University of Oxford

MMath in Mathematics and Statistics (Part C)

Dissertation on a Topic in Statistics

A Recursive Distributional Equation for the Stable Tree

Candidate Number: 570698

supervised by
Dr. Matthias Winkel

Hilary Term 2017

Contents

List of Figures

A Recursive Distributional Equation for the Stable Tree

Candidate Number: 570698

Hilary Term 2017

Abstract: Given a parameter $\alpha \in (1, 2]$, we employ Marchal's random growth algorithm for Duquesne and Le Gall's α-stable tree to demonstrate that its law is a fixpoint solution of a stochastic equation $\mathcal{T} \overset{d}{=} g\left(\xi, \mathcal{T}_i : i \geq 0\right)$, where $\xi = (\xi_i : i \geq 0)$ are scaling factors, $\mathcal{T}_i$ are i.i.d copies of the α-stable tree, and g is a concatenation operator. This is a new recursive distributional equation in the general framework of Aldous and Bandyopadhyay. We will also discuss other fixpoint solutions to this recursive distributional equation. We highlight connections to other recursive constructions in the literature when $\alpha = 2$, namely, Aldous' Brownian continuum random tree.

1 Introduction

Real trees, or $\mathbb{R}$-trees, constitute a class of loop-free length spaces which frequently turn up as scaling limits of many discrete trees [27]. In their own right, $\mathbb{R}$-trees have multiple applications, such as, phylogenetic models [35] and rough path integration theory [42]. Following Aldous' introduction of the Brownian continuum random tree (BCRT)[5, 6, 7], significant attention turned to random $\mathbb{R}$-trees. Naturally, the BCRT manifests in the asymptotics of discrete tree-like structures, including uniform random labelled trees [5, 7] and critical Galton-Watson trees with finite offspring variance [5]. Bewilderingly, recent applications of the BCRT have surpassed objects not overtly tree-like, for example: random recursive triangulations [22], random planar quadrangulations [55], and Liouville quantum gravity [25].

The BCRT was further generalised by Duquesne and Le Gall's α-stable trees [29, 31], parameterised by $\alpha \in (1, 2]$. The α-stable trees are themselves a special case of Le Gall and Le Jan's Lévy trees [49]. The family of α-stable trees represents the genealogies of continuous-state branching processes with branching mechanism $\psi(\lambda) = \lambda^\alpha$. When $\alpha = 2$, we recover the BCRT. Akin to the BCRT, the family of α-stable trees constitutes all possible scaling limits of Galton-Watson trees, conditioned on its total progeny, whose offspring distribution lie in the domain of attraction of an α-stable law [26]. Likewise, α-stable trees emerge in scaling limits of numerous discrete tree structures: vertex-cut Galton-Watson trees [24] and conditioned stable Lévy forests [17]. Pursuing a dedicated approach with Lévy processes gives links to superprocesses [29, 49], and beta-coalescents in genetic models [1, 11]. Particular aspects of α-stable trees, such as, invariance under uniform re-rooting [41], Hausdorff and packing measures [28, 30, 31], spectral dimensions [20], heights and diameters [32], and an embedding property of stable trees [21], have also been closely studied.

However, we wish to emphasise a crucial self-similarity property of α-stable trees. This property plausibly explains the prevalence of α-stable trees in such diverse contexts, especially in problems of a recursive nature. Decomposing an α-stable tree above a certain height or at appropriate nodes results in the connected components after decomposition forming independent rescaled copies of the original tree. This observation was first formalised by Miermont [52, 53], building upon Bertoin's self-similar fragmentation theory [12].

In this dissertation, we express the self-similarity of the α-stable tree by a new *recursive distributional equation* (RDE) in the setting of Aldous and Bandyopadhyay's survey paper [9]. Given a random variable $\mathcal{T}$ valued in a Polish metric space $(\mathbb{T}, d)$, an RDE is a stochastic equation of the form

$$\mathcal{T} \overset{d}{=} g\left(\xi, \mathcal{T}_i : i \geq 0\right) \quad \text{ on } \mathbb{T},$$

where $\mathcal{T}_i \overset{i.i.d}{\sim} \mathcal{T}$ for $i \geq 0$, g is a measurable mapping, and $\xi = (\xi_i : i \geq 0)$ is independent of $(\mathcal{T}_i : i \geq 0)$. RDEs are pertinent in various contexts with recursive structures, including, Galton-Watson branching processes [9], Poisson weighted infinite trees [10], and Quicksort algorithms [63].

RDEs have been employed in the recursive construction of the BCRT by Albenque and Goldschmidt [4]. Broutin and Sulzbach [14] extended this to further recursive combinatorial structures, and $\mathbb{R}$-trees, under a finite concatenation operation. Rembart and Winkel [61] did so with continuum random trees under a countably infinite concatenation operation.

Paralleling their approaches, we regard g as a concatenation operator acting on (countably) infinitely many $\mathbb{R}$-trees $\mathcal{T}_i \overset{i.i.d}{\sim} \mathcal{T}$ rescaled by $\xi_i > 0$, which then gives us a version of $\mathcal{T}$. In Section 4, Theorem 4.3.4 shows that the law of the α-stable tree is a fixpoint solution of an RDE. Our primary argument appeals to Marchal's random growth algorithm [51], which provides a recursive method of constructing α-stable trees as a scaling limit. However, we do not have uniqueness of solution for this RDE. Other fixpoints can be obtained by rescaling the metric by a multiplicative constant, or by decorating the α-stable tree with massless branches described in Proposition 4.3.5.

In Section 5, we demonstrate how our results relate to the recursive constructions in [4] and [14] for the BCRT. Our contribution lies in extending their finite concatenation operations to handle α-stable trees. While [61] presents an RDE for which the law of the α-stable tree is a unique and attractive fixpoint, we do not follow their approach of employing bead-splitting processes of [59]. Our arguments only require Poisson-Dirichlet and Chinese Restaurant Processes, which are intrinsic in bead-splitting processes. This gives a less technical recursive construction of α-stable trees that elucidates how mass partitions in α-stable trees relate to urn models and partition-valued processes. Furthermore, we prove the self-similarity property of α-stable trees decomposed at a branch point solely via the recursive nature of Marchal's algorithm, without need for Miermont's fragmentation tree theory [53].

Structure of the dissertation: In Section 2, we state background results pertaining to $\mathbb{R}$-trees and α-stable trees. Section 3 develops tools required to obtain our results. Namely, rigorous setups of RDEs, a Pólya urn model, the Chinese Restaurant Process, and Marchal's algorithm. Section 4 is dedicated to establishing an RDE for the law of the α-stable tree and studying other fixpoint solutions. In Section 5, we connect approaches used for the recursive construction of the BCRT in [4] and [14]. We supply further lines of inquiry that stem from this dissertation in Section 6.

2 Preliminaries on Trees

We introduce several background formalisms and theories on $\mathbb{R}$-trees, metrics applied to $\mathbb{R}$-trees, and α-stable trees.

2.1 $\mathbb{R}$-trees

Definition 2.1.1. *A metric space $(\mathcal{T}, d)$ is an $\mathbb{R}$-tree if for every $a, b \in \mathcal{T}$,*

- *There exists a unique isometry $f_{a,b}\colon [0, d(a,b)] \to \mathcal{T}$ such that*

$$f_{a,b}(0) = a \text{ and } f_{a,b}(d(a,b)) = b.$$

  Denote the image $f_{a,b}([0, d(a,b)])$ by $[\![a, b]\!]$.

- *If $g\colon [0,1] \to \mathcal{T}$ is a continuous injective map with $g(0) = a$ and $g(1) = b$, then*

$$g([0,1]) = [\![a, b]\!].$$

  The only non self-intersecting path from a to b is $[\![a, b]\!]$.

A *rooted* $\mathbb{R}$-tree is an $\mathbb{R}$-tree with a distinguished vertex $\rho \in \mathcal{T}$ called a *root*, denoted by $(\mathcal{T}, d, \rho)$. The *degree* of a vertex $a \in \mathcal{T}$ is the number of connected components in $\mathcal{T} \setminus \{a\}$. A *leaf* is a vertex $a \in \mathcal{T} \setminus \{\rho\}$ which has degree 1. Denote the set of leaves in $\mathcal{T}$ by $\mathcal{L}(\mathcal{T})$, and call $\mathcal{T}$ *leaf-dense* if $\overline{\mathcal{L}(\mathcal{T})} = \mathcal{T}$. We say that $a \in \mathcal{T} \setminus \{\rho\}$ is a *branch point* if $\mathcal{T} \setminus \{a\}$ has at least three connected components. Finally, for any $a \in \mathcal{T}$, we define the *height* of a as $d(\rho, a)$, and the *height* of $\mathcal{T}$ as $\mathrm{ht}(\mathcal{T}) := \sup_{a \in \mathcal{T}} d(\rho, a)$.

An $\mathbb{R}$-tree is *compact* if $(\mathcal{T}, d)$ is compact. Two rooted $\mathbb{R}$-trees $(\mathcal{T}, d, \rho)$ and $(\mathcal{T}', d', \rho')$ are *GH-equivalent* if there exists an isometry $f\colon \mathcal{T} \to \mathcal{T}'$ such that $f(\rho) = \rho'$. The set of equivalence classes of compact rooted $\mathbb{R}$-trees is denoted by $\mathbb{T}$. We focus on compact rooted $\mathbb{R}$-trees hereafter. $\mathbb{R}$-trees may also be characterised by the *four points condition.*

Theorem 2.1.2. *[15, Theorem 1] Let $(\mathcal{T}, d)$ be a compact and path-connected metric space, then $(\mathcal{T}, d)$ is a compact $\mathbb{R}$-tree if and only if it satisfies the four points condition:*

$$\forall x, y, z, w \in \mathcal{T} : d(x,y) + d(z,w) \leq \max\{(d(x,z) + d(y,w)), (d(x,w) + d(y,z))\}.$$

2.2 Metrics on Compact $\mathbb{R}$-trees

Given a metric space (X, δ), define the *Hausdorff metric* between compact subsets $K, K' \subseteq X$ as

$$\delta_H(K, K') := \inf\{\epsilon > 0 : K \subseteq U_\epsilon(K') \text{ and } K' \subseteq U_\epsilon(K)\}, \tag{1}$$

where $U_\epsilon(K) := \{x \in X : \inf_{y \in K} \delta(x,y) \leq \epsilon\}$ is the ϵ-*halo* of $K \subseteq X$.

Suppose $(\mathcal{T}, d, \rho)$ and $(\mathcal{T}', d', \rho')$ are compact rooted $\mathbb{R}$-trees, the *Gromov-Hausdorff metric* is defined as

$$d_{GH}((\mathcal{T}, d, \rho), (\mathcal{T}', d', \rho')) := \inf_{\phi, \phi'} \left(\delta_H(\phi(\mathcal{T}), \phi'(\mathcal{T}')) \vee \delta(\phi(\rho), \phi'(\rho'))\right), \tag{2}$$

where the infimum is taken over all metric spaces (X, δ) and all isometric embeddings $\phi\colon \mathcal{T} \to X$ and $\phi'\colon \mathcal{T}' \to X$. By virtue of this, the Gromov-Hausdorff distance between $\mathcal{T}$ and $\mathcal{T}'$ only depends on the GH-equivalence classes of $(\mathcal{T}, d, \rho)$ and $(\mathcal{T}', d', \rho')$, thereby inducing a metric on $\mathbb{T}$, which we also denote by d_{GH}.

There is an alternative characterisation of the Gromov-Hausdorff metric [16, Theorem 7.3.25]. Given two compact metric spaces (X, δ) and (X', δ'), a *correspondence* between X and X' is a subset

$\mathcal{R} \subseteq X \times X'$ such that for every $x \in X$, there exists at least one $x' \in X'$ such that $(x, x') \in \mathcal{R}$, and conversely, for every $y' \in X'$, there exists at least one $y \in X$ such that $(y, y') \in \mathcal{R}$. The *distortion* of this correspondence $\mathcal{R}$ is defined as

$$\operatorname{dis}(\mathcal{R}) := \sup \left\{ \left|\delta(x, y) - \delta'(x', y')\right| : (x, x'), (y, y') \in \mathcal{R} \right\}. \tag{3}$$

Applying the above to our compact rooted $\mathbb{R}$-trees $(\mathcal{T}, d, \rho)$ and $(\mathcal{T}', d', \rho')$, and defining $\mathcal{C}(\mathcal{T}, \mathcal{T}')$ as the set of all correspondences between $(\mathcal{T}, d, \rho)$ and $(\mathcal{T}', d', \rho')$ which have (ρ, ρ') in correspondence, then

$$d_{GH}((\mathcal{T}, d, \rho), (\mathcal{T}', d', \rho')) = \frac{1}{2} \inf_{\mathcal{R} \in \mathcal{C}(\mathcal{T}, \mathcal{T}')} \operatorname{dis}(\mathcal{R}). \tag{4}$$

We note that by definition of a correspondence, (4) does not depend on the representatives chosen in each GH-equivalence class.

In some instances, we may want to specify marked points on a compact rooted $\mathbb{R}$-tree. We will only consider compact rooted $\mathbb{R}$-trees marked at a single point x and denote this space by $(\mathcal{T}, d, \rho, x)$. We refer the reader to [54, Section 6.4] for further extensions. Given two marked compact rooted $\mathbb{R}$-trees $(\mathcal{T}, d, \rho, x)$ and $(\mathcal{T}', d', \rho', x')$, the *marked Gromov-Hausdorff metric* is defined as

$$d_{GH}^m\left((\mathcal{T}, d, \rho, x), (\mathcal{T}', d', \rho', x')\right) := \inf_{\phi, \phi'} \left(\delta_H\left(\phi(\mathcal{T}), \phi'(\mathcal{T}')\right) \vee \delta\left(\phi(\rho), \phi'(\rho')\right) \vee \delta\left(\phi(x), \phi'(x')\right)\right), \tag{5}$$

where the infimum is taken over all metric spaces (X, δ) and all isometric embeddings $\phi\colon \mathcal{T} \to X$ and $\phi'\colon \mathcal{T}' \to X$. We say two marked compact rooted $\mathbb{R}$-trees are *GH^m-equivalent* if there exists an isometry $f\colon \mathcal{T} \to \mathcal{T}'$ such that $f(\rho) = \rho'$ and $f(x) = x'$. Denote the set of equivalence classes of marked compact rooted $\mathbb{R}$-trees by $\mathbb{T}_m$. The marked Gromov-Hausdorff distance only depends on the GH^m-equivalence classes of $(\mathcal{T}, d, \rho, x)$, thereby inducing a metric on $\mathbb{T}_m$, which we also denote by d_{GH}^m.

We also have a characterisation of the marked Gromov-Hausdorff metric via distortions [54, Proposition 9(i)]. For marked compact rooted $\mathbb{R}$-trees $(\mathcal{T}, d, \rho, x)$ and $(\mathcal{T}', d', \rho', x')$, denote by $\mathcal{C}^m(\mathcal{T}, \mathcal{T}')$ the set of all correspondences between $(\mathcal{T}, d, \rho, x)$ and $(\mathcal{T}', d', \rho', x')$ which have (ρ, ρ') and (x, x') in correspondence, then

$$d_{GH}^m((\mathcal{T}, d, \rho, x), (\mathcal{T}', d', \rho', x')) = \frac{1}{2} \inf_{\mathcal{R} \in \mathcal{C}^m(\mathcal{T}, \mathcal{T}')} \operatorname{dis}(\mathcal{R}). \tag{6}$$

Suppose now that (X, δ) is a complete metric space, then (X, δ, μ) is a *measured metric space* if it is further equipped with a Borel probability measure with respect to δ on X. The *Prokhorov metric* between two measured metric spaces (X, δ, μ) and (X, δ, μ') is defined as

$$\delta_P(\mu, \mu') := \inf\{\epsilon > 0 : \mu(D) \leq \mu'(U_\epsilon(D)) + \epsilon \text{ and } \mu'(D) \leq \mu(U_\epsilon(D)) + \epsilon \text{ for all } D \subseteq X \text{ closed}\}. \tag{7}$$

Define a *weighted* $\mathbb{R}$-tree as a compact rooted $\mathbb{R}$-tree $(\mathcal{T}, d, \rho)$ equipped with a Borel probability measure μ. The *Gromov-Prokhorov metric* between two weighted $\mathbb{R}$-trees $(\mathcal{T}, d, \rho, \mu)$ and $(\mathcal{T}', d', \rho', \mu')$ is defined as

$$d_{GP}((\mathcal{T}, d, \rho, \mu), (\mathcal{T}', d', \rho', \mu')) := \inf_{\phi, \phi'} \left(\delta(\phi(\rho), \phi'(\rho')) \vee \delta_P(\phi_*\mu, \phi'_*\mu')\right), \tag{8}$$

where the infimum is taken over all metric spaces (X, δ) and all isometric embeddings $\phi\colon \mathcal{T} \to X$ and $\phi'\colon \mathcal{T}' \to X$, δ_P is as in (7), and $\phi_*\mu, \phi'_*\mu'$ are push-forwards of μ, μ' under ϕ, ϕ' respectively.

Two weighted $\mathbb{R}$-trees $(\mathcal{T}, d, \rho, \mu)$ and $(\mathcal{T}', d', \rho', \mu')$ are *GHP-equivalent* if there is an isometry $f\colon (\mathcal{T}, d, \rho, \mu) \to (\mathcal{T}', d', \rho', \mu')$ such that $f(\rho) = \rho'$ and μ' is the push-forward of μ under f. Denote the set of equivalence classes of weighted $\mathbb{R}$-trees by $\mathbb{T}_w$. Define the *Gromov-Hausdorff-Prokhorov metric* between two weighted $\mathbb{R}$-trees $(\mathcal{T}, d, \rho, \mu)$ and $(\mathcal{T}', d', \rho', \mu')$ as

$$d_{GHP}((\mathcal{T}, d, \rho, \mu), (\mathcal{T}', d', \rho', \mu')) := \inf_{\phi, \phi'} \left(\delta_H(\phi(\mathcal{T}), \phi'(\mathcal{T}')) \vee \delta(\phi(\rho), \phi'(\rho')) \vee \delta_P(\phi_*\mu, \phi'_*\mu') \right), \quad (9)$$

where the infimum is taken over all metric spaces (X, δ) and all isometric embeddings $\phi\colon \mathcal{T} \to X$ and $\phi'\colon \mathcal{T}' \to X$, δ_H is as in (1), δ_P is as in (7), and $\phi_*\mu, \phi'_*\mu'$ are push-forwards of μ, μ' under ϕ, ϕ' respectively. The Gromov-Hausdorff-Prokhorov metric only depends on the GHP-equivalence classes of $(\mathcal{T}, d, \rho, \mu)$ and $(\mathcal{T}', d', \rho', \mu')$ in $\mathbb{T}_w$, thereby inducing a metric on $\mathbb{T}_w$, also denoted by d_{GHP}. We state a key result about the metric spaces we introduced.

Theorem 2.2.1. *The following are Polish metric spaces, that is, complete separable metric spaces:*

(i) [34, Theorem 4.23] $(\mathbb{T}, d_{GH})$,

(ii) [54, Proposition 9(ii)] $(\mathbb{T}_m, d^m_{GH})$,

(iii) [2, Theorem 2.7] $(\mathbb{T}_w, d_{GHP})$.

2.3 Coding $\mathbb{R}$-trees with Continuous Height Functions

The formalisms required to define $\mathbb{R}$-trees might make it seem a rather abstract entity. We may encode $\mathbb{R}$-trees by a continuous height function over a real parameter, allowing us to construe $\mathbb{R}$-trees in the familiar setting of function spaces. Denote $\mathcal{H}_M$ as the set of non-negative continuous functions h on $[0, M]$ such that $h(0) = h(M) = 0$. The following results hold for non-negative càdlàg functions on $[0, M]$ with non-positive jumps. We do not consider this here, and refer the reader to [27]. Denote $\mathcal{H} := \bigcup_{M \geq 0} \mathcal{H}_M$ as the set of *height functions*. Fix $h \in \mathcal{H}_M$, then given any $0 \leq s \leq t \leq M$, define:

$$\begin{aligned} m_h(s,t) &:= \inf_{r \in [s,t]} h(r) \quad \text{and} \\ d_h(s,t) &:= h(s) + h(t) - 2m_h(s,t). \end{aligned} \quad (10)$$

By definition, $d_h(s,t)$ is non-negative and symmetric. The triangle inequality follows by the properties of the infimum. However, this is a pseudo-metric on $[0, M]$. For example, we might have $h(s) = h(t) = m_h(s,t)$ for s, t distinct. To make this a metric, define an equivalence relation on elements in $[0, M]$ by $s \sim_h t \iff d_h(s,t) = 0$. Define $T_h := [0, M]/\sim_h$ and let $\tau_h\colon [0, M] \to T_h$ be the canonical projection onto a quotient space. Finally, let $\tilde{d}_h$ be the mapping induced by d_h on T_h. Now, $\tilde{d}_h$ is a bona-fide metric on T_h. Where no ambiguity exists, drop the identification $\tilde{d}_h$ and denote it by d_h. It can be further checked that τ_h is a continuous mapping between $[0, M]$ equipped with the Euclidean norm and T_h equipped with d_h. Thus, should we require a measured metric space, the construction provides a measure μ_h on Borel sets of (T_h, d_h) which is the push-forward of the Lebesgue measure on $[0, M]$ under the mapping τ_h. To link this construction back to our discussion on $\mathbb{R}$-trees, we have the following result.

Theorem 2.3.1. *[47, Theorem 2.2] Suppose $h \in \mathcal{H}$, then the metric space (T_h, d_h) is an $\mathbb{R}$-tree. Set the root as $\rho_h = \tau_h(0)$.*
[27, Lemma 2.1] Additionally, (T_h, d_h) is compact.

Thus, we henceforth denote T_h by $\mathcal{T}_h$ to represent the tree *coded* by h. Conversely, it was proved in [27, Theorem 1.1] that, subject to mild conditions on its Borel measure and ordering of vertices, any weighted $\mathbb{R}$-tree may be encoded by a non-negative càdlàg function on $[0, M]$ with non-positive jumps. Functional encoding can be visualised as applying "glue" under the graph of h, then compressing the graph horizontally to form a tree-like structure representing $\mathcal{T}_h$. Local infima, namely $m_h(s,t)$, become branch points, and the metric $d_h(s,t)$ measures the total *vertical* distance traversed moving from $h(s)$ to $m_h(s,t)$ to $h(t)$. This is illustrated in Figure 1.

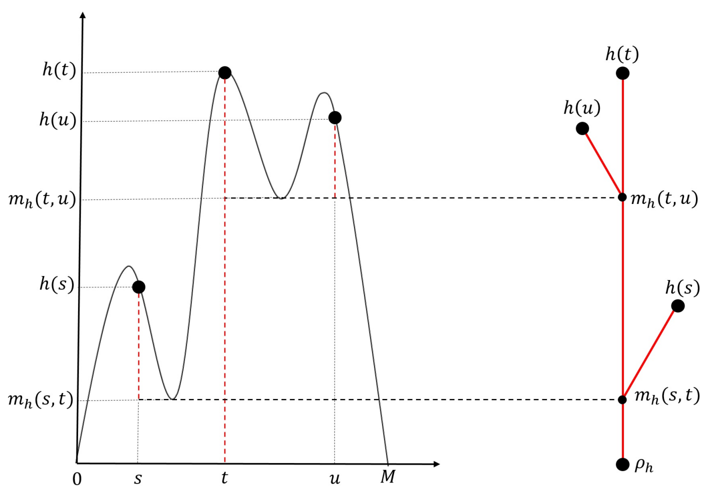

Figure 1: A representation of a sub-tree of $\mathcal{T}_h$ corresponding to $h \in \mathcal{H}_M$ spanned by $s, t, u \in [0, M]$.

2.4 The Continuum Random Tree

In [5, 6, 7], Aldous originally built his theory of a special class of weighted $\mathbb{R}$-trees with infinitesimally short edges, known as *continuum trees*, under an embedding into $l_1(\mathbb{N})$ space, endowing compact subsets with the Hausdorff metric. However, [7, Theorem 3] connects Aldous' $l_1(\mathbb{N})$ embedding and our setup of weighted $\mathbb{R}$-trees. So, we may make the following definition.

Definition 2.4.1. *A weighted $\mathbb{R}$-tree $(\mathcal{T}, d, \rho, \mu)$ is a continuum tree if the Borel probability measure μ satisfies:*

(i) $\mu(\mathcal{L}(\mathcal{T})) = 1$, *that is, μ is supported by the leaves of $\mathcal{T}$,*

(ii) μ *is non-atomic, that is, if $a \in \mathcal{L}(\mathcal{T})$, then $\mu(\{a\}) = 0$,*

(iii) For every $a \in \mathcal{T} \setminus \mathcal{L}(\mathcal{T})$, $\mu(\mathcal{T}_a) > 0$, where $\mathcal{T}_a := \{\sigma \in \mathcal{T} : a \in [\![\rho, \sigma]\!]\}$ is the sub-tree above a in $\mathcal{T}$.

A continuum random tree (CRT) is a random variable valued in the space of continuum trees.

Conditions (i) and (ii) above imply that a continuum tree has uncountably many leaves. Conditions (i) and (iii) imply a continuum tree is leaf-dense [7]. However, simply by the definition of a CRT, it is not obvious how to determine the distributions of CRTs. It is useful to have a notion of *reduced sub-trees.*

Definition 2.4.2. *Given a CRT $(\mathcal{T}, d, \rho, \mu)$ and $m \geq 1$, let $V_1, \ldots, V_m \overset{i.i.d}{\sim} \mu$. Call this a uniform sample of m points according to measure μ. An m-th reduced sub-tree of $(\mathcal{T}, d, \rho, \mu)$ is the sub-tree of $\mathcal{T}$ spanned by $V_1, \ldots, V_m \overset{i.i.d}{\sim} \mu$ and ρ.*

The distribution of an m-th reduced sub-tree is fully specified by its *tree-shape* when regarded as a discrete, graph-theoretic, rooted tree with m labelled leaves, and by its *edge-lengths.* In this sense, the distribution of an m-th reduced sub-tree may be regarded as a random finite-dimensional distribution of a CRT [4]. For example, if $h(s)$, $h(t)$ and $h(u)$ were a uniform sample of 3 points according to μ_h, Figure 1 would represent a 3-rd reduced sub-tree of $\mathcal{T}_h$.

2.5 The Brownian Continuum Random Tree

Following from the previous two sections, define the BCRT to be an $\mathbb{R}$-tree encoded by $h(t) = \sqrt{2}e(t)$ for $0 \leq t \leq 1$, where $(e(t), 0 \leq t \leq 1)$ is a normalised Brownian excursion. As Aldous used $2e(t)$ to define the BCRT, when citing results from [7], we account for this appropriately.

Theorem 2.5.1. *[7, Theorem 13] Let $h\colon [0, 1] \to [0, \infty)$ satisfy:*

(i) $h \in \mathcal{H}_1$ and there is at most one zero of h in $(0, 1)$,

(ii) The set of times of strict local minima of h is dense in $[0, 1]$,

(iii) If $t_1 < t_2$ are strict local minima with $h(t_1) = h(t_2)$, then $\inf_{t_1<t<t_2} h(t) < h(t_1)$,

(iv) The set of times of one-sided local minima of h has Lebesgue measure 0.

Then, the corresponding weighted $\mathbb{R}$-tree $(\mathcal{T}_h, d_h, \rho_h, \mu_h)$ coded by h under the metric d_h defined in (10), rooted at $\rho_h = \tau_h(0)$, with mass measure μ_h being the push-forward of the Lebesgue measure on $[0, 1]$, also satisfies the hypotheses for a CRT in Definition 2.4.1.

Since every local minimum of Brownian motion is a strict local minimum [56, Proposition 9.1] and the set of local minima forms a countable dense set [56, Corollary 9.3], a normalised Brownian excursion satisfies (ii). (iv) follows, since otherwise, with positive probability, we sample a uniform point $u \in [0, 1]$ such that u is a one-sided local minimum. This contradicts the standard fact that Brownian sample paths are of infinite variation on any interval almost surely. As the local minima of Brownian motion are almost surely distinct [56, Lemma 9.2], (iii) is fulfilled. Thus, $h(t) = \sqrt{2}e(t)$ encodes a CRT in Definition 2.4.1. Furthermore, [56, Lemma 9.2] implies that the tree-shapes of reduced sub-trees of $\mathcal{T}_2$ are almost surely binary. Hence, the BCRT is almost surely *binary.*

In [6], Aldous provides three equivalent ways for which the BCRT arises as a scaling limit of discrete structures under appropriate notions of convergence. Of particular interest is a construction, which we shall term Aldous' *line-breaking construction*, outlined below.

(i) Denote by $C_1, C_2, \ldots$ the event times of an inhomogeneous Poisson process on $[0, \infty)$ with rate $\lambda(t) = 2t$. Denote $C_0 = 0$.

(ii) "Cut" $(0, \infty)$ into the intervals $(C_{i-1}, C_i]$ for $i \geq 1$.

(iii) Make 0 the root and start with the interval $[0, C_1]$. Then, inductively for $i = 2, 3, \ldots$, choose a point J_i uniformly over the length of the existing tree edges. Identify C_{i-1} and J_i, regard $(C_{i-1}, C_i]$ as a branch connected at J_i.

(iv) The desired process is the closure of the union of all branches regarded in $\mathbb{T}$. A length measure on the BCRT may be recovered as the almost sure weak limit of the normalised length measure on the edges at each step.

This construction enables us to construct a family of random discrete binary trees.

Proposition 2.5.2. *[7, Lemma 21] There exists a family $(\mathcal{R}(k) : k \geq 1)$ of random binary trees such that $\mathcal{R}(k)$ has tree-shape $\hat{t}$ and edge-lengths $(x_1, \ldots, x_{2k-1})$, with joint density on $\{(\hat{t}, x_1, \ldots, x_{2k-1}) : x_1 > 0, \ldots, x_{2k-1} > 0, \hat{t} \in \hat{T}_k\}$ given by*

$$f\left(\hat{t}, x_1, \ldots, x_{2k-1}\right) = 2^k s \exp\left(-s^2\right),$$

where $s = \sum_{i=1}^{2k-1} x_i$ and $\hat{T}_k$ is the set of all rooted binary tree-shapes with k labelled leaves. The density is well-defined by the identity

$$\int_0^\infty \cdots \int_0^\infty 2^k s \exp\left(-s^2\right) dx_1 \cdots dx_{2k-1} = \prod_{i=1}^{k-1} \frac{1}{2i-1},$$

and the fact that there are $\prod_{i=1}^{k-1}(2i-1)$ rooted binary tree-shapes with k labelled leaves. In particular, the edge-lengths have an exchangeable joint distribution, independent of the tree-shape.

Conversely, the BCRT may be defined as the CRT which has this family as its reduced sub-trees [7, Corollary 22]. In particular, the distance between two points chosen uniformly according to the mass measure of the BCRT has a *Rayleigh distribution with scale parameter* $1/\sqrt{2}$, denoted by Rayleigh $\left(1/\sqrt{2}\right)$. It has density

$$2x \exp(-x^2) \quad \text{for } x > 0, \tag{11}$$

with expectation $\sqrt{\pi}/2$. As the BCRT is invariant under uniform re-rooting [41, Theorem 11], the root may be regarded as a uniform sample from the mass measure itself. Hence, (11) is the distribution of the distance between the root and a uniformly sampled point on the BCRT.

2.6 α-stable Trees

As with the BCRT, an α-stable tree may be functionally encoded. We provide a brief summary, and refer to [29, Chapter 1] for technical details. Fix $\alpha \in (1, 2]$. Consider a stable Lévy process $X = (X_t, t \geq 0)$ given by the Laplace exponent $\psi(\lambda) = \log \mathbb{E}\left[\exp\left(-\lambda X_1\right)\right] = \lambda^\alpha$. From [46, Section 1.2.6], an important property of X is the *scaling property*

$$\left(\lambda^{-1/\alpha} X_{\lambda t}, t \geq 0\right) \stackrel{d}{=} (X_t, t \geq 0) \quad \text{for all } \lambda > 0. \tag{12}$$

For every $t > 0$, let $\hat{X}^{(t)}$ be the process

$$\hat{X}_s^{(t)} = \begin{cases} X_t - X_{(t-s)^-} & \text{if } 0 \leq s < t, \\ X_t & \text{if } s = t, \end{cases} \tag{13}$$

and denote $\hat{S}^{(t)}$ as its associated supremum process. The *height process* $H = (H_t, t \geq 0)$ is the real-valued process such that $H_0 = 0$, and H_t is the local time at level 0 at $t > 0$ of the process $\hat{X}^{(t)} - \hat{S}^{(t)}$. Informally, H_t measures the set $\{s \leq t : X_s = \inf_{r \in [s,t]} X_r\}$. By [29, Theorem 1.4.3], our choice of ψ guarantees that, almost surely, the height process admits a continuous modification. As a consequence of the scaling property (12), the height process also satisfies an analogous scaling property, from [29, Corollary 3.3.2],

$$\left(\lambda^{1/\alpha - 1} H_{\lambda t}, t \geq 0\right) \stackrel{d}{=} (H_t, t \geq 0) \quad \text{for all } \lambda > 0. \tag{14}$$

The scaling property in (14) enables definition of the *law of the normalised excursion* of H, denoted by $N^{(1)}$. We do not delve into details here. As H is continuous, we may define the quantities

$$L := \sup\{t \in (0,1) : H_t = 0\} \quad \text{and} \quad R := \inf\{t > 1 : H_t = 0\}. \tag{15}$$

We take $N^{(1)}$ to be the law of

$$\frac{1}{(R-L)^{1-1/\alpha}} \left| H_{L+(R-L)t} \right| \quad \text{for } t \in (0,1). \tag{16}$$

Under this law, sample paths are in $\mathcal{H}_1$, and so by Theorem 2.3.1, encode a compact $\mathbb{R}$-tree almost surely. We recover $\sqrt{2}$ times a normalised Brownian excursion when $\alpha = 2$.

Definition 2.6.1. *Fix $\alpha \in (1, 2]$. An $\mathbb{R}$-tree coded by the height process H under $N^{(1)}$ is an α-stable tree (which is normalised). We call the measure μ_α, the usual push-foward of the Lebesgue measure on $[0,1]$ under the canonical projection, the mass measure. We denote the above by $(\mathcal{T}_\alpha, d_\alpha, \rho_\alpha, \mu_\alpha)$, or simply by $\mathcal{T}_\alpha$, when no ambiguity exists.*

It is verified in [28, Display (28)] that $N^{(1)}$-almost surely, an α-stable tree equipped with its mass measure satisfies the conditions of a CRT in Definition 2.4.1.

Furthermore, (14) extends to an appropriate scaling for an α-stable tree [48, Proposition 4.3]. Where the context is clear, given $\alpha \in (1, 2]$, denote $\beta := 1 - 1/\alpha \in (0, 1/2]$. We deduce from (14) that rescaling the mass measure by $\xi > 0$ leads to lengths being rescaled by ξ^β, and *vice versa*. Denote by $\xi^\beta \mathcal{T}_\alpha$, a *rescaled α-stable tree*, an α-stable tree with associated parameters $\left(\mathcal{T}_\alpha, \xi^\beta d_\alpha, \rho_\alpha, \xi \mu_\alpha\right)$.

The quantity β arises as the negative of the *index* in Miermont's self-similar fragmentation process [52, 53], and the scaling factor for the aforementioned Galton-Watson trees conditioned on fixed progeny n. Write such a tree as T_n^{GW} and endow it with the graph distance. Aldous in [7, Theorem 23] proves for the BCRT, if the offspring distribution η has mean 1 and variance $\sigma^2 \in (0, \infty)$, that

$$\frac{T_n^{GW}}{\sqrt{n}} \xrightarrow{d} \frac{\sqrt{2}}{\sigma} \mathcal{T}_2 \quad \text{as } n \to \infty, \tag{17}$$

in the Gromov-Hausdorff topology. Duquesne in [26, Theorem 3.1] proves in general that if the offspring distribution has mean 1 and $\eta(k) \sim Ck^{-1-\alpha}$ as $k \to \infty$, then

$$\frac{T_n^{GW}}{n^\beta} \xrightarrow{d} \left(\frac{\alpha(\alpha-1)}{C\Gamma(2-\alpha)}\right)^{1/\alpha} \mathcal{T}_\alpha \quad \text{as } n \to \infty, \tag{18}$$

in the Gromov-Hausdorff topology.

One might then wonder if there is a corresponding line-breaking construction for α-stable trees. Goldschmidt and Haas in [36] provide two equivalent line-breaking constructions reminiscent of Aldous' construction. They recover the length measure of an α-stable tree by generalising the inhomogeneous Poisson process for the BCRT to a Mittag-Leffler Markov chain. We introduce the Mittag-Leffler distribution in Section 3.2. Marchal's random growth algorithm in Section 3.3 provides another construction that gives an α-stable tree as a scaling limit with its mass measure in Gromov-Hausdorff-Prokhorov topology.

Another key property of α-stable trees is that its distribution is fully specified by the joint distribution of its reduced sub-trees with points sampled according to mass measure μ_α. The BCRT case was verified in [7, Corollary 22] and the general α-stable tree case in [29, Theorem 2.2.1]. For the BCRT, we had an explicit result on the distribution of its reduced sub-trees in Proposition 2.5.2. We cite an analogous result for the α-stable tree, which gives exchangeability of edge-lengths, but with dependence on tree-shape.

Theorem 2.6.2. *[29, Theorem 3.3.3] Suppose $\alpha \in (1,2)$, $m \geq 2$, and $\mathbf{t}_m$ is the tree-shape of a given rooted tree with m labelled leaves. The probability of the tree-shape of an m-th reduced sub-tree of $\mathcal{T}_\alpha$ being $\mathbf{t}_m$ is*

$$\frac{\prod_{v\in\mathbf{t}_m:\deg(v)\geq 3}\left|\prod_{i=1}^{\deg(v)-2}(\alpha-i)\right|}{\prod_{i=1}^{m-1}(i\alpha-1)}.$$

Conditional on tree-shape $\mathbf{t}_m$, denoting the edge set of $\mathbf{t}_m$ by $\mathcal{E}$, the edge-lengths have a joint Lebesgue density supported on $\mathbb{R}_+^{|\mathcal{E}|}$ given by

$$\frac{\Gamma\left(m-1/\alpha\right)}{\Gamma\left(\theta\right)}\alpha^{|\mathcal{E}|}\int_0^1 u^{\theta-1}q(\alpha s,1-u)du,$$

where $\theta = m-\beta|\mathcal{E}|-1/\alpha > 0$, s denotes the sum of edge-lengths, and $q(v,u)$ is the continuous density at time v of the stable subordinator with exponent β, which is characterised by

$$\int_0^\infty e^{-\lambda u}q(v,u)du = \exp\left(-v\lambda^\beta\right).$$

In particular, conditional on the tree-shape of a reduced sub-tree of $\mathcal{T}_\alpha$, the edge-lengths have an exchangeable joint distribution.

[29, Theorem 3.3.3] also determines that, when $m = 1$ (so that tree-shape is deterministic), the distance between the root and a uniformly sampled point has a distribution with density

$$\alpha\Gamma(\beta)q(\alpha x,1) \quad \text{for } x > 0. \tag{19}$$

When $\alpha = 2$, $q(v,u)$ has an explicit form, given in [29] by

$$q(v,u) = \frac{v}{2\sqrt{\pi}u^{3/2}}\exp\left(-\frac{v^2}{4u}\right) \Rightarrow 2\Gamma\left(\frac{1}{2}\right)q(2x,1) = 2x\exp(-x^2).$$

Thus, (19) agrees with (11) earlier.

We conclude with a pivotal difference between the BCRT and an α-stable tree with $\alpha \in (1,2)$. Recall that the BCRT is binary almost surely. However, when $\alpha \in (1,2)$, the degree of a branch point in an α-stable tree is infinite almost surely, see [48, Proposition 5.2]. We call such a tree *infinitary*. This fact can be proved by invoking properties of the height process H [52, Section 2.2]. However, we provide an alternative proof of this using Marchal's random growth algorithm and the Chinese Restaurant Process in Proposition 4.1.2.

3 Background on Tools Required

We give rigorous setups of RDEs, a specific Pólya urn model, the Chinese Restaurant Process and Marchal's random growth algorithm.

3.1 Recursive Distributional Equations

It is instructive to review RDEs in their full generality, as presented in [9, Section 2.1], since we work over abstract topological spaces. Denote our underlying probability space by $(\Omega, \mathcal{F}, \mathbb{P})$. Given two measurable spaces $(\mathbb{S}, \mathcal{F}_{\mathbb{S}})$ and $(\Theta, \mathcal{F}_{\Theta})$, construct the product space

$$\Theta^* := \Theta \times \bigcup_{0 \leq m \leq \infty} \mathbb{S}^m, \tag{20}$$

where the union is a disjoint union over $\mathbb{S}^m$, the space of $\mathbb{S}$-valued sequences of lengths $0 \leq m \leq \infty$, where $\mathbb{S}^0 := \{\Delta\}$ is the singleton set and $\mathbb{S}^\infty$ constructed as a typical sequence space.

Equip Θ^* with the product sigma-algebra. Let $g \colon \Theta^* \to \mathbb{S}$ be a measurable map, and define random variables $(\mathcal{S}_i : i \geq 0) \in \mathbb{S}^\infty$, $(\xi, N) \in \Theta \times \overline{\mathbb{N}} := \Theta \times \{0, 1, \ldots; \infty\}$ as follows

(i) $(\xi, N) \sim \nu$, where ν is a probability measure on $\Theta \times \overline{\mathbb{N}}$,

(ii) $\mathcal{S}_i \overset{i.i.d}{\sim} \eta$, where η is a a probability measure on $\mathbb{S}$,

(iii) (ξ, N) and $(\mathcal{S}_i : i \geq 0)$ are independent.

Denote by $\mathcal{P}(\mathbb{S})$ the set of probability measures on $(\mathbb{S}, \mathcal{F}_{\mathbb{S}})$. Given the distribution ν on $\Theta \times \overline{\mathbb{N}}$, we obtain a mapping

$$\begin{aligned} &\Phi \colon \mathcal{P}(\mathbb{S}) \to \mathcal{P}(\mathbb{S}), \qquad \eta \mapsto \Phi(\eta), \\ &\text{where } \Phