}(\mathrm{d}\rho_i|\Delta\le\delta) F(\rho \circledast_{i=1}^{k+1} (s_i,\rho_i))\mathbf{1}_{\{\Delta<\delta\}}\right].\end{aligned} \tag{3.5.24}$$

In particular, if $\delta > 0$ *is not an atom of the Lévy measure* π, *we have:*

$$\mathbf{N}^{\psi}[F(\rho)|\Delta=\delta] = \int_{\mathbb{R}_+\times\mathscr{D}} \mathbf{P}^{\psi}_{\delta}(\mathrm{d}s,\mathrm{d}\tilde{\rho}) \int_{\mathscr{D}} \mathbb{P}^{\psi}_{\delta}(\mathrm{d}\hat{\rho}|\Delta\le\delta) F(\tilde{\rho}\circledast(s,\hat{\rho})). \tag{3.5.25}$$

If $\delta>0$ *is an atom of* π, *we have:*

$$\mathbf{N}^{\psi}[F(\rho)|\Delta=\delta] = \int_{\mathscr{D}} \mathbf{P}^{\psi,\mathrm{a}}_{\delta}(\mathrm{d}\tilde{\rho})\, F(\tilde{\rho}). \tag{3.5.26}$$

Remark 3.5.8. Let E_δ be the event that the maximal degree is δ and there is a unique first-generation node with mass δ. We have:

$$\mathbf{N}^{\psi}\left[F(\rho)\middle|E_\delta\right] = \int_{\mathbb{R}_+\times\mathscr{D}} \mathbf{P}^{\psi}_{\delta}(\mathrm{d}s,\mathrm{d}\tilde{\rho}) \int_{\mathscr{D}} \mathbb{P}^{\psi}_{\delta}(\mathrm{d}\hat{\rho}|\Delta\le\delta) F(\tilde{\rho}\circledast(s,\hat{\rho})). \tag{3.5.27}$$

When δ is an atom of π, this can be proved by taking $k=0$ in (3.5.24). Indeed, we have:

$$\mathbf{N}^{\psi}\left[F(\rho)\mathbf{1}_{E_\delta}\middle|\Delta=\delta\right] = \frac{1}{\mathbf{N}^{\psi}[M_\delta \mathbf{1}_{\{\Delta<\delta\}}]} \mathbf{N}^{\psi}\left[\int \mathbf{1}_{[0,\sigma]}(s)\,\mathrm{d}s\,\mathbb{P}^{\psi}_{\delta}(\mathrm{d}\hat{\rho}|\Delta\le\delta) F(\rho\circledast(s,\hat{\rho}))\right],$$

and the result follows by conditionning. When δ is not an atom of π, this follows from Theorem 3.5.7 together with the fact that, conditionally on $\Delta = \delta$, there is a unique node with mass δ (see Corollary 3.5.9 below). In other words, conditioning the exploration process by E_δ when δ is an atom of π yields the same distribution as conditioning by $\Delta = \delta$ when δ is not an atom of π.

Proof. Assume that $\delta \in \operatorname{supp}(\pi)$ is an atom of π. Then the event $\{\Delta = \delta\}$ has positive $\mathbf{N}^\psi$-measure (see Corollary 3.3.2) and it follows from Theorem 3.4.1 that ρ conditioned on $\Delta = \delta$ has distribution $\mathbf{P}_\delta^{\psi,\mathrm{a}}$.

Assume then that $\delta \in \operatorname{supp}(\pi)$ is not an atom of π and let $F\colon \mathscr{D} \to \mathbb{R}$ be continuous and bounded. Applying Lemma 3.5.1 and using the fact $E_{\delta,\varepsilon} \subset \{\delta - \varepsilon < \Delta < \delta + \varepsilon\}$, we have as $\varepsilon \to 0$:

$$\mathbf{N}^\psi[F(\rho)|\delta - \varepsilon < \Delta < \delta + \varepsilon] \sim \mathbf{N}^\psi[F(\rho)|E_{\delta,\varepsilon}].$$

But, thanks to Lemma 3.5.2, we have:

$$\begin{aligned} \mathbf{N}^\psi[F(\rho)|E_{\delta,\varepsilon}] &= \mathbf{N}^\psi[F(\rho^{\delta-\varepsilon,-} \circledast (T_{\delta-\varepsilon}, \rho^{\delta-\varepsilon,+})|E_{\delta,\varepsilon}] \\ &= \int_{\mathbb{R}_+ \times \mathscr{D}} \mathbf{P}^\psi_{(\delta-\varepsilon)+}(\mathrm{d}s, \mathrm{d}\tilde\rho) \int_{\mathscr{D}} \mathbb{Q}^\psi_{\delta-\varepsilon}(\mathrm{d}\hat\rho|\Delta < \delta + \varepsilon) F(\tilde\rho \circledast (s, \hat\rho)). \end{aligned} \tag{3.5.28}$$

Recall from Lemma 3.5.4 that for every fixed $(\nu, s) \in \mathscr{D} \times (0, \infty)$, the mapping $\mu \mapsto \nu \circledast (s, \mu)$ is continuous from $\mathscr{D}_0$ to $\mathscr{D}$. Together with Lemma 3.5.3 and Lemma 3.5.5, this gives:

$$\lim_{\varepsilon \to 0} \mathbf{N}^\psi\left[F(\rho)|\delta - \varepsilon < \Delta < \delta + \varepsilon\right] = \int_{\mathbb{R}_+ \times \mathscr{D}} \mathbf{P}^\psi_\delta(\mathrm{d}s, \mathrm{d}\tilde\rho) \int_{\mathscr{D}} \mathbb{P}^\psi_\delta(\mathrm{d}\hat\rho|\Delta \le \delta) F(\tilde\rho \circledast (s, \hat\rho)).$$

A standard result on measure differentiation, see e.g. [68, Theorem 1.30], yields the desired result. □

Corollary 3.5.9. *Assume that $\delta > 0$ is not an atom of the Lévy measure π. Then, under $\mathbf{N}^\psi$ and conditionally on $\Delta = \delta$, there is a unique node with mass δ.*

Proof. Notice that $\mathbf{P}^\psi_\delta$-a.s. $\Delta(\rho) < \delta$ by definition. Thus, thanks to Theorem 3.5.7, it is enough to show that $\mathbb{P}^\psi_\delta(\cdot|\Delta \le \delta)$-a.s. there is a unique node with mass δ. We shall use the Poissonian decomposition from Lemma 3.2.6. Let $\sum_{i\in I} \delta_{(\ell_i, \rho^i)}$ be a point measure with distribution $\mathbb{P}^\psi_\delta$, that is a Poisson point measure with intensity $\mathbf{1}_{[0,\delta]}(\ell)\,\mathrm{d}\ell\,\mathbf{N}^\psi[\mathrm{d}\rho]$. Then it suffices to check that, conditionally on $\sup_{i\in I} \Delta(\rho^i) \le \delta$, it holds that $\sup_{i \in I} \Delta(\rho^i) < \delta$.

Since $\mathbf{N}^\psi[\Delta > \delta/2] < \infty$, only finitely many ρ^i are such that $\Delta(\rho^i) > \delta/2$. We deduce that

$$\mathbb{P}\left(\sup_{i\in I} \Delta(\rho^i) < \delta\right) = \mathbb{P}\left(\sup_{i\in I} \Delta(\rho^i) \le \delta,\, \Delta(\rho^i) \neq \delta \text{ for all } i \in I\right).$$

Thanks to Corollary 3.3.2, we have $\mathbf{N}^\psi[\Delta = \delta] = 0$, which implies that $\Delta(\rho^i) \neq \delta$ for all $i \in I$ almost surely. Therefore we get:

$$\mathbb{P}\left(\sup_{i\in I} \Delta(\rho^i) < \delta\right) = \mathbb{P}\left(\sup_{i\in I} \Delta(\rho^i) \leq \delta\right).$$

This proves the result. □

As an application of Theorem 3.5.7, we can compute the joint distribution of the degree Δ of the exploration process when the Lévy measure π is diffuse and the height H_Δ of the (unique) node with mass Δ. We start by determining the distribution of $H(\rho_s)$ under $\mathbf{P}^\psi_\delta(\mathrm{d}s, \mathrm{d}\rho)$. Recall from (3.5.1) the definition of w.

Lemma 3.5.10. *Under $\mathbf{P}^\psi_\delta(\mathrm{d}s, \mathrm{d}\rho)$ (resp. $\mathbf{P}^\psi_{\delta+}(\mathrm{d}s, \mathrm{d}\rho)$), the random variable $H(\rho_s)$ is exponentially distributed with mean $w(\delta)$ (resp. $w_+(\delta)$).*

Proof. We only prove the result under $\mathbf{P}^\psi_\delta$, the other being similar. By definition, we have:

$$\begin{aligned}\int_{\mathbb{R}_+\times\mathscr{D}} \mathbf{1}_{\{H(\rho_s)>h\}}\, \mathbf{P}^\psi_\delta(\mathrm{d}s, \mathrm{d}\rho) &= \frac{1}{w(\delta)} \mathbf{N}^\psi\left[\int_0^\sigma \mathbf{1}_{\{H_s>h\}}\, \mathrm{d}s\, \mathbf{1}_{\{\Delta<\delta\}}\right] \\ &= \frac{1}{w(\delta)} \mathbf{N}^{\psi_{\delta-}}\left[\mathrm{e}^{-\pi[\delta,\infty)\sigma} \int_0^\sigma \mathbf{1}_{\{H_s>h\}}\right],\end{aligned}$$

where we used Corollary 3.4.3 for the last equality.

Thanks to Bismut's decomposition, see e.g. [3, Theorem 2.1], we have for every $\lambda > 0$:

$$\begin{aligned}&\mathbf{N}^{\psi_{\delta-}}\left[\mathrm{e}^{-\lambda\sigma} \int_0^\sigma \mathbf{1}_{\{H_s>h\}}\, \mathrm{d}s\right] \\ &= \int_h^\infty \mathrm{d}t \exp\left\{-t\left[\psi'_{\delta-}(0) + 2\beta \mathbf{N}^{\psi_{\delta-}}[1-\mathrm{e}^{-\lambda\sigma}] + \int_{(0,\delta)} r\,\pi(\mathrm{d}r)\, \mathbb{P}^{\psi_{\delta-}}_r(1-\mathrm{e}^{-\lambda\sigma})\right]\right\} \\ &= \int_h^\infty \mathrm{d}t \exp\left\{-t\left[\psi'_{\delta-}(0) + 2\beta\psi^{-1}_{\delta-}(\lambda) + \int_{(0,\delta)} r(1-\mathrm{e}^{-r\psi^{-1}_{\delta-}(\lambda)})\,\pi(\mathrm{d}r)\right]\right\} \\ &= \int_h^\infty \mathrm{d}t\, \mathrm{e}^{-t\psi'_{\delta-}\circ\psi^{-1}_{\delta-}(\lambda)} \\ &= \frac{1}{\psi'_{\delta-}\circ\psi^{-1}_{\delta-}(\lambda)} \mathrm{e}^{-h\psi'_{\delta-}\circ\psi^{-1}_{\delta-}(\lambda)}.\end{aligned} \tag{3.5.29}$$

Applying this to $\lambda = \pi[\delta,\infty)$ and using (3.5.2), it follows that

$$\int_{\mathbb{R}_+\times\mathscr{D}} \mathbf{1}_{\{H(\rho_s)>h\}}\, \mathbf{P}^\psi_\delta(\mathrm{d}s, \mathrm{d}\rho) = \mathrm{e}^{-h/w(\delta)}.$$

This proves the result. □

Let

$$T_\Delta = \inf\{t \geq 0 \colon \ \Delta(\rho_t) = \Delta\} \tag{3.5.30}$$

be the first time that ρ contains an atom with mass Δ and let $H_\Delta = H(\rho_{T_\Delta})$ be the value of the height process at that time. We shall determine the joint distribution of (Δ, H_Δ) assuming that the Lévy measure π is diffuse.

Proposition 3.5.11. *Assume that the Lévy measure π is diffuse. Then, $\mathbf{N}^\psi$-a.e. there is a unique node with mass Δ. Furthermore, for every $\delta, h > 0$, we have:*

$$\mathbf{N}^\psi[\Delta > \delta, H_\Delta > h] = \int_{(\delta,\infty)} \mathrm{e}^{-h/w(r)} \, \mathbf{N}^\psi[\Delta \in \mathrm{d}r]. \tag{3.5.31}$$

In other words, under $\mathbf{N}^\psi$ and conditionally on $\Delta = \delta$, H_Δ is exponentially distributed with mean $w(\delta)$.

Question 3.5.12. If δ is an atom of π, what is the distribution of the height of the MRCA of the nodes with mass exactly δ under $\mathbf{N}^\psi$, conditionally on $\Delta = \delta$?

Proof. The first part follows from Corollary 3.5.9. Then, using Theorem 3.5.7, we have:

$$\mathbf{N}^\psi[\Delta > \delta, H_\Delta > h] = \int_{(\delta,\infty)} \mathbf{N}^\psi[H_\Delta > h | \Delta = r] \, \mathbf{N}^\psi[\Delta \in \mathrm{d}r].$$

Now under $\mathbf{N}^\psi[\cdot|\Delta = r]$, H_Δ is distributed as $H_s = H(\rho_s)$ under $\mathbf{P}^\psi_r(\mathrm{d}s, \mathrm{d}\rho)$. Lemma 3.5.10 allows to conclude. □

3.6 Conditioning on $\Delta = \delta$ and $H_\Delta = h$

In this section, we assume that the Lévy measure π is diffuse. Recall then from Proposition 3.5.11 that there is a unique node with mass Δ and H_Δ is its height. The goal of this section is to make sense of the conditional measure $\mathbf{N}^\psi[\cdot|\Delta = \delta, H_\Delta = h]$. Let

$$F_{\delta,\varepsilon} = \{\delta - \varepsilon < \Delta < \delta + \varepsilon, \, Z_0^{\delta-\varepsilon} = 1, \, h - \varepsilon < H(\rho_{T_{\delta-\varepsilon}}) < h + \varepsilon\} \tag{3.6.1}$$

be the event that the maximal degree is between $\delta - \varepsilon$ and $\delta + \varepsilon$, there is a unique first-generation node with mass larger than $\delta - \varepsilon$ and its height is between $h - \varepsilon$ and $h + \varepsilon$. Recall from (3.5.1) the definition of w.

Lemma 3.6.1. *Assume that the Lévy measure π is diffuse. For every $\delta \in \operatorname{supp}(\pi)$ such that $\bar{\pi}(\delta) > 0$ and $h > 0$, we have as $\varepsilon \to 0$:*

$$\begin{aligned} \mathbf{N}^\psi[\delta - \varepsilon < \Delta < \delta + \varepsilon, h - \varepsilon < H_\Delta < h + \varepsilon] &\sim \mathbf{N}^\psi[F_{\delta,\varepsilon}] \\ &\sim 2\varepsilon \, \mathbb{Q}^\psi_{\delta-\varepsilon}(\Delta < \delta + \varepsilon) \bar{\pi}(\delta) \mathrm{e}^{-h/w(\delta)}. \end{aligned} \tag{3.6.2}$$

Proof. By Proposition 3.5.11, we have:

$$\mathbf{N}^{\psi}[\delta-\varepsilon<\Delta<\delta+\varepsilon, h-\varepsilon<H_{\Delta}<h+\varepsilon] = \int_{(\delta-\varepsilon,\delta+\varepsilon)} \mathbf{N}^{\psi}[\Delta\in \mathrm{d}r]\, w(r)^{-1}\int_{h-\varepsilon}^{h+\varepsilon} \mathrm{e}^{-t/w(r)}\,\mathrm{d}t.$$

A straightforward application of the dominated convergence theorem gives:

$$\mathbf{N}^{\psi}[\delta-\varepsilon<\Delta<\delta+\varepsilon, h-\varepsilon<H_{\Delta}<h+\varepsilon] \sim 2\varepsilon\int_{(\delta-\varepsilon,\delta+\varepsilon)} \mathbf{N}^{\psi}[\Delta\in \mathrm{d}r]\, w(r)^{-1}\mathrm{e}^{-h/w(r)}.$$

Since π is diffuse, observe that $\mathbf{N}^{\psi}[\sigma\mathbf{1}_{\{\Delta=r\}}]=0$ for every $r>0$ thanks to Corollary 3.3.2. This implies that w is continuous and we deduce that

$$\mathbf{N}^{\psi}[\delta-\varepsilon<\Delta<\delta+\varepsilon, h-\varepsilon<H_{\Delta}<h+\varepsilon] \sim 2\varepsilon\mathbf{N}^{\psi}[\delta-\varepsilon<\Delta<\delta+\varepsilon]\, w(\delta)^{-1}\mathrm{e}^{-h/w(\delta)}.$$

But Lemma 3.5.1 gives:

$$\mathbf{N}^{\psi}[\delta-\varepsilon<\Delta<\delta+\varepsilon] \sim \mathbf{N}^{\psi}[E_{\delta,\varepsilon}].$$

Moreover, thanks to (3.5.11) and the continuity of w, we have:

$$\mathbf{N}^{\psi}[E_{\delta,\varepsilon}] = \bar{\pi}(\delta-\varepsilon)\, w(\delta-\varepsilon)\, \mathbb{Q}^{\psi}_{\delta-\varepsilon}(\Delta<\delta+\varepsilon) \sim \bar{\pi}(\delta)\, w(\delta)\, \mathbb{Q}^{\psi}_{\delta-\varepsilon}(\Delta<\delta+\varepsilon).$$

We deduce that

$$\mathbf{N}^{\psi}[\delta-\varepsilon<\Delta<\delta+\varepsilon, h-\varepsilon<H_{\Delta}<h+\varepsilon] \sim 2\varepsilon\,\mathbb{Q}^{\psi}_{\delta-\varepsilon}(\Delta<\delta+\varepsilon)\bar{\pi}(\delta)\mathrm{e}^{-h/w(\delta)}.$$

On the other hand, thanks to Theorem 3.4.1, we have:

$$\mathbf{N}^{\psi}[F_{\delta,\varepsilon}] = \mathbf{N}^{\psi_{\delta-\varepsilon}}\left[\bar{\pi}(\delta-\varepsilon)\mathrm{e}^{-\bar{\pi}(\delta-\varepsilon)\sigma}\int_0^{\sigma}\mathbf{1}_{\{h-\varepsilon<H_s<h+\varepsilon\}}\,\mathrm{d}s\right]\mathbb{Q}^{\psi}_{\delta-\varepsilon}(\Delta<\delta+\varepsilon). \tag{3.6.3}$$

Using Bismut's decomposition as in (3.5.29) , we get:

$$\mathbf{N}^{\psi_{\delta-\varepsilon}}\left[\mathrm{e}^{-\bar{\pi}(\delta-\varepsilon)\sigma}\int_0^{\sigma}\mathbf{1}_{\{h-\varepsilon<H_s<h+\varepsilon\}}\,\mathrm{d}s\right] = \int_{h-\varepsilon}^{h+\varepsilon}\mathrm{e}^{-t/w(\delta-\varepsilon)}\,\mathrm{d}t \sim 2\varepsilon\mathrm{e}^{-h/w(\delta)}, \tag{3.6.4}$$

where again we used the continuity of w. It follows that

$$\mathbf{N}^{\psi}[F_{\delta,\varepsilon}] \sim 2\varepsilon\,\mathbb{Q}^{\psi}_{\delta-\varepsilon}(\Delta<\delta+\varepsilon)\bar{\pi}(\delta)\mathrm{e}^{-h/w(\delta)}.$$

□

For every $\delta, h>0$, denote by $\mathbf{P}^{\psi}_{\delta,h}$ the probability measure on the space $\mathbb{R}_+\times\mathscr{D}$ defined by:

$$\int_{\mathbb{R}_+\times\mathscr{D}} F\,\mathrm{d}\mathbf{P}^{\psi}_{\delta,h} = \frac{1}{\mathbf{N}^{\psi}[L^h_{\sigma}\mathbf{1}_{\{\Delta<\delta\}}]}\mathbf{N}^{\psi}\left[\int_0^{\sigma}F(s,\rho)\,L^h(\mathrm{d}s)\,\mathbf{1}_{\{\Delta<\delta\}}\right], \tag{3.6.5}$$

for every $F \in \mathscr{B}_+(\mathbb{R}_+ \times \mathscr{D})$. Since we are assuming that the Lévy measure π is diffuse, we may replace the event $\{\Delta < \delta\}$ by $\{\Delta \le \delta\}$ thanks to Corollary 3.3.2. Thus, using Corollary 3.4.3, [58, Theorem 4.5] and (3.5.3), we have:

$$\mathbf{N}^\psi[L_\sigma^h \mathbf{1}_{\{\Delta<\delta\}}] = \mathbf{N}^{\psi_\delta}\left[L_\sigma^h \mathrm{e}^{-\bar{\pi}(\delta)\sigma}\right] = \mathrm{e}^{-h\psi_\delta'(\mathbf{N}^\psi[\Delta>\delta])} = \mathrm{e}^{-h/w(\delta)}. \tag{3.6.6}$$

In particular, the following identity relating the measures $\mathbf{P}_\delta^\psi$ and $\mathbf{P}_{\delta,h}^\psi$ holds:

$$\frac{1}{w(\delta)}\int_0^\infty \mathrm{d}h\, \mathrm{e}^{-h/w(\delta)}\, \mathbf{P}_{\delta,h}^\psi(\mathrm{d}s, \mathrm{d}\rho) = \mathbf{P}_\delta^\psi(\mathrm{d}s, \mathrm{d}\rho),$$

where we used that $\mathbf{1}_{[0,\sigma]}(s)\,\mathrm{d}s = \int_0^a \mathrm{d}a\, L^a(\mathrm{d}s)$, see Section 3.2.5. The next lemma gives an approximation of the measure $\mathbf{P}_{\delta,h}^\psi$.

Lemma 3.6.2. *Let $F\colon \mathbb{R}_+ \times \mathscr{D} \to \mathbb{R}$ be measurable and bounded. We have for every $\delta, h > 0$:*

$$\lim_{\varepsilon\to 0}\frac{1}{2\varepsilon}\mathbf{N}^\psi\left[\int_0^\sigma F(s,\rho)\mathbf{1}_{\{h-\varepsilon<H_s<h+\varepsilon\}}\,\mathrm{d}s\,\mathbf{1}_{\{\Delta\le\delta-\varepsilon\}}\right] = \mathbf{N}^\psi\left[\int_0^\sigma F(s,\rho)\,L^h(\mathrm{d}s)\,\mathbf{1}_{\{\Delta<\delta\}}\right]. \tag{3.6.7}$$

Proof. Recall from Section 3.2.5 that the measure L^a is supported on the set $\{s \in [0,\sigma]\colon\ H_s = a\}$. Thus, we have:

$$\frac{1}{2\varepsilon}\int_0^\sigma F(s,\rho)\mathbf{1}_{\{h-\varepsilon<H_s<h+\varepsilon\}}\,\mathrm{d}s = \frac{1}{2\varepsilon}\int_{h-\varepsilon}^{h+\varepsilon}\mathrm{d}a\int_0^\sigma F(s,\rho)\,L^a(\mathrm{d}s). \tag{3.6.8}$$

Furthermore, h is a jump time for the local time process $a \mapsto L^a$ if and only if it is a jump time for the total mass process $a \mapsto L_\sigma^a$. But, under $\mathbf{N}^\psi$, the process $(L_\sigma^a,\ a \ge 0)$ is a ψ-CB process. In particular, it has no fixed jump times. As a result, $\mathbf{N}^\psi$-a.e. the mapping $a \mapsto L^a$ is continuous at h. We deduce that the following convergence holds $\mathbf{N}^\psi$-a.e.:

$$\lim_{\varepsilon\to 0}\frac{1}{2\varepsilon}\int_0^\sigma F(s,\rho)\mathbf{1}_{\{h-\varepsilon<H_s<h+\varepsilon\}}\,\mathrm{d}s\,\mathbf{1}_{\{\Delta\le\delta-\varepsilon\}} = \int_0^\sigma F(s,\rho)\,L^h(\mathrm{d}s)\,\mathbf{1}_{\{\Delta<\delta\}}.$$

Next, using (3.6.8), we have:

$$\frac{1}{2\varepsilon}\left|\int_0^\sigma F(s,\rho)\mathbf{1}_{\{h-\varepsilon<H_s<h+\varepsilon\}}\,\mathrm{d}s\right|\mathbf{1}_{\{\Delta\le\delta-\varepsilon\}} \le \frac{\|F\|_\infty}{2\varepsilon}\int_{h-\varepsilon}^{h+\varepsilon} L_\sigma^a\,\mathrm{d}a,$$

where the last term converges $\mathbf{N}^\psi$-a.e. to $\|F\|_\infty L_\sigma^h$ thanks to the continuity of $a \mapsto L_\sigma^a$ at h. Furthermore, by [58, Eq. (12)] we have the convergence:

$$\lim_{\varepsilon\to 0}\frac{1}{2\varepsilon}\mathbf{N}^\psi\left[\int_{h-\varepsilon}^{h+\varepsilon} L_\sigma^a\,\mathrm{d}a\right] = \mathbf{N}^\psi[L_\sigma^h].$$

Thus, the generalized dominated convergence theorem yields (3.6.7). □

The main result of this section is the following description of the exploration process conditioned on having maximal degree δ at height h.

Theorem 3.6.3. *Assume that the Lévy measure π is diffuse. There exists a conditional probability measure $\mathbf{N}^{\psi}[\cdot|\Delta=\delta, H_\Delta=h]$ for $\delta\in\operatorname{supp}(\pi)$. Furthermore, for every $F\in\mathscr{B}_+(\mathscr{D})$, we have:*

$$\mathbf{N}^{\psi}[F(\rho)|\Delta=\delta,H_\Delta=h]=\int_{\mathbb{R}_+\times\mathscr{D}}\mathbf{P}^{\psi}_{\delta,h}(\mathrm{d}s,\mathrm{d}\tilde{\rho})\int_{\mathscr{D}}\mathbb{P}^{\psi}_{\delta}(\mathrm{d}\hat{\rho}|\Delta\le\delta)F(\tilde{\rho}\circledast(s,\hat{\rho})). \tag{3.6.9}$$

Assuming the Grey condition, this can be interpreted as follows in terms of trees. Under $\mathbf{N}^{\psi}$, conditionally on $\Delta=\delta$, H_Δ is exponentially distributed with mean $w(\delta)$. Moreover, conditionally on $\Delta=\delta$ and $H_\Delta=h$, the Lévy tree can be constructed as follows: start with $\widetilde{\mathscr{T}}$ with distribution $\mathbf{N}^{\psi}[L^h_\sigma\mathbf{1}_{\{\Delta<\delta\}}]^{-1}\mathbf{N}^{\psi_\delta}[L^h_\sigma\mathbf{1}_{\{\Delta<\delta\}}\,\mathrm{d}\mathscr{T}]$, choose a leaf uniformly at random in $\widetilde{\mathscr{T}}$ at height h (i.e. according to the probability measure $L^h(\mathrm{d}x)/L^h_\sigma$) and on this leaf graft an independent Lévy forest with initial mass δ conditioned to have degree $\le\delta$. Notice that this result generalizes Theorem 3.5.7 when the Lévy measure π is diffuse. In particular, one can recover the latter simply by integrating with respect to h.

Proof. Let $\delta\in\operatorname{supp}(\pi)$ and $h>0$. Thanks to Lemma 3.6.1, we have as $\varepsilon\to 0$:

$$\mathbf{N}^{\psi}\left[F(\rho)\middle|\delta-\varepsilon<\Delta<\delta+\varepsilon,h-\varepsilon<H_\Delta<h+\varepsilon\right]\sim\mathbf{N}^{\psi}\left[F(\rho)\middle|F_{\delta,\varepsilon}\right]. \tag{3.6.10}$$

Recall from (3.5.14) and (3.5.15) the definitions of $\rho^{\delta-\varepsilon,-}$ and $\rho^{\delta-\varepsilon,+}$. Using the Poissonian decomposition from Theorem 3.4.1 and Corollary 3.4.3, we have:

$$\begin{aligned}\mathbf{N}^{\psi}\left[F(\rho)\mathbf{1}_{F_{\delta,\varepsilon}}\right]&=\mathbf{N}^{\psi}\left[F\left(\rho^{\delta-\varepsilon,-}\circledast(T_{\delta-\varepsilon},\rho^{\delta-\varepsilon,+})\right)\mathbf{1}_{F_{\delta,\varepsilon}}\right]\\&=\mathbf{N}^{\psi_{\delta-\varepsilon}}\left[\bar{\pi}(\delta-\varepsilon)\mathrm{e}^{-\bar{\pi}(\delta-\varepsilon)\sigma}\int\mathbf{1}_{[0,\sigma]}(s)\,\mathrm{d}s\,\mathbb{Q}^{\psi}_{\delta-\varepsilon}(\mathbf{1}_{\{\Delta<\delta+\varepsilon\}}\,\mathrm{d}\hat{\rho})F(\rho\circledast(s,\hat{\rho}))\mathbf{1}_{\{h-\varepsilon<H_s<h+\varepsilon\}}\right]\\&=\bar{\pi}(\delta-\varepsilon)\mathbf{N}^{\psi}\left[\int\mathbf{1}_{[0,\sigma]}(s)\,\mathrm{d}s\,\mathbb{Q}^{\psi}_{\delta-\varepsilon}(\mathbf{1}_{\{\Delta<\delta+\varepsilon\}}\,\mathrm{d}\hat{\rho})F(\rho\circledast(s,\hat{\rho}))\mathbf{1}_{\{h-\varepsilon<H_s<h+\varepsilon,\Delta\le\delta-\varepsilon\}}\right].\end{aligned}$$

By conditioning, it follows from (3.6.3) that

$$\begin{aligned}\mathbf{N}^{\psi}\left[F(\rho)\middle|F_{\delta,\varepsilon}\right]&=\frac{1}{\mathbf{N}^{\psi}\left[\int_0^\sigma\mathbf{1}_{\{h-\varepsilon<H_s<h+\varepsilon\}}\,\mathrm{d}s\,\mathbf{1}_{\{\Delta\le\delta-\varepsilon\}}\right]}\\&\quad\times\mathbf{N}^{\psi}\left[\int\mathbf{1}_{[0,\sigma]}(s)\,\mathrm{d}s\,\mathbb{Q}^{\psi}_{\delta-\varepsilon}(\mathrm{d}\hat{\rho}|\Delta<\delta+\varepsilon)F(\rho\circledast(s,\hat{\rho}))\mathbf{1}_{\{h-\varepsilon<H_s<h+\varepsilon,\Delta\le\delta-\varepsilon\}}\right].\end{aligned} \tag{3.6.11}$$

Therefore, using Lemma 3.6.2, Lemma 3.5.4 and Lemma 3.5.5, we deduce that

$$\begin{aligned}&\lim_{\varepsilon\to0}\mathbf{N}^{\psi}\left[F(\rho)\middle|\delta-\varepsilon<\Delta<\delta+\varepsilon,h-\varepsilon<H_\Delta<h+\varepsilon\right]\\&\qquad=\int_{\mathbb{R}_+\times\mathscr{D}}F\,\mathrm{d}\mathbf{P}^{\psi}_{\delta,h}\times\mathbb{P}^{\psi}_{\delta}(G(\rho)|\Delta\le\delta),\end{aligned} \tag{3.6.12}$$

and the result readily follows by using [68, Theorem 1.30]. □

3.7 Local limit of the Lévy tree conditioned on large maximal degree

In this section, we shall investigate the behavior of the exploration process conditionally on $\Delta = \delta$ as $\delta \to \infty$. We start with the subcritical case. Then recall from (3.2.13) that $\mathbf{N}^\psi[\sigma] = \alpha^{-1} < \infty$. We define a probability measure $\mathbf{P}_\infty^\psi$ on the space $\mathbb{R}_+ \times \mathscr{D}$ by setting:

$$\int_{\mathbb{R}_+ \times \mathscr{D}} F \, \mathrm{d}\mathbf{P}_\infty^\psi = \alpha \, \mathbf{N}^\psi \left[\int_0^\sigma F(s, \rho) \, \mathrm{d}s \right], \tag{3.7.1}$$

for every $F \in \mathscr{B}_+(\mathbb{R}_+ \times \mathscr{D})$.

Lemma 3.7.1. *Assume that ψ is subcritical. The probability measure $\mathbf{P}_\delta^\psi$ converges to $\mathbf{P}_\infty^\psi$ in total variation distance on the space $\mathbb{R}_+ \times \mathscr{D}$ as $\delta \to \infty$.*

Proof. Let $F : \mathbb{R}_+ \times \mathscr{D}$ be measurable and bounded. We have:

$$\left| \mathbf{N}^\psi \left[\int_0^\sigma F(s, \rho) \, \mathrm{d}s \, \mathbf{1}_{\{\Delta < \delta\}} \right] - \mathbf{N}^\psi \left[\int_0^\sigma F(s, \rho) \, \mathrm{d}s \right] \right| \le \|F\|_\infty \mathbf{N}^\psi[\sigma \mathbf{1}_{\{\Delta \ge \delta\}}].$$

Since ψ is subcritical, we have $\mathbf{N}^\psi[\sigma] < \infty$ and the right-hand side converges to 0 as $\delta \to \infty$. This proves the result. □

For every measure-valued process $\mu = (\mu_t, t \ge 0) \in \mathscr{D}$, we define the measure-valued process $R_0(\mu)$ obtained from μ by removing any atoms at 0:

$$R_0(\mu)_t = \mu_t - \mu_t(0)\delta_0. \tag{3.7.2}$$

Denote by $\mathbb{P}^\psi$ the distribution of the exploration process ρ with branching mechanism ψ starting from 0.

Lemma 3.7.2. *Assume that ψ is subcritical and that π is unbounded. Under $\mathbb{P}_\delta^\psi(\cdot | \Delta \le \delta)$, the process $R_0(\rho)$ converges in distribution to $\mathbb{P}^\psi$ in the space $(\mathscr{D}, d_S)$ as $\delta \to \infty$.*

Proof. Recall from (3.5.8) that $\mathbb{P}_\delta^\psi(\Delta \le \delta) = \mathrm{e}^{-\delta \mathbf{N}^\psi[\Delta > \delta]}$. Since ψ is subcritical, by [87, Proposition 3.8], we have as $\delta \to \infty$:

$$\mathbf{N}^\psi[\Delta > \delta] \sim \frac{\bar{\pi}(\delta)}{\alpha}. \tag{3.7.3}$$

But $\delta \bar{\pi}(\delta) \le \int_{(\delta, \infty)} r \, \pi(\mathrm{d}r)$ and the last term goes to 0 as $\delta \to \infty$. It follows that $\lim_{\delta \to \infty} \delta \mathbf{N}^\psi[\Delta > \delta] = 0$ and

$$\lim_{\delta \to \infty} \mathbb{P}_\delta^\psi(\Delta \le \delta) = 1. \tag{3.7.4}$$

Thus, it suffices to show that for every continuous and bounded function $F: \mathscr{D} \to \mathbb{R}$, the following convergence holds:

$$\lim_{\delta\to\infty} \mathbb{P}^{\psi}_{\delta}(F \circ R_0(\rho)) = \mathbb{P}^{\psi}(F(\rho)). \tag{3.7.5}$$

Let ρ be the exploration process with branching mechanism ψ starting from 0, that is ρ has distribution $\mathbb{P}^{\psi}$. Then, the process $\tilde{\rho}^{(\delta)}$ defined in (3.2.20) has distribution $\mathbb{P}^{\psi}_{\delta}$. Notice that we have $R_0(\tilde{\rho}^{(\delta)})_t = \rho_t \mathbf{1}_{\{L^0_t \le r\}}$, which implies that

$$d_S(R_0(\tilde{\rho}^{(\delta)}), \rho) \le \sup_{t\ge0} d_{\mathrm{BL}}(R_0(\tilde{\rho}^{(\delta)})_t, \rho_t) = \sup_{L^0_t>\delta} \langle \rho_t, 1\rangle.$$

Recall that $\langle \rho_t, 1\rangle = X_t - I_t$. Since the Lévy measure π satisfies the integrability assumption $\int_{(0,\infty)} (r \wedge r^2)\, \pi(\mathrm{d}r) < \infty$, the process X does not drift to ∞; see e.g. [29, Chapter VII]. This implies that the following convergence holds a.s.:

$$\lim_{\delta\to\infty} \sup_{L^0_t>\delta} (X_t - I_t) = 0.$$

Therefore, the process $R_0(\tilde{\rho}^{(\delta)})$ converges a.s. to ρ for the Skorokhod topology. This proves (3.7.5) and the proof is complete. □

Remark 3.7.3. It should be clear from (3.2.20) that the mass $\tilde{\rho}^{(\delta)}_0(0)$ of the atom at 0 goes to ∞ as $\delta \to \infty$. This corresponds to the condensation phenomenon: a node with infinite mass appears at the limit. By introducing the operator R_0, we remove this mass which allows us to study the limiting behavior above the condensation node.

Similarly to what was done in Section 3.5 (see (3.5.14) and (3.5.15)), we split the path of the exploration process into two parts around the first node with mass Δ: $\rho^{\Delta,-}$ is the pruned exploration process (that is the exploration process minus the first node with mass Δ) and $\rho^{\Delta,+}$ is the path of the exploration process above the first node with mass Δ. Notice that $\rho^{\Delta,+}_0$ is equal to Δ times the Dirac measure at 0. Let

$$E_\delta = \{\Delta = \delta, \Delta(\rho^{\Delta,-}) < \delta\} \tag{3.7.6}$$

be the event that the maximal degree is equal to δ and there is a unique first-generation node with mass δ. Recall from (3.5.23) the definition of $\mathfrak{g}$.

Lemma 3.7.4. *Assume that ψ is subcritical and that the set of atoms of the Lévy measure π is unbounded. The following holds as $\delta \to \infty$ along the set of atoms of π:*

$$\mathbf{N}^{\psi}[\Delta = \delta] \sim \mathbf{N}^{\psi}[E_\delta] \sim \frac{\mathfrak{g}(\delta)}{\alpha}. \tag{3.7.7}$$

Proof. Under $\mathbf{N}^{\psi_{\delta\text{-}}}$ and conditionally on ρ, let $\sum_{i=1}^{N} \delta_{(s_i,\rho_i)}$ be a Poisson point measure with intensity $\mathrm{d}s \int_{[\delta,\infty)} \pi(\mathrm{d}r)\,\mathbb{P}_r^{\psi}(\mathrm{d}\tilde{\rho})$. Thanks to Theorem 3.4.1, we have:

$$\mathbf{N}^{\psi}[E_\delta] = \mathbf{N}^{\psi_{\delta\text{-}}}\left[N = 1, \Delta(\rho_1) \le \delta\right].$$

But, under $\mathbf{N}^{\psi_{\delta\text{-}}}$ and conditionally on ρ, N has Poisson distribution with parameter $\pi[\delta,\infty)\sigma$, ρ_1 has distribution $\pi[\delta,\infty)^{-1} \int_{[\delta,\infty)} \pi(\mathrm{d}r)\,\mathbb{P}_r^{\psi}(\mathrm{d}\tilde{\rho})$ and they are independent. It follows that

$$\begin{aligned}\mathbf{N}^{\psi}[E_\delta] &= \mathbf{N}^{\psi_{\delta\text{-}}}\left[\sigma \mathrm{e}^{-\pi[\delta,\infty)\sigma} \int_{[\delta,\infty)} \pi(\mathrm{d}r)\,\mathbb{P}_r^{\psi}(\Delta \le \delta)\right] \\ &= \mathbf{N}^{\psi_{\delta\text{-}}}\left[\sigma \mathrm{e}^{-\pi[\delta,\infty)\sigma} \pi(\delta) \mathrm{e}^{-\delta \mathbf{N}^{\psi}[\Delta > \delta]}\right] \\ &= \mathfrak{g}(\delta)\, w(\delta), \end{aligned} \tag{3.7.8}$$

where we used (3.5.8) for the second equality and (3.5.2) for the last.

Recall from (3.5.1) the definition of w. Since ψ is subcritical, it follows from (3.2.13) that $\lim_{\delta\to\infty} w(\delta) = \alpha^{-1}$. This proves that

$$\mathbf{N}^{\psi}[E_\delta] \sim \frac{\mathfrak{g}(\delta)}{\alpha}.$$

A similar computation yields:

$$\begin{aligned}\mathbf{N}^{\psi}[\Delta = \delta] &= \mathbf{N}^{\psi_{\delta\text{-}}}\left[N \ge 1, \Delta(\rho_i) \le \delta, \forall 1 \le i \le N\right] \\ &= \mathbf{N}^{\psi_{\delta\text{-}}}\left[\mathrm{e}^{-\pi[\delta,\infty)\sigma}\left(\mathrm{e}^{\mathfrak{g}(\delta)\sigma} - 1\right)\right] \\ &= \mathbf{N}^{\psi}\left[\left(\mathrm{e}^{\mathfrak{g}(\delta)\sigma} - 1\right)\mathbf{1}_{\{\Delta<\delta\}}\right], \end{aligned} \tag{3.7.9}$$

where we used Corollary 3.4.3 for the last equality.

Observe that since $\pi(1,\infty) < \infty$, $\pi(\delta)$ (and thus also $\mathfrak{g}(\delta)$) converges to 0 as $\delta \to \infty$. It is clear that

$$\lim_{\delta\to\infty} \frac{\mathrm{e}^{\mathfrak{g}(\delta)\sigma} - 1}{\mathfrak{g}(\delta)} \mathbf{1}_{\{\Delta<\delta\}} = \sigma.$$

Furthermore, since ψ is subcritical, there exists $\lambda_0 > 0$ such that $\mathbf{N}^{\psi}\left[\sigma \mathrm{e}^{\lambda_0 \sigma}\right] < \infty$. Thus, using Taylor's inequality, we have for $\delta > 0$ large enough:

$$\frac{\mathrm{e}^{\mathfrak{g}(\delta)\sigma} - 1}{\mathfrak{g}(\delta)} \mathbf{1}_{\{\Delta<\delta\}} \le \sigma \mathrm{e}^{\mathfrak{g}(\delta)\sigma} \le \sigma \mathrm{e}^{\lambda_0\sigma},$$

Thanks to the dominated convergence theorem, we deduce that

$$\lim_{\delta\to\infty} \frac{\mathbf{N}^{\psi}[\Delta = \delta]}{\mathfrak{g}(\delta)} = \lim_{\delta\to\infty} \frac{1}{\mathfrak{g}(\delta)} \mathbf{N}^{\psi}\left[\left(\mathrm{e}^{\mathfrak{g}(\delta)\sigma} - 1\right)\mathbf{1}_{\{\Delta<\delta\}}\right] = \mathbf{N}^{\psi}[\sigma] = \alpha^{-1}.$$

This finishes the proof. □

The first main result of this section concerns the limit of the subcritical Lévy tree conditioned on having a large maximal degree. Then there is a condensation phenomenon: the limit consists of a size-biased Lévy tree onto which one grafts – at a uniformly chosen leaf – an independent Lévy forest with infinite mass. In particular, the height of the condensation node is exponentially distributed. Recall from (3.5.30) that T_Δ is the first time that the exploration process contains an atom with mass Δ. Recall also that $\rho^{\Delta,-}$ denotes the path of the exploration process after removing the first node with mass Δ while $\rho^{\Delta,+}$ denotes the path of the exploration process above that node. Finally, recall that $\mathbb{P}^\psi$ is the distribution of the exploration process with branching mechanism ψ starting from 0.

Theorem 3.7.5. *Assume that ψ is subcritical and that π is unbounded. Let $F\colon \mathbb{R}_+ \times \mathscr{D} \to \mathbb{R}$ and $G\colon \mathscr{D} \to \mathbb{R}$ be continuous and bounded. We have:*

$$\lim_{\delta\to\infty} \mathbf{N}^\psi \left[F(T_\Delta, \rho^{\Delta,-}) G\circ R_0(\rho^{\Delta,+}) | \Delta = \delta\right] = \alpha \mathbf{N}^\psi \left[\int_0^\sigma F(s,\rho)\,\mathrm{d}s\right] \mathbb{P}^\psi(G(\rho)). \tag{3.7.10}$$

Proof. When $\delta \to \infty$ along the set of non-atoms $\{\delta > 0\colon \pi(\delta) = 0\}$, the convergence is a direct consequence of Theorem 3.5.7, Lemma 3.7.1 and Lemma 3.7.2.

Now assume that $\delta > 0$ is an atom of π. Thanks to Lemma 3.7.4 and since the inclusion $E_\delta \subset \{\Delta = \delta\}$ holds, it is enough to show that the result holds when conditionning by E_δ. But, thanks to Remark 3.5.8, we have:

$$\mathbf{N}^\psi \left[F(T_\Delta, \rho^{\Delta,-}) G\circ R_0(\rho^{\Delta,+}) \middle| E_\delta\right] = \int_{\mathbb{R}_+\times\mathscr{D}} F(s,\tilde{\rho})\, \mathbf{P}^\psi_\delta(\mathrm{d}s, \mathrm{d}\tilde{\rho})\, \mathbb{P}^\psi_\delta(G\circ R_0(\rho) | \Delta \le \delta).$$

The result readily follows from Lemma 3.7.1 and Lemma 3.7.2. □

Next, we consider the critical case. Recall from (3.5.1) the definition of w. The next lemma is a key ingredient in the proof of the local convergence of the critical Lévy tree.

Lemma 3.7.6. *Assume that ψ is critical and that the Lévy measure π is unbounded. For every $h > 0$, we have*

$$\lim_{\delta\to\infty} \frac{1}{w(\delta)} \mathbf{N}^\psi \left[\sigma F(r_h(\rho)) \mathbf{1}_{\{\Delta<\delta\}}\right] = \mathbf{N}^\psi \left[L^h_\sigma F(r_h(\rho))\right]. \tag{3.7.11}$$

Proof. We shall use the decomposition of the exploration process above level h, see Section 3.2.5. Let $(\rho^i, i \in I_h)$ be the excursions of the exploration process above level h. For every $i \in I_h$, let σ^i (resp. Δ^i) be the lifetime (resp. the maximal degree) of ρ^i. Similarly, denote by σ_h (resp. Δ_h) the lifetime (resp. the maximal degree) of $r_h(\rho)$. Thanks to Proposition 3.2.4, we have:

$$\begin{aligned}\mathbf{N}^\psi \left[\sigma F(r_h(\rho)) \mathbf{1}_{\{\Delta<\delta\}}\right] &= \mathbf{N}^\psi \left[\left(\sigma_h + \sum_{i\in I_h} \sigma^i\right) F(r_h(\rho)); \Delta_h < \delta, \Delta^i < \delta, \forall i \in I_h\right] \\ &= \mathbf{N}^\psi \left[F(r_h(\rho)) \mathbf{1}_{\{\Delta_h<\delta\}} \mathbf{N}^\psi \left[\sigma_h + \sum_{i\in I_h} \sigma^i; \Delta^i < \delta, \forall i \in I_h \middle| \mathscr{E}_h\right]\right].\end{aligned}$$

Thanks to the Mecke formula for Poisson random measures, see e.g. [115, Chapter 4, Theorem 4.1], we get:

$$\mathbf{N}^{\psi}\left[\sigma_h + \sum_{i\in I_h} \sigma^i ; \Delta^i < \delta, \forall i \in I_h \middle| \mathscr{E}_h\right] = \left(\sigma_h + L_\sigma^h \mathbf{N}^{\psi}[\sigma \mathbf{1}_{\{\Delta<\delta\}}]\right) \mathrm{e}^{-L_\sigma^h \mathbf{N}^{\psi}[\Delta\geq\delta]}.$$

We deduce that

$$\frac{1}{w(\delta)} \mathbf{N}^{\psi}\left[\sigma F(r_h(\rho))\mathbf{1}_{\{\Delta<\delta\}}\right] = \mathbf{N}^{\psi}\left[F(r_h(\rho))\mathbf{1}_{\{\Delta_h<\delta\}}\left(L_\sigma^h + w(\delta)^{-1}\sigma_h\right)\mathrm{e}^{-L_\sigma^h \mathbf{N}^{\psi}[\Delta\geq\delta]}\right]. \tag{3.7.12}$$

Notice that $w(\delta)\to\infty$ as $\delta\to\infty$ since ψ is critical. Furthermore, it is clear that $\sigma_h = \int_0^h L_\sigma^a \,\mathrm{d}a$. Now letting $\delta\to\infty$ in (3.6.6) gives that $\mathbf{N}^{\psi}[L_\sigma^a] = 1$. It follows that $\mathbf{N}^{\psi}[\sigma_h] = h < \infty$. Thus, the dominated convergence theorem applies and we obtain the desired result by letting $\delta\to\infty$ in (3.7.12). □

Recall from Theorem 3.5.7 that when the Lévy measure π has an atom $\delta > 0$, the exploration process conditioned on $\Delta = \delta$ has a random number of first-generation nodes with mass δ. The next lemma gives a sufficient condition for there to be exactly one with high probability as $\delta\to\infty$. Recall from (3.5.1) the definition of w. Recall also from (3.7.6) that E_δ denotes the event that the maximal degree is equal to δ and there is a unique first-generation node with mass δ.

Lemma 3.7.7. *Assume that ψ is critical and that the Lévy measure π is unbounded. Furthermore, assume that*

$$\lim_{\delta\to\infty} \frac{\pi(\delta)}{w(\delta)\bar{\pi}(\delta)\int_{[\delta,\infty)} r\,\pi(\mathrm{d}r)} = 0. \tag{3.7.13}$$

We have as $\delta\to\infty$ along the set of atoms of π:

$$\mathbf{N}^{\psi}[\Delta=\delta] \sim \mathbf{N}^{\psi}[E_\delta] = \mathfrak{g}(\delta)\,w(\delta). \tag{3.7.14}$$

Proof. Recall from (3.7.8) and (3.7.9) that

$$\mathbf{N}^{\psi}[E_\delta] = \mathfrak{g}(\delta)\,w(\delta) \quad \text{and} \quad \mathbf{N}^{\psi}[\Delta=\delta] = \mathbf{N}^{\psi}\left[\left(\mathrm{e}^{\mathfrak{g}(\delta)\sigma} - 1\right)\mathbf{1}_{\{\Delta<\delta\}}\right].$$

Using Taylor's inequality, we deduce that

$$1 \leq \frac{\mathbf{N}^{\psi}[\Delta=\delta]}{\mathbf{N}^{\psi}[E_\delta]} \leq \frac{w_1(\delta)}{w(\delta)}, \tag{3.7.15}$$

where we set $w_1(\delta) = \mathbf{N}^{\psi}\left[\sigma \mathrm{e}^{\mathfrak{g}(\delta)\sigma}\mathbf{1}_{\{\Delta<\delta\}}\right]$.

Using (3.5.2) and the inequality $\mathrm{e}^{-x} \geq 1-x$ for every $x\geq 0$, we have:

$$w(\delta) = \mathbf{N}^{\psi_{\delta-}}\left[\sigma \mathrm{e}^{-\pi[\delta,\infty)\sigma}\right] \geq \mathbf{N}^{\psi_{\delta-}}\left[\sigma(1-\pi(\delta)\sigma)\mathrm{e}^{-\bar{\pi}(\delta)\sigma}\right].$$

But thanks to Corollary 3.4.3, observe that

$$w_1(\delta) = \mathbf{N}^{\psi_{\delta-}}\left[\sigma \mathrm{e}^{(\mathfrak{g}(\delta)-\pi[\delta,\infty))\sigma}\right] \le \mathbf{N}^{\psi_{\delta-}}\left[\sigma \mathrm{e}^{-\bar{\pi}(\delta)\sigma}\right],$$

where we used that $\mathfrak{g}(\delta) \le \pi(\delta)$ for the inequality. Furthermore, using that the function $x \mapsto x\mathrm{e}^{-x}$ is bounded on $\mathbb{R}_+$ by some constant $M > 0$, we have:

$$\mathbf{N}^{\psi_{\delta-}}\left[\sigma^2 \mathrm{e}^{-\bar{\pi}(\delta)\sigma}\right] \le \frac{M}{\bar{\pi}(\delta)} \mathbf{N}^{\psi_{\delta-}}[\sigma] = \frac{M}{\bar{\pi}(\delta)\int_{[\delta,\infty)} r\,\pi(\mathrm{d}r)}.$$

We deduce that

$$w(\delta) \ge w_1(\delta) - \frac{M\pi(\delta)}{\bar{\pi}(\delta)\int_{[\delta,\infty)} r\,\pi(\mathrm{d}r)}.$$

It follows from (3.7.15) that

$$1 \le \frac{\mathbf{N}^{\psi}[\Delta = \delta]}{\mathbf{N}^{\psi}[E_\delta]} \le 1 + \frac{M\pi(\delta)}{w(\delta)\bar{\pi}(\delta)\int_{[\delta,\infty)} r\,\pi(\mathrm{d}r)},$$

and the result readily follows by using (3.7.13). □

In the critical case, the Lévy tree conditioned on having a large maximal degree converges locally to the immortal Lévy tree. Intuitively, the condensation node goes to infinity and thus becomes invisible to local convergence.

Theorem 3.7.8. *Assume that ψ is critical and that π is unbounded. Furthermore, assume that (3.7.13) holds. Let $F\colon \mathscr{D} \to \mathbb{R}$ be continuous and bounded. For every $h > 0$, we have:*

$$\lim_{\delta\to\infty} \mathbf{N}^{\psi}\left[F(r_h(\rho))\middle|\Delta = \delta\right] = \mathbf{N}^{\psi}\left[L_\sigma^h F(r_h(\rho))\right]. \tag{3.7.16}$$

Proof. First assume that $\delta > 0$ is not an atom of π. Thanks to Theorem 3.5.7, conditionally on $\Delta = \delta$, ρ is distributed as $\tilde{\rho} \circledast (s, \hat{\rho})$, where $(s, \tilde{\rho})$ has distribution $\mathbf{P}_\delta^\psi$, $\hat{\rho}$ has distribution $\mathbb{P}_\delta^\psi(\cdot|\Delta \le \delta)$ and they are independent.

Next, assume that $\delta > 0$ is an atom of π. Recall that E_δ denotes the event that $\Delta = \delta$ and there is a unique first-generation node with mass δ. Thanks to Lemma 3.7.7, since $E_\delta \subset \{\Delta = \delta\}$, the two conditionings are equivalent and it is enough to show that the result holds when conditioning on E_δ. But Remark 3.5.8 gives that, conditionally on E_δ, ρ is again distributed as $\tilde{\rho} \circledast (s, \hat{\rho})$.

Thus, in all cases, it is enough to show that

$$\lim_{\delta\to\infty} \int_{\mathbb{R}_+\times\mathscr{D}} \mathbf{P}_\delta^\psi(\mathrm{d}s, \mathrm{d}\tilde{\rho}) \int_{\mathscr{D}} \mathbb{P}_\delta^\psi(\mathrm{d}\hat{\rho}|\Delta \le \delta) F(r_h(\tilde{\rho} \circledast (s, \hat{\rho}))) = \mathbf{N}^{\psi}\left[L_\sigma^h F(r_h(\rho))\right].$$

Now, Lemma 3.5.10 gives that the height $H(\tilde{\rho}_s)$ at which $\hat{\rho}$ is grafted is exponentially distributed with mean $w(\delta)$. Since ψ is critical, it holds that $\lim_{\delta\to\infty} w(\delta) = \infty$. Thus, we deduce that

$H(\tilde{\rho}_s) > h$ with high probability as $\delta \to \infty$ under $\mathbf{P}^{\psi}_{\delta}(\mathrm{d}s, \mathrm{d}\tilde{\rho})$, i.e. we have:

$$\lim_{\delta\to\infty} \int_{\mathbb{R}_+ \times \mathcal{D}} \mathbf{P}^{\psi}_{\delta}(\mathrm{d}s, \mathrm{d}\tilde{\rho}) \mathbf{1}_{\{H(\tilde{\rho}_s) \leq h\}} = 0.$$

Furthermore, on the event $\{H(\tilde{\rho}_s) > h\}$, it holds that $r_h(\tilde{\rho} \circledast (s, \hat{\rho})) = r_h(\tilde{\rho})$, and the proof reduces to showing the following convergence:

$$\lim_{\delta\to\infty} \int_{\mathbb{R}_+ \times \mathcal{D}} \mathbf{P}^{\psi}_{\delta}(\mathrm{d}s, \mathrm{d}\tilde{\rho}) F(r_h(\tilde{\rho})) = \mathbf{N}^{\psi}\left[L^h_{\sigma} F(r_h(\rho))\right]. \tag{3.7.17}$$

Recalling from (3.5.4) the definition of $\mathbf{P}^{\psi}_{\delta}$, Lemma 3.7.6 yields (3.7.17) and the proof is complete. □

We end this section with the following result dealing with the asymptotic behavior of the exploration process conditioned on having a large maximal degree at a fixed height h. Notice that this conditioning does not allow the condensation node to escape to infinity (even in the critical case as opposed to the conditioning of large maximal degree) and forces condensation to occur at a finite height. The limit consists of a Lévy tree biased by the population size at level h onto which one grafts – at a leaf chosen uniformly at random at height h – an independent Lévy forest with infinite mass.

Theorem 3.7.9. *Assume that ψ is (sub)critical and that π is unbounded and diffuse. Let $F: \mathbb{R}_+ \times \mathcal{D} \to \mathbb{R}$ and $G: \mathcal{D} \to \mathbb{R}$ be continuous and bounded. We have:*

$$\lim_{\delta\to\infty} \mathbf{N}^{\psi}\left[F(T_\Delta, \rho^{\Delta,-}) G(\rho^{\Delta,+}) | \Delta = \delta, H_\Delta = h\right] = \mathrm{e}^{\alpha h} \mathbf{N}^{\psi}\left[\int_0^{\sigma} F(s,\rho)\, L^h(\mathrm{d}s)\right] \mathbb{P}^{\psi}(G(\rho)). \tag{3.7.18}$$

Proof. Letting $\delta \to \infty$ in (3.6.6), we have that $\lim_{\delta\to\infty} \mathbf{N}^{\psi}[L^h_{\sigma} \mathbf{1}_{\{\Delta<\delta\}}] = \mathrm{e}^{-\alpha h}$. Furthermore, the dominated convergence theorem yields:

$$\lim_{\delta\to\infty} \mathbf{N}^{\psi}\left[\int_0^{\sigma} F(s,\rho)\, L^h(\mathrm{d}s)\, \mathbf{1}_{\{\Delta<\delta\}}\right] = \mathbf{N}^{\psi}\left[\int_0^{\sigma} F(s,\rho)\, L^h(\mathrm{d}s)\right].$$

This proves that the following convergence holds:

$$\lim_{\delta\to\infty} \int_{\mathbb{R}_+ \times \mathcal{D}} F \,\mathrm{d}\mathbf{P}^{\psi}_{\delta,h} = \mathrm{e}^{\alpha h} \mathbf{N}^{\psi}\left[\int_0^{\sigma} F(s,\rho)\, L^h(\mathrm{d}s)\right].$$

The result is then a direct consequence of Theorem 3.6.3 and Lemma 3.7.2. □

3.8 Other conditionings of large maximal degree

In this section, we look at other conditionings of large maximal degree. Recall from Section 3.4.2 that Z^{δ}_0 denotes the number of first-generation nodes with mass larger than δ while W^{δ} denotes the total number of nodes with mass larger than δ. Specifically, we study the conditionings $\Delta > \delta$ (which is equal to $Z_0^\delta \geq 1$ or $W^\delta \geq 1$), $Z_0^\delta = 1$ and $W^\delta = 1$. We shall see that, in the subcritical and critical cases, all three give rise to the same asymptotic behavior as conditioning by $\Delta = \delta$.

Notice that $\{W^\delta = 1\}$ (resp. $\{Z_0^\delta = 1\}$) is the event that ρ contains exactly one node (resp. one first-generation node) with mass larger than δ. To begin, we compute the measure of these two events. In the subcritical case, they are equivalent in $\mathbf{N}^\psi$-measure to $\{\Delta > \delta\}$. However, this is no longer the case for critical branching mechanisms, see Proposition 3.9.2 for the (critical) stable case.

Proposition 3.8.1. *We have:*

$$\mathbf{N}^\psi[Z_0^\delta = 1] = \frac{\bar{\pi}(\delta)}{\psi_\delta'(\mathbf{N}^\psi[\Delta > \delta])}, \tag{3.8.1}$$

$$\mathbf{N}^\psi[W^\delta = 1] = \frac{1}{\psi_\delta'(\mathbf{N}^\psi[\Delta > \delta])} \int_{(\delta,\infty)} \mathrm{e}^{-r\mathbf{N}^\psi[\Delta > \delta]} \pi(\mathrm{d}r). \tag{3.8.2}$$

In particular, assuming that ψ is subcritical and that π is unbounded, we have as $\delta \to \infty$:

$$\mathbf{N}^\psi[Z_0^\delta = 1] \sim \mathbf{N}^\psi[W^\delta = 1] \sim \mathbf{N}^\psi[\Delta > \delta] \sim \frac{\bar{\pi}(\delta)}{\alpha}. \tag{3.8.3}$$

Since we have the inclusions $\{W^\delta = 1\} \subset \{Z_0^\delta = 1\} \subset \{\Delta > \delta\}$, Proposition 3.8.1 entails that, in the subcritical case, the three conditionings are equivalent as $\delta \to \infty$. In particular, conditionally on $\Delta > \delta$, there is exactly one node with mass larger than δ with probability tending to 1 as $\delta \to \infty$.

Proof. Notice that $\{W^\delta = 1\}$ is the event that ρ contains only one first-generation node with mass larger than δ and that this node has no descendants with mass larger than δ. Thus, using the Poissonian decomposition of Theorem 3.4.1, we get $\mathbf{N}^\psi[Z_0^\delta = 1] = \mathbf{N}^{\psi_\delta}[\zeta = 1]$ and

$$\mathbf{N}^\psi\left[W^\delta = 1\right] = \mathbf{N}^{\psi_\delta}[\zeta = 1]\, \mathbb{Q}_\delta^\psi(\Delta \leq \delta). \tag{3.8.4}$$

Recall that under $\mathbf{N}^{\psi_\delta}$ and conditionally on ρ, ζ has Poisson distribution with parameter $\bar{\pi}(\delta)\sigma$. Thus we have

$$\mathbf{N}^{\psi_\delta}[\zeta = 1] = \mathbf{N}^{\psi_\delta}\left[\bar{\pi}(\delta)\sigma \mathrm{e}^{-\bar{\pi}(\delta)\sigma}\right] = \frac{\bar{\pi}(\delta)}{\psi_\delta' \circ \psi_\delta^{-1}(\bar{\pi}(\delta))} = \frac{\bar{\pi}(\delta)}{\psi_\delta'(\mathbf{N}^\psi[\Delta > \delta])}, \tag{3.8.5}$$

where we used (3.3.4) for the last equality. This proves (3.8.1).

Moreover, using the Poissonian decomposition of Proposition 3.4.4 together with the fact that, under $\mathbb{Q}_\delta^{\psi_\delta}$ and conditionally on ρ, ξ has Poisson distribution with parameter $\bar{\pi}(\delta)\sigma$, we get:

$$\mathbb{Q}_\delta^{\psi}(\Delta \le \delta) = \mathbb{Q}_\delta^{\psi_\delta}(\xi = 0) = \mathbb{Q}_\delta^{\psi_\delta}(\mathrm{e}^{-\bar{\pi}(\delta)\sigma}) = \frac{1}{\bar{\pi}(\delta)} \int_{(\delta,\infty)} \pi(\mathrm{d}r) \mathbb{P}_r^{\psi_\delta}(\mathrm{e}^{-\bar{\pi}(\delta)\sigma}).$$

Thus, it follows from (3.4.10) and (3.3.4) that

$$\mathbb{Q}_\delta^{\psi}(\Delta \le \delta) = \frac{1}{\bar{\pi}(\delta)} \int_{(\delta,\infty)} \mathrm{e}^{-r\psi_\delta^{-1}(\bar{\pi}(\delta))} \pi(\mathrm{d}r) = \frac{1}{\bar{\pi}(\delta)} \int_{(\delta,\infty)} \mathrm{e}^{-r\mathbf{N}^\psi[\Delta>\delta]} \pi(\mathrm{d}r). \tag{3.8.6}$$

Finally, combining (3.8.4), (3.8.5) and (3.8.6), we deduce (3.8.2).

Now assume that ψ is subcritical and that π is unbounded. Recall from (3.7.3) that

$$\mathbf{N}^\psi[\Delta > \delta] \sim \frac{\bar{\pi}(\delta)}{\alpha}.$$

On the other hand, differentiating (3.1.8), we get:

$$\psi_\delta'(\mathbf{N}^\psi[\Delta > \delta]) = \psi'(\mathbf{N}^\psi[\Delta > \delta]) + \int_{(\delta,\infty)} r\mathrm{e}^{-r\mathbf{N}^\psi[\Delta>\delta]} \pi(\mathrm{d}r).$$

Since $\int_{(1,\infty)} r\,\pi(\mathrm{d}r) < \infty$, the dominated convergence theorem shows that the last integral converges to 0 as $\delta \to \infty$. It follows that

$$\lim_{\delta\to\infty} \psi_\delta'(\mathbf{N}^\psi[\Delta > \delta]) = \psi'(0) = \alpha. \tag{3.8.7}$$

In particular, we get that $\mathbf{N}^\psi[Z_0^\delta = 1] \sim \alpha^{-1}\bar{\pi}(\delta)$.

Furthermore, we have:

$$\begin{aligned} 0 \le 1 - \frac{1}{\bar{\pi}(\delta)} \int_{(\delta,\infty)} \mathrm{e}^{-r\mathbf{N}^\psi[\Delta>\delta]} \pi(\mathrm{d}r) &= \frac{1}{\bar{\pi}(\delta)} \int_{(\delta,\infty)} \left(1 - \mathrm{e}^{-r\mathbf{N}^\psi[\Delta>\delta]}\right) \pi(\mathrm{d}r) \\ &\le \frac{\mathbf{N}^\psi[\Delta > \delta]}{\bar{\pi}(\delta)} \int_{(\delta,\infty)} r\,\pi(\mathrm{d}r). \end{aligned}$$

The dominated convergence theorem gives $\lim_{\delta\to\infty} \int_{(\delta,\infty)} r\,\pi(\mathrm{d}r) = 0$. Since $\lim_{\delta\to\infty} \mathbf{N}^\psi[\Delta > \delta]/\bar{\pi}(\delta) = \alpha^{-1}$, we deduce that

$$\lim_{\delta\to\infty} \frac{1}{\bar{\pi}(\delta)} \int_{(\delta,\infty)} \mathrm{e}^{-r\mathbf{N}^\psi[\Delta>\delta]} \pi(\mathrm{d}r) = 1.$$

Together with (3.8.2) and (3.8.7), this yields $\mathbf{N}^\psi[W^\delta = 1] \sim \alpha^{-1}\bar{\pi}(\delta)$. This concludes the proof. □

In the subcritical case, the three conditionings $\Delta > \delta$, $Z_0^\delta = 1$ and $W^\delta = 1$ are equivalent as $\delta \to \infty$ and thus they yield the same asymptotic behavior: a condensation phenomenon occurs at the limit just like in Theorem 3.7.5 where we condition by $\Delta = \delta$. Recall from (3.5.30) that T_Δ is the first time that the exploration process contains an atom with mass Δ. Recall also from Section 3.7 that $\rho^{\Delta,-}$ denotes the path of the exploration process after removing the first node with Δ while $\rho^{\Delta,+}$ denotes the path of the exploration process above that node.

Theorem 3.8.2. *Assume that ψ is subcritical and that π is unbounded. Let $F\colon \mathbb{R}_+ \times \mathscr{D} \to \mathbb{R}$ and $G\colon \mathscr{D} \to \mathbb{R}$ be continuous and bounded and let A_δ be equal to $\{\Delta > \delta\}$, $\{Z_0^\delta = 1\}$ or $\{W^\delta = 1\}$. We have:*

$$\lim_{\delta\to\infty} \mathbf{N}^\psi \left[F(T_\Delta, \rho^{\Delta,-}) G \circ R_0(\rho^{\Delta,+}) \middle| A_\delta\right] = \alpha \mathbf{N}^\psi \left[\int_0^\sigma F(s,\rho)\,\mathrm{d}s\right] \mathbb{P}^\psi(G(\rho)). \tag{3.8.8}$$

Proof. As the three events are equivalent it is enough to show the result for $A_\delta = \{\Delta > \delta\}$. Disintegrating with respect to Δ, we have:

$$\begin{aligned}
&\mathbf{N}^\psi \left[F(T_\Delta, \rho^{\Delta,-}) G \circ (\rho^{\Delta,+}) \middle| \Delta > \delta\right] \\
&\qquad = \frac{1}{\mathbf{N}^\psi[\Delta > \delta]} \int_{(\delta,\infty)} \mathbf{N}^\psi[\Delta \in \mathrm{d}r]\, \mathbf{N}^\psi \left[F(T_\Delta, \rho^{\Delta,-}) G \circ R_0(\rho^{\Delta,+}) \middle| \Delta = r\right].
\end{aligned}$$

The conclusion follows from Theorem 3.7.5. □

Recall from (3.5.12) that T_δ is the first time ρ contains a node with mass larger than δ. Also recall from (3.5.14) and (3.5.15) that $\rho^{\delta,-}$ denotes the path of the exploration process after removing the first node with mass larger than δ while $\rho^{\delta,+}$ denotes the path of the exploration process above that node. We shall determine the joint distribution of $(T_\delta, \rho^{\delta,-}, \rho^{\delta,+})$ conditionally on $Z_0^\delta = 1$ and $W^\delta = 1$. Recall from (3.5.1) the definition of w_+.

Lemma 3.8.3. *Assume that ψ is (sub)critical and let $F \in \mathscr{B}_+(\mathbb{R}_+ \times \mathscr{D})$ and $G \in \mathscr{B}_+(\mathscr{D})$. We have:*

$$\mathbf{N}^\psi \left[F(T_\delta, \rho^{\delta,-}) G(\rho^{\delta,+}) \middle| Z_0^\delta = 1\right] = \frac{1}{w_+(\delta)} \mathbf{N}^\psi \left[\int_0^\sigma F(s,\rho)\,\mathrm{d}s\, \mathbf{1}_{\{\Delta \le \delta\}}\right] \mathbb{Q}_\delta^\psi(G(\rho)), \tag{3.8.9}$$

$$\mathbf{N}^\psi \left[F(T_\delta, \rho^{\delta,-}) G(\rho^{\delta,+}) \middle| W^\delta = 1\right] = \frac{1}{w_+(\delta)} \mathbf{N}^\psi \left[\int_0^\sigma F(s,\rho)\,\mathrm{d}s\, \mathbf{1}_{\{\Delta \le \delta\}}\right] \mathbb{Q}_\delta^\psi(G(\rho) | \Delta \le \delta). \tag{3.8.10}$$

Proof. We only prove the first identity, the second one being similar. Theorem 3.4.1 gives:

$$\mathbf{N}^\psi \left[F(T_\delta, \rho^{\delta,-}) G(\rho^{\delta,+}) \mathbf{1}_{\{Z_0^\delta = 1\}}\right] = \mathbf{N}^{\psi_\delta} \left[F(U,\rho) G(\mathscr{F}^\delta) \mathbf{1}_{\{\zeta = 1\}}\right],$$

where under $\mathbf{N}^{\psi_\delta}$ and conditionally on ρ, ρ^δ has distribution $\mathbb{Q}_\delta^\psi$, U is uniformly distributed on $[0,\sigma]$, ζ has Poisson distribution with parameter $\bar\pi(\delta)\sigma$ and they are independent. Therefore, conditioning on ρ in the last term, we get:

$$\begin{aligned}
\mathbf{N}^\psi \left[F(T_\delta, \rho^{\delta,-}) G(\rho^{\delta,+}) \mathbf{1}_{\{Z_0^\delta = 1\}}\right] &= \bar\pi(\delta) \mathbf{N}^{\psi_\delta} \left[\mathrm{e}^{-\bar\pi(\delta)\sigma} \int_0^\sigma F(s,\rho)\,\mathrm{d}s\right] \mathbb{Q}_\delta^\psi(G(\rho)) \\
&= \bar\pi(\delta) \mathbf{N}^\psi \left[\int_0^\sigma F(s,\rho)\,\mathrm{d}s\, \mathbf{1}_{\{\Delta \le \delta\}}\right] \mathbb{Q}_\delta^\psi(G(\rho)),
\end{aligned}$$

where we used Corollary 3.4.3 for the last equality. This in conjunction with (3.8.1) and (3.5.3) yields the desired result. □

In the critical case, the three conditionings $\Delta > \delta$, $Z_0^\delta = 1$ and $W^\delta = 1$ are not equivalent but they still yield the same asymptotic behavior: local convergence to the immortal Lévy tree just like in Theorem 3.7.8 where we condition by $\Delta = \delta$.

Theorem 3.8.4. *Assume that ψ is critical and that π is unbounded. Let $F\colon \mathscr{D} \to \mathbb{R}$ be continuous and bounded and let A_δ be equal to $\{\Delta > \delta\}$, $\{Z_0^\delta = 1\}$ or $\{W^\delta = 1\}$. We have*

$$\lim_{\delta\to\infty} \mathbf{N}^\psi \left[F(r_h(\rho)) \middle| A_\delta\right] = \mathbf{N}^\psi \left[L_\sigma^h F(r_h(\rho))\right]. \tag{3.8.11}$$

Proof. Since the conditioning by $\Delta > \delta$ was already treated in [86], we only consider the other two. The proof uses similar arguments to that of Theorem 3.7.8 and we only give a sketch. By Lemma 3.8.3, under $\mathbf{N}^\psi$ and conditionally on $Z_0^\delta = 1$, ρ is distributed as $\tilde\rho \circledast (s, \hat\rho)$, where $(s, \tilde\rho)$ has distribution

$$\mathbb{E}\left[F(s,\tilde\rho)\right] = \frac{1}{w_+(\delta)} \mathbf{N}^\psi \left[\int_0^\sigma F(s,\rho)\, \mathrm{d}s\, \mathbf{1}_{\{\Delta\le\delta\}}\right],$$

$\hat\rho$ has distribution $\mathbb{Q}_\delta^\psi$ and they are independent. But Lemma 3.5.10 gives that the height $H(\tilde\rho_s)$ is exponentially distributed with mean $w_+(\delta)$. Since ψ is critical, this last quantity goes to ∞ as $\delta \to \infty$. In particular, it holds that $H(\tilde\rho_s) > h$ with high probability as $\delta \to \infty$. Furthermore, on the event $\{H(\tilde\rho_s) > h\}$, we have that

$$r_h(\tilde\rho \circledast (s,\hat\rho)) = r_h(\tilde\rho). \tag{3.8.12}$$

As a consequence, in order to show the result, it is enough to prove that

$$\lim_{\delta\to\infty} \frac{1}{w_+(\delta)} \mathbf{N}^\psi \left[\sigma F(r_h(\rho)) \mathbf{1}_{\{\Delta\le\delta\}}\right] = \mathbf{N}^\psi \left[L_\sigma^h F(r_h(\rho))\right].$$

This last convergence holds by adapting the proof of Lemma 3.7.6. Finally, when conditioning on $W^\delta = 1$, the only change is that $\hat\rho$ has distribution $\mathbb{Q}_\delta^\psi(\cdot|\Delta\le\delta)$ but this does not contribute to the limit because of (3.8.12). This completes the proof. □

3.9 Stable case

We consider the stable case $\psi(\lambda) = \lambda^\gamma$ with $\gamma \in (1,2)$. Notice that the branching mechanism is critical with $\alpha = \beta = 0$ and the Lévy measure π is given by:

$$\pi(\mathrm{d}r) = a_\gamma r^{-1-\gamma}\, \mathrm{d}r, \quad \text{where} \quad a_\gamma = \frac{\gamma(\gamma-1)}{\Gamma(2-\gamma)}\cdot$$

Then we have:

$$\bar\pi(\delta) = \pi(\delta,\infty) = \frac{a_\gamma}{\gamma}\delta^{-\gamma}. \tag{3.9.1}$$

Furthermore, the Grey condition (3.2.21) is satisfied and we can speak of the Lévy tree $\mathscr{T}$, see Section 3.2.9.

We recall the scaling property of the stable tree. For every $\gamma \in (1,2)$, define the mapping $R_\gamma : \mathbb{T} \times (0,\infty) \to \mathbb{T}$ by:

$$R_\gamma((T,\emptyset,d,\mu),a) = (T,\emptyset,ad,a^{\gamma/(\gamma-1)}\mu), \quad \forall T \in \mathbb{T}. \tag{3.9.2}$$

In words, the real tree $R_\gamma((T,\emptyset,d,\mu),a)$ is obtained from $(T,\emptyset,d,\mu)$ by multiplying the metric by a and the measure by $a^{\gamma/(\gamma-1)}$. The choice of the exponent is justified by the following identity: for every $a>0$,

$$R_\gamma(\mathscr{T},a) \quad \text{under } \mathbf{N}^\psi \quad \overset{(d)}{=} \quad \mathscr{T} \quad \text{under } a^{1/(\gamma-1)}\mathbf{N}^\psi. \tag{3.9.3}$$

Using this, one can define a regular conditional probability measure $\mathbf{N}^\psi[\cdot|\sigma=a]$ such that $\mathbf{N}^\psi[\cdot|\sigma=a]$-a.s. $\sigma=a$ and

$$\mathbf{N}^\psi[\mathrm{d}\mathscr{T}] = \frac{1}{\gamma\Gamma(1-1/\gamma)}\int_0^\infty \frac{\mathrm{d}a}{a^{1+1/\gamma}}\mathbf{N}^\psi[\mathrm{d}\mathscr{T}|\sigma=a]. \tag{3.9.4}$$

Furthermore, under $\mathbf{N}^\psi[\cdot|\sigma=a]$, $\mathscr{T}$ is distributed as $R_\gamma(\mathscr{T},a^{1-1/\gamma})$ under $\mathbf{N}^\psi[\cdot|\sigma=1]$. We shall now establish the scaling property of the degree.

Proposition 3.9.1. *Let $\psi(\lambda)=\lambda^\gamma$ with $\gamma\in(1,2)$. Then, under $\mathbf{N}^\psi[\cdot|\Delta=\delta]$, the stable tree $\mathscr{T}$ is distributed as $R_\gamma(\mathscr{T},\delta^{\gamma-1})$ under $\mathbf{N}^\psi[\cdot|\Delta=1]$.*

Proof. Thanks to [58, Theorem 4.7], we can write the degree of the stable tree $\mathscr{T}$ as

$$\Delta(\mathscr{T}) = \sup_{x\in\mathscr{T}}\left(\lim_{\varepsilon\to 0}((\gamma-1)\varepsilon)^{-1/(\gamma-1)}n_{\mathscr{T}}(x,\varepsilon)\right),$$

where $n_{\mathscr{T}}(x,\varepsilon)$ is the number of subtrees originating from x with height greater than ε. In particular, it is straightforward to check that $\Delta(R_\gamma(\mathscr{T},a)) = a^{1/(\gamma-1)}\Delta(\mathscr{T})$. Then the conclusion readily follows from (3.9.3). □

Denote by $\Gamma(s,y)$ the upper incomplete gamma function:

$$\Gamma(s,y) = \int_y^\infty t^{s-1}\mathrm{e}^{-t}\,\mathrm{d}t, \quad \forall s\in\mathbb{R}, y>0.$$

Then the Laplace exponent ψ_δ is given by:

$$\psi_\delta(\lambda) = \lambda^\gamma + a_\gamma\int_\delta^\infty (1-\mathrm{e}^{-\lambda r})\frac{\mathrm{d}r}{r^{1+\gamma}} = \lambda^\gamma(1-a_\gamma\Gamma(-\gamma,\lambda\delta)) + \gamma^{-1}a_\gamma\delta^{-\gamma}. \tag{3.9.5}$$

We will aslo need its derivative:

$$\psi_\delta'(\lambda) = \lambda^{\gamma-1}(\gamma + a_\gamma \Gamma(1-\gamma, \lambda\delta)).$$

Proposition 3.9.2. *In the stable case $\psi(\lambda) = \lambda^\gamma$, we have:*

$$\mathbf{N}^\psi[\Delta > \delta] = c_\gamma \delta^{-1}, \tag{3.9.6}$$

$$\mathbf{N}^\psi[Z_0^\delta = 1] = \frac{c_\gamma}{\gamma} \mathrm{e}^{c_\gamma} \delta^{-1}, \tag{3.9.7}$$

$$\mathbf{N}^\psi[W^\delta = 1] = \left(c_\gamma - \frac{\gamma c_\gamma^{\gamma+1}}{a_\gamma} \mathrm{e}^{c_\gamma} \right) \delta^{-1}, \tag{3.9.8}$$

where $c_\gamma \in (0,\infty)$ is such that $\Gamma(-\gamma, c_\gamma) = a_\gamma^{-1}$.

Proof. Thanks to (3.3.4), we have $\psi_\delta(\mathbf{N}^\psi[\Delta > \delta]) = \bar{\pi}(\delta)$. Together with (3.9.5), this implies that $\delta \mathbf{N}^\psi[\Delta > \delta]$ is solution to $\Gamma(-\gamma, x) = a_\gamma^{-1}$. This proves (3.9.6).

To prove the remaining two identities, notice that

$$\psi_\delta'(\mathbf{N}^\psi[\Delta > \delta]) = \psi_\delta'(c_\gamma \delta^{-1}) = c_\gamma^{\gamma-1} \delta^{1-\gamma} \left(\gamma + a_\gamma \Gamma(1-\gamma, c_\gamma) \right) = \frac{a_\gamma}{c_\gamma} \mathrm{e}^{-c_\gamma} \delta^{1-\gamma},$$

where we used the identity $\Gamma(s+1, x) = s\Gamma(s, x) + x^s \mathrm{e}^{-x}$ together with the definition of c_γ for the last equality. The result readily follows from Proposition 3.8.1 by a straightforward computation. □

Lemma 3.9.3. *For every $\lambda \geq 0$, there exists a constant $c_\gamma(\lambda) \in (0,\infty)$ such that*

$$\psi_\delta^{-1}\left((1 - \mathrm{e}^{-\lambda}) \bar{\pi}(\delta) \right) = \frac{c_\gamma(\lambda)}{\delta}. \tag{3.9.9}$$

Moreover, $c_\gamma(\lambda)$ is the unique positive solution to $x^\gamma(a_\gamma \Gamma(-\gamma, x) - 1) = \gamma^{-1} a_\gamma \mathrm{e}^{-\lambda}$.

Proof. Fix $\lambda \geq 0$ and let

$$u_\gamma^\lambda(x) = x^\gamma(1 - a_\gamma \Gamma(-\gamma, x)) + \gamma^{-1} a_\gamma \mathrm{e}^{-\lambda}, \quad \forall x \geq 0.$$

Using the estimate $\Gamma(-\gamma, x) \sim \gamma^{-1} x^\gamma$ as $x \to 0$, elementary analysis gives that u_γ^λ has a unique root which we denote by $c_\gamma(\lambda)$. Thanks to (3.9.5), we get:

$$\psi_\delta(\delta^{-1} c_\gamma(\lambda)) = (1 - \mathrm{e}^{-\lambda}) \gamma^{-1} a_\gamma \delta^{-\gamma},$$

and the conclusion readily follows from (3.9.1). □

In the stable case, we can make explicit the distribution of the Bienaymé-Galton-Watson forest $\mathbf{t}_\delta$.

Proposition 3.9.4. *Under $\mathbf{N}^\psi$, conditionally on $\Delta > \delta$, the random forest $\mathbf{t}_\delta$ consisting of nodes with mass larger than δ is a critical (Z_0^δ, ξ^δ)-Bienaymé-Galton-Watson forest, where*

$$\mathbf{N}^\psi\left[1-\mathrm{e}^{-\lambda Z_0^\delta}\middle|\Delta>\delta\right] = \frac{c_\gamma(\lambda)}{c_\gamma} \quad \textit{and} \quad \mathbf{N}^\psi\left[\mathrm{e}^{-\lambda\xi^\delta}\middle|\Delta>\delta\right] = \mathrm{e}^{-\lambda} + \frac{\gamma}{a_\gamma}c_\gamma(\lambda). \tag{3.9.10}$$

In particular, conditionally on $\Delta > \delta$, the distribution of $\mathbf{t}_\delta$ is independent of δ.

Proof. Under $\mathbf{N}^{\psi_\delta}$ and conditionally on $\mathscr{T}$, let ζ be a Poisson random variable with parameter $\bar{\pi}(\delta)\sigma$. Notice that conditionally on $\Delta>\delta$, Z_0^δ is distributed as ζ under $\mathbf{N}^{\psi_\delta}$ conditionally on $\zeta \geq 1$. Thus we have:

$$\mathbf{N}^\psi\left[1-\mathrm{e}^{-\lambda Z_0^\delta}\middle|\Delta>\delta\right] = \frac{\mathbf{N}^{\psi_\delta}\left[(1-\mathrm{e}^{-\lambda\zeta})\mathbf{1}_{\{\zeta\geq 1\}}\right]}{\mathbf{N}^{\psi_\delta}\left[\zeta\geq 1\right]} = \frac{\mathbf{N}^{\psi_\delta}\left[1-\mathrm{e}^{-\lambda\zeta}\right]}{\mathbf{N}^{\psi_\delta}\left[\zeta\geq 1\right]}. \tag{3.9.11}$$

Since $\mathbf{N}^{\psi_\delta}[\zeta\geq 1] = \mathbf{N}^\psi[\Delta>\delta]$ thanks to Theorem 3.4.1, it follows from (3.4.7) that

$$\mathbf{N}^\psi\left[1-\mathrm{e}^{-\lambda Z_0^\delta}\middle|\Delta>\delta\right] = \frac{\psi_\delta^{-1}\left((1-\mathrm{e}^{-\lambda})\bar{\pi}(\delta)\right)}{\mathbf{N}^\psi[\Delta>\delta]}.$$

Combining (3.9.6) and (3.9.9), we deduce that

$$\mathbf{N}^\psi\left[1-\mathrm{e}^{-\lambda Z_0^\delta}\middle|\Delta>\delta\right] = \frac{c_\gamma(\lambda)}{c_\gamma}.$$

Next, thanks to Theorem 3.4.1, it is easy to see than under $\mathbf{N}^\psi$, the random variables ξ^δ and $\mathbf{1}_{\{\Delta>\delta\}} = \mathbf{1}_{\{Z_0^\delta\geq 1\}}$ are independent. It follows from (3.4.8) and Lemma 3.9.3 that

$$\mathbf{N}^\psi\left[\mathrm{e}^{-\lambda\xi^\delta}\middle|\Delta>\delta\right] = \mathbf{N}^\psi\left[\mathrm{e}^{-\lambda\xi^\delta}\right] = \mathrm{e}^{-\lambda} + \frac{\gamma}{a_\gamma}c_\gamma(\lambda).$$

□

✂ ....

Index of notation

Spaces

$\mathscr{M}_f(E)$	space of finite measures on E
$\mathscr{D}$	space of càdlàg functions from $\mathbb{R}_+$ to $\mathscr{M}_f(\mathbb{R}_+)$
$\mathscr{D}_0$	space of càdlàg excursions from $\mathbb{R}_+$ to $\mathscr{M}_f(\mathbb{R}_+)$

Random variables

ρ_t	exploration process
η_t	dual process
H_t	height process
σ	lifetime of the exploration process
$L^h(\mathrm{d}s)$	local time at level h
Δ	maximal degree of the exploration process
T_δ	first time the exploration process contains a node with mass larger than δ
$\rho^{\delta,-}$	path of the exploration process after removing the first node with mass larger than δ
$\rho^{\delta,+}$	path of the exploration process above the first node with mass larger than δ
$\mathbf{t}_\delta$	discrete tree consisting of nodes with mass larger than δ
W^δ	number of nodes with mass larger δ
Z_0^δ	number of first-generation nodes with mass larger than δ
T_Δ	first time the exploration process contains a node with mass Δ
H_Δ	height of the first node with mass Δ
$\rho^{\Delta,-}$	path of the exploration process after removing the first node with mass Δ
$\rho^{\Delta,+}$	path of the exploration process above the first node with mass Δ

Measures

$\mathbb{P}^\psi$	distribution of the exploration process starting from 0
$\mathbf{N}^\psi$	excursion measure of the exploration process
$\mathbb{P}_\nu^{\psi,*}$	distribution of the exploration process starting at ν and killed when it first reaches 0
$\mathbb{P}_r^\psi$	distribution of the exploration process with initial degree r

$\mathbb{Q}^{\psi}_{\delta}$	distribution of the exploration process with random initial degree, (3.2.19)
$\mathbf{P}^{\psi}_{\delta}$	distribution of a marked exploration process with degree restriction, (3.5.4)
$\mathbf{P}^{\psi}_{\delta,h}$	distribution of a marked (at level h) exploration process with degree restriction, (3.6.5)

Functions

$\bar{\pi}(\delta)$	tail of the Lévy measure π
$w(\delta)$	$\mathbf{N}^{\psi}[\sigma \mathbf{1}_{\{\Delta<\delta\}}]$
$w_+(\delta)$	$\mathbf{N}^{\psi}[\sigma \mathbf{1}_{\{\Delta\leq\delta\}}]$
$\mathfrak{g}(\delta)$	$\pi(\delta)\mathrm{e}^{-\delta \mathbf{N}^{\psi}[\Delta>\delta]}$

Bibliography

[1] R. Abraham and J.-F. Delmas, *Fragmentation associated with Lévy processes using snake*, Probab. Theory Related Fields, 141 (2008), pp. 113–154.

[2] ——, *Williams' decomposition of the Lévy continuum random tree and simultaneous extinction probability for populations with neutral mutations*, Stochastic Process. Appl., 119 (2009), pp. 1124–1143.

[3] ——, *The forest associated with the record process on a Lévy tree*, Stochastic Process. Appl., 123 (2013), pp. 3497–3517.

[4] ——, *Local limits of conditioned Galton-Watson trees: the condensation case*, Electron. J. Probab., 19 (2014), pp. no. 56, 29.

[5] R. Abraham, J.-F. Delmas, and P. Hoscheit, *A note on the Gromov-Hausdorff-Prokhorov distance between (locally) compact metric measure spaces*, Electron. J. Probab., 18 (2013), pp. no. 14, 21.

[6] R. Abraham, J.-F. Delmas, and M. Nassif, *Conditioning (sub)critical Lévy trees by their maximal degree: decomposition and local limit*, 2022. arXiv:2211.02317.

[7] R. Abraham, J.-F. Delmas, and M. Nassif, *Global regime for general additive functionals of conditioned Bienaymé-Galton-Watson trees*, Probab. Theory Related Fields, 182 (2022), pp. 277–351.

[8] R. Abraham, J.-F. Delmas, and G. Voisin, *Pruning a Lévy continuum random tree*, Electron. J. Probab., 15 (2010), pp. no. 46, 1429–1473.

[9] R. Abraham, J.-F. O. Delmas, and H. He, *Brownian continuum random tree conditioned to be large*, 2022. arXiv:2202.10258.

[10] R. Abraham and L. Serlet, *Poisson snake and fragmentation*, Electron. J. Probab., 7 (2002), pp. no. 17, 15.

[11] L. Addario-Berry and M. Albenque, *The scaling limit of random simple triangulations and random simple quadrangulations*, Ann. Probab., 45 (2017), pp. 2767–2825.

[12] ———, *Convergence of non-bipartite maps via symmetrization of labeled trees*, Ann. H. Lebesgue, 4 (2021), pp. 653–683.

[13] L. Addario-Berry, N. Broutin, C. Goldschmidt, and G. Miermont, *The scaling limit of the minimum spanning tree of the complete graph*, Ann. Probab., 45 (2017), pp. 3075–3144.

[14] L. Addario-Berry, L. Devroye, and S. Janson, *Sub-Gaussian tail bounds for the width and height of conditioned Galton-Watson trees*, Ann. Probab., 41 (2013), pp. 1072–1087.

[15] P.-M. Agapow and A. Purvis, *Power of eight tree shape statistics to detect nonrandom diversification: a comparison by simulation of two models of cladogenesis*, Systematic biology, 51 (2002), pp. 866–872.

[16] D. Aldous, *The continuum random tree. I*, Ann. Probab., (1991), pp. 1–28.

[17] ———, *The continuum random tree. II. An overview*, in Stochastic analysis (Durham, 1990), vol. 167 of London Math. Soc. Lecture Note Ser., Cambridge Univ. Press, Cambridge, 1991, pp. 23–70.

[18] ———, *The continuum random tree. III*, Ann. Probab., 21 (1993), pp. 248–289.

[19] D. Aldous and J. Pitman, *The standard additive coalescent*, Ann. Probab., 26 (1998), pp. 1703–1726.

[20] ———, *The standard additive coalescent*, Ann. Probab., 26 (1998), pp. 1703–1726.

[21] C. D. Aliprantis and K. C. Border, *Infinite dimensional analysis*, Springer, Berlin, third ed., 2006. A hitchhiker's guide.

[22] J. Ambjørn, B. Durhuus, and T. Jonsson, *Quantum geometry*, Cambridge Monographs on Mathematical Physics, Cambridge University Press, Cambridge, 1997. A statistical field theory approach.

[23] K. B. Athreya and P. E. Ney, *Branching processes*, Springer-Verlag, New York-Heidelberg, 1972. Die Grundlehren der mathematischen Wissenschaften, Band 196.

[24] S. Athreya, W. Löhr, and A. Winter, *The gap between Gromov-vague and Gromov-Hausdorff-vague topology*, Stochastic Process. Appl., 126 (2016), pp. 2527–2553.

[25] N. Bacaër, *A short history of mathematical population dynamics*, Springer-Verlag London, Ltd., London, 2011.

[26] J. Berestycki and N. Berestycki, *Kingman's coalescent and Brownian motion*, ALEA Lat. Am. J. Probab. Math. Stat., 6 (2009), pp. 239–259.

[27] J. Berestycki, N. Berestycki, and J. Schweinsberg, *Beta-coalescents and continuous stable random trees*, Ann. Probab., 35 (2007), pp. 1835–1887.

[28] J. BERESTYCKI, M. C. FITTIPALDI, AND J. FONTBONA, *Ray-Knight representation of flows of branching processes with competition by pruning of Lévy trees*, Probab. Theory Related Fields, 172 (2018), pp. 725–788.

[29] J. BERTOIN, *Lévy Processes*, vol. 121, Cambridge University Press, Cambridge, 1996.

[30] ———, *A fragmentation process connected to Brownian motion*, Probab. Theory Related Fields, 117 (2000), pp. 289–301.

[31] ———, *Homogeneous fragmentation processes*, Probab. Theory Related Fields, 121 (2001), pp. 301–318.

[32] ———, *Self-similar fragmentations*, Ann. Inst. H. Poincaré Probab. Statist., 38 (2002), pp. 319–340.

[33] ———, *Random fragmentation and coagulation processes*, vol. 102 of Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, 2006.

[34] ———, *On the maximal offspring in a critical branching process with infinite variance*, J. Appl. Probab., 48 (2011), pp. 576–582.

[35] J. BETTINELLI, *Scaling limit of random planar quadrangulations with a boundary*, Ann. Inst. Henri Poincaré Probab. Stat., 51 (2015), pp. 432–477.

[36] J. BETTINELLI, E. JACOB, AND G. MIERMONT, *The scaling limit of uniform random plane maps, via the Ambjørn-Budd bijection*, Electron. J. Probab., 19 (2014), pp. no. 74, 16.

[37] P. BIANE, J. PITMAN, AND M. YOR, *Probability laws related to the Jacobi theta and Riemann zeta functions, and Brownian excursions*, Bull. Amer. Math. Soc. (N.S.), 38 (2001), pp. 435–465.

[38] P. BILLINGSLEY, *Convergence of probability measures*, Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons, Inc., New York, second ed., 1999. A Wiley-Interscience Publication.

[39] J.-M. BISMUT, *Last exit decompositions and regularity at the boundary of transition probabilities*, Z. Wahrsch. Verw. Gebiete, 69 (1985), pp. 65–98.

[40] V. I. BOGACHEV, *Measure theory. Vol. I, II*, Springer-Verlag, Berlin, 2007.

[41] N. BROUTIN, T. DUQUESNE, AND M. WANG, *Limits of multiplicative inhomogeneous random graphs and Lévy trees: limit theorems*, Probab. Theory Related Fields, 181 (2021), pp. 865–973.

[42] ———, *Limits of multiplicative inhomogeneous random graphs and Lévy trees: The continuum graphs*, Ann. Appl. Probab., 32 (2022), pp. 2448–2503.

[43] N. BROUTIN AND J.-F. MARCKERT, *Asymptotics of trees with a prescribed degree sequence and applications*, Random Structures Algorithms, 44 (2014), pp. 290–316.

[44] A. CARACENI, *The scaling limit of random outerplanar maps*, Ann. Inst. Henri Poincaré Probab. Stat., 52 (2016), pp. 1667–1686.

[45] P. CHASSAING AND G. SCHAEFFER, *Random planar lattices and integrated superBrownian excursion*, Probab. Theory Related Fields, 128 (2004), pp. 161–212.

[46] N. CURIEN AND B. HAAS, *The stable trees are nested*, Probab. Theory Related Fields, 157 (2013), pp. 847–883.

[47] N. CURIEN, B. HAAS, AND I. KORTCHEMSKI, *The CRT is the scaling limit of random dissections*, Random Structures Algorithms, 47 (2015), pp. 304–327.

[48] D. A. DAWSON, *Measure-valued Markov processes*, in École d'Été de Probabilités de Saint-Flour XXI—1991, vol. 1541 of Lecture Notes in Math., Springer, Berlin, 1993, pp. 1–260.

[49] D. A. DAWSON AND E. A. PERKINS, *Measure-valued processes and renormalization of branching particle systems*, in Stochastic partial differential equations: six perspectives, vol. 64 of Math. Surveys Monogr., Amer. Math. Soc., Providence, RI, 1999, pp. 45–106.

[50] C. DELLACHERIE AND P.-A. MEYER, *Probabilités et potentiel*, Publications de l'Institut de Mathématique de l'Université de Strasbourg, No. XV, Hermann, Paris, 1975. Chapitres I à IV, Édition entièrement refondue.

[51] J.-F. DELMAS, *Fragmentation at height associated with Lévy processes*, Stochastic Process. Appl., 117 (2007), pp. 297–311.

[52] J.-F. DELMAS, J.-S. DHERSIN, AND M. SCIAUVEAU, *Cost functionals for large (uniform and simply generated) random trees*, Electron. J. Probab., 23 (2018), p. 36 pp.

[53] M. DRMOTA AND B. GITTENBERGER, *On the profile of random trees*, Random Structures Algorithms, 10 (1997), pp. 421–451.

[54] T. DUQUESNE, *A limit theorem for the contour process of condidtioned Galton-Watson trees*, Ann. Probab., 31 (2003), pp. 996–1027.

[55] ——, *The coding of compact real trees by real valued functions*, 2006. arXiv:math/0604106.

[56] ——, *Continuum random trees and branching processes with immigration*, Stochastic Process. Appl., 119 (2009), pp. 99–129.

[57] T. DUQUESNE AND J.-F. LE GALL, *Random trees, Lévy processes and spatial branching processes*, Astérisque, (2002), pp. vi+147.

[58] ——, *Probabilistic and fractal aspects of Lévy trees*, Probab. Theory Related Fields, 131 (2005), pp. 553–603.

[59] ——, *On the re-rooting invariance property of Lévy trees*, Electron. Commun. Probab., 14 (2009), pp. 317–326.

[60] T. DUQUESNE AND M. WANG, *Decomposition of Lévy trees along their diameter*, Ann. Inst. Henri Poincaré Probab. Stat., 53 (2017), pp. 539–593.

[61] T. DUQUESNE AND M. WINKEL, *Mass erasure of Lévy trees.* In preparation.

[62] ——, *Growth of Lévy trees*, Probab. Theory Related Fields, 139 (2007), pp. 313–371.

[63] M. DWASS, *The total progeny in a branching process and a related random walk*, J. Appl. Probability, 6 (1969), pp. 682–686.

[64] E. B. DYNKIN, *A probabilistic approach to one class of nonlinear differential equations*, Probab. Theory Related Fields, 89 (1991), pp. 89–115.

[65] ——, *Superprocesses and partial differential equations*, Ann. Probab., 21 (1993), pp. 1185–1262.

[66] K. B. ERICKSON AND R. A. MALLER, *Finiteness of integrals of functions of Lévy processes*, Proc. Lond. Math. Soc. (3), 94 (2007), pp. 386–420.

[67] A. M. ETHERIDGE, *An introduction to superprocesses*, vol. 20 of University Lecture Series, American Mathematical Society, Providence, RI, 2000.

[68] L. C. EVANS AND R. F. GARIEPY, *Measure theory and fine properties of functions*, Textbooks in Mathematics, CRC Press, Boca Raton, FL, revised ed., 2015.

[69] S. N. EVANS, *Probability and real trees*, vol. 1920 of Lecture Notes in Mathematics, Springer, Berlin, 2008. Lectures from the 35th Summer School on Probability Theory held in Saint-Flour, July 6–23, 2005.

[70] S. N. EVANS, J. PITMAN, AND A. WINTER, *Rayleigh processes, real trees, and root growth with re-grafting*, Probab. Theory Related Fields, 134 (2006), pp. 81–126.

[71] P.-H. FABRE, L. HAUTIER, D. DIMITROV, AND E. J. DOUZERY, *A glimpse on the pattern of rodent diversification: a phylogenetic approach*, BMC evolutionary biology, 12 (2012), p. 88.

[72] W. FELLER, *An introduction to probability theory and its applications. Vol. II*, John Wiley & Sons, Inc., New York-London-Sydney, second ed., 1971.

[73] J. A. FILL AND S. JANSON, *The sum of powers of subtree sizes for conditioned Galton–Watson trees*, Electronic Journal of Probability, 27 (2022), pp. 1 – 77.

[74] J. A. FILL AND N. KAPUR, *Limiting distributions for additive functionals on Catalan trees*, Theoret. Comput. Sci., 326 (2004), pp. 69–102.

[75] F. S. GALTON AND H. W. WATSON, *On the probability of the extinction of families*, The Journal of the Anthropological Institute of Great Britain and Ireland, 4 (1875), pp. 138–144.

[76] C. GOLDSCHMIDT AND B. HAAS, *Behavior near the extinction time in self-similar fragmentations. I. The stable case*, Ann. Inst. Henri Poincaré Probab. Stat., 46 (2010), pp. 338–368.

[77] ———, *A line-breaking construction of the stable trees*, Electron. J. Probab., 20 (2015), pp. no. 16, 24.

[78] A. GREVEN, P. PFAFFELHUBER, AND A. WINTER, *Convergence in distribution of random metric measure spaces (Λ-coalescent measure trees)*, Probab. Theory Related Fields, 145 (2009), pp. 285–322.

[79] M. GROMOV, *Metric structures for Riemannian and non-Riemannian spaces*, vol. 152 of Progress in Mathematics, Birkhäuser Boston, Inc., Boston, MA, 1999. Based on the 1981 French original [MR0682063 (85e:53051)], With appendices by M. Katz, P. Pansu and S. Semmes, Translated from the French by Sean Michael Bates.

[80] B. HAAS, *Fragmentation processes with an initial mass converging to infinity*, J. Theoret. Probab., 20 (2007), pp. 721–758.

[81] B. HAAS AND G. MIERMONT, *The genealogy of self-similar fragmentations with negative index as a continuum random tree*, Electron. J. Probab., 9 (2004), pp. no. 4, 57–97.

[82] ———, *Scaling limits of Markov branching trees with applications to Galton-Watson and random unordered trees*, Ann. Probab., 40 (2012), pp. 2589–2666.

[83] B. HAAS, G. MIERMONT, J. PITMAN, AND M. WINKEL, *Continuum tree asymptotics of discrete fragmentations and applications to phylogenetic models*, Ann. Probab., 36 (2008), pp. 1790–1837.

[84] B. HAAS, J. PITMAN, AND M. WINKEL, *Spinal partitions and invariance under re-rooting of continuum random trees*, Ann. Probab., 37 (2009), pp. 1381–1411.

[85] X. HE, *Conditioning Galton-Watson trees on large maximal outdegree*, J. Theoret. Probab., 30 (2017), pp. 842–851.

[86] ———, *Local convergence of critical random trees and continuous-state branching processes*, J. Theoret. Probab., 35 (2022), pp. 685–713.

[87] X. HE AND Z. LI, *The maximal jump and local convergence of continuous-state branching processes*, 2014. arXiv:1412.4615.

[88] ———, *Distributions of jumps in a continuous-state branching process with immigration*, J. Appl. Probab., 53 (2016), pp. 1166–1177.

[89] C. C. HEYDE AND E. SENETA, *Studies in the history of probability and statistics. XXXI. The simple branching process, a turning point test and a fundamental inequality: a historical note on I. J. Bienaymé*, Biometrika, 59 (1972), pp. 680–683.

[90] I. A. IBRAGIMOV AND Y. V. LINNIK, *Independent and stationary sequences of random variables*, Wolters-Noordhoff Publishing, Groningen, 1971. With a supplementary chapter by I. A. Ibragimov and V. V. Petrov, Translation from the Russian edited by J. F. C. Kingman.

[91] F. JABOT AND J. CHAVE, *Inferring the parameters of the neutral theory of biodiversity using phylogenetic information and implications for tropical forests*, Ecology letters, 12 (2009), pp. 239–248.

[92] J. JACOD AND A. N. SHIRYAEV, *Limit theorems for stochastic processes*, vol. 288 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin, second ed., 2003.

[93] S. JANSON, *Stable distributions*, 2011. arXiv:1112.0220.

[94] S. JANSON, *Simply generated trees, conditioned Galton-Watson trees, random allocations and condensation*, Probab. Surv., 9 (2012), pp. 103–252.

[95] ———, *Asymptotic normality of fringe subtrees and additive functionals in conditioned Galton-Watson trees*, Random Structures Algorithms, 48 (2016), pp. xiii+694.

[96] S. JANSON AND S. O. STEFÁNSSON, *Scaling limits of random planar maps with a unique large face*, Ann. Probab., 43 (2015), pp. 1045–1081.

[97] M. JIŘINA, *Stochastic branching processes with continuous state space*, Czechoslovak Math. J., 8(83) (1958), pp. 292–313.

[98] T. JONSSON AND S. O. STEFÁNSSON, *Condensation in nongeneric trees*, J. Stat. Phys., 142 (2011), pp. 277–313.

[99] O. KALLENBERG, *Random measures, theory and applications*, vol. 77 of Probability Theory and Stochastic Modelling, Springer, Cham, 2017.

[100] J. H. B. KEMPERMAN, *The General One-Dimensional Random Walk with Absorbing Barriers with Applications to Sequential Analysis*, Excelsiors Foto-Offset, 's-Gravenhage, 1950.

[101] J. H. B. KEMPERMAN, *The passage problem for a stationary Markov chain*, Statistical Research Monographs, Vol. I, University of Chicago Press, Chicago, Ill., 1961.

[102] D. G. KENDALL, *Branching processes since 1873*, J. London Math. Soc., 41 (1966), pp. 385–406. (1 plate).

[103] ———, *The genealogy of genealogy: branching processes before (and after) 1873*, Bull. London Math. Soc., 7 (1975), pp. 225–253. With a French appendix containing Bienaymé's paper of 1845.

[104] A. KHEZELI, *Metrization of the Gromov-Hausdorff (-Prokhorov) topology for boundedly-compact metric spaces*, Stochastic Process. Appl., 130 (2020), pp. 3842–3864.

[105] J. F. C. KINGMAN, *The coalescent*, Stochastic Process. Appl., 13 (1982), pp. 235–248.

[106] ——, *On the genealogy of large populations*, J. Appl. Probab., (1982), pp. 27–43. Essays in statistical science.

[107] M. KIRXPATRICK AND M. SLATKIN, *Searching for evolutionary patterns in the shape of a phylogenetic tree*, Evolution, 47 (1993), pp. 1171–1181.

[108] S. KOCHEN AND C. STONE, *A note on the Borel-Cantelli lemma*, Illinois J. Math., 8 (1964), pp. 248–251.

[109] I. KORTCHEMSKI, *Invariance principles for Galton-Watson trees conditioned on the number of leaves*, Stochastic Process. Appl., 122 (2012), pp. 3126–3172.

[110] ——, *A simple proof of Duquesne's theorem on contour processes of conditioned Galton-Watson trees*, in Séminaire de Probabilités XLV, vol. 2078 of Lecture Notes in Math., Springer, Cham, 2013, pp. 537–558.

[111] ——, *Limit theorems for conditioned non-generic Galton-Watson trees*, Ann. Inst. Henri Poincaré Probab. Stat., 51 (2015), pp. 489–511.

[112] ——, *Sub-exponential tail bounds for conditioned stable Bienaymé–Galton–Watson trees*, Probab. Theory Related Fields, 168 (2017), pp. 1–40.

[113] J. LAMPERTI, *Continuous state branching processes*, Bull. Amer. Math. Soc., 73 (1967), pp. 382–386.

[114] ——, *The limit of a sequence of branching processes*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete, 7 (1967), pp. 271–288.

[115] G. LAST AND M. PENROSE, *Lectures on the Poisson process*, vol. 7 of Institute of Mathematical Statistics Textbooks, Cambridge University Press, Cambridge, 2018.

[116] J.-F. LE GALL, *Random trees and applications*, Probab. Surv., 2 (2005), pp. 245–311.

[117] ——, *The topological structure of scaling limits of large planar maps*, Invent. Math., 169 (2007), pp. 621–670.

[118] ——, *Uniqueness and universality of the Brownian map*, Ann. Probab., 41 (2013), pp. 2880–2960.

[119] J.-F. LE GALL AND Y. LE JAN, *Branching processes in Lévy processes: Laplace functionals of snakes and superprocesses*, Ann. Probab., 26 (1998), pp. 1407–1432.

[120] Z. LI, E. PARDOUX, AND A. WAKOLBINGER, *The height process of a continuous-state branching process with interaction*, J. Theoret. Probab., 35 (2022), pp. 142–185.

[121] W. LÖHR, *Equivalence of Gromov-Prohorov- and Gromov's $\underline{\square}_\lambda$-metric on the space of metric measure spaces*, Electron. Commun. Probab., 18 (2013), pp. no. 17, 10.

[122] P. MARCHAL, *A note on the fragmentation of a stable tree*, in Fifth Colloquium on Mathematics and Computer Science, Discrete Math. Theor. Comput. Sci. Proc., AI, Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2008, pp. 489–499.

[123] J.-F. MARCKERT AND G. MIERMONT, *The CRT is the scaling limit of unordered binary trees*, Random Structures Algorithms, 38 (2011), pp. 467–501.

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

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

[126] ——, *Invariance principles for spatial multitype Galton-Watson trees*, Ann. Inst. Henri Poincaré Probab. Stat., 44 (2008), pp. 1128–1161.

[127] ——, *The Brownian map is the scaling limit of uniform random plane quadrangulations*, Acta Math., 210 (2013), pp. 319–401.

[128] M. NASSIF, *Zooming in at the root of the stable tree*, Electron. J. Probab., 27 (2022), pp. 1–38.

[129] J. NEVEU, *Arbres et processus de Galton-Watson*, Ann. Inst. H. Poincaré Probab. Statist., 22 (1986), pp. 199–207.

[130] R. OTTER, *The multiplicative process*, Ann. Math. Statistics, 20 (1949), pp. 206–224.

[131] K. PANAGIOTOU AND B. STUFLER, *Scaling limits of random Pólya trees*, Probab. Theory Related Fields, 170 (2018), pp. 801–820.

[132] K. PANAGIOTOU, B. STUFLER, AND K. WELLER, *Scaling limits of random graphs from subcritical classes*, Ann. Probab., 44 (2016), pp. 3291–3334.

[133] F. PAULIN, *The Gromov topology on* $\mathbf{R}$*-trees*, Topology Appl., 32 (1989), pp. 197–221.

[134] E. PERKINS, *Dawson-Watanabe superprocesses and measure-valued diffusions*, in Lectures on probability theory and statistics (Saint-Flour, 1999), vol. 1781 of Lecture Notes in Math., Springer, Berlin, 2002, pp. 125–324.

[135] J. PITMAN, *Coalescents with multiple collisions*, Ann. Probab., 27 (1999), pp. 1870–1902.

[136] ——, *The SDE solved by local times of a Brownian excursion or bridge derived from the height profile of a random tree or forest*, Ann. Probab., 27 (1999), pp. 261–283.

[137] ——, *Combinatorial stochastic processes*, vol. 1875 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, 2006. Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002.

[138] A. F. POON, *Phylodynamic inference with kernel ABC and its application to HIV epidemiology*, Molecular biology and evolution, 32 (2015), pp. 2483–2495.

[139] D. RALAIVAOSAONA, M. ŠILEIKIS, AND S. WAGNER, *A central limit theorem for almost local additive tree functionals*, Algorithmica, 82 (2020), pp. 642–679.

[140] S. I. RESNICK, *Heavy-tail phenomena*, Springer Series in Operations Research and Financial Engineering, Springer, New York, 2007. Probabilistic and statistical modeling.

[141] N. ROSS AND Y. WEN, *Scaling limits for some random trees constructed inhomogeneously*, Electron. J. Probab., 23 (2018), pp. Paper No. 5, 35.

[142] S. SAGITOV, *The general coalescent with asynchronous mergers of ancestral lines*, J. Appl. Probab., 36 (1999), pp. 1116–1125.

[143] J. SCOTT, P. MAINI, A. ANDERSON, AND A. FLETCHER, *Inferring tumor proliferative organization from phylogenetic tree measures in a computational model*, Systematic biology, 69 (2020), pp. 623–637.

[144] K.-T. SHAO AND R. R. SOKAL, *Tree balance*, Systematic Zoology, 39 (1990), pp. 266–276.

[145] B. STUFLER, *Scaling limits of random outerplanar maps with independent link-weights*, Ann. Inst. Henri Poincaré Probab. Stat., 53 (2017), pp. 900–915.

[146] ———, *On the maximal offspring in a subcritical branching process*, Electron. J. Probab., 25 (2020), pp. Paper No. 104, 62.

[147] L. SZÉKELY, S. WAGNER, AND H. WANG, *Problems related to graph indices in trees*, in Recent trends in combinatorics, vol. 159 of IMA Vol. Math. Appl., Springer, [Cham], 2016, pp. 3–30.

[148] L. TAKÁCS, *On the total heights of random rooted binary trees*, J. Combin. Theory Ser. B, 61 (1994), pp. 155–166.

[149] S. TAVARÉ, *Ancestral inference in population genetics*, in Lectures on probability theory and statistics, vol. 1837 of Lecture Notes in Math., Springer, Berlin, 2004, pp. 1–188.

[150] M. TYRAN-KAMIŃSKA, *Convergence to Lévy stable processes under some weak dependence conditions*, Stochastic Process. Appl., 120 (2010), pp. 1629–1650.

[151] C. VILLANI, *Optimal transport*, vol. 338 of Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin, 2009. Old and new.

[152] G. VOISIN, *Dislocation measure of the fragmentation of a general Lévy tree*, ESAIM Probab. Stat., 15 (2011), pp. 372–389.

[153] S. WAGNER, *Central limit theorems for additive tree parameters with small toll functions*, Combin. Probab. Comput., 24 (2015), pp. 329–353.

[154] S. J. WOLFE, *On moments of probability distribution functions*, in Fractional calculus and its applications (Proc. Internat. Conf., Univ. New Haven, West Haven, Conn., 1974), 1975, pp. 306–316. Lecture Notes in Math., Vol. 457.

[155] V. M. ZOLOTAREV, *One-dimensional stable distributions*, vol. 65 of Translations of Mathematical Monographs, American Mathematical Society, Providence, RI, 1986. Translated from the Russian by H. H. McFaden, Translation edited by Ben Silver.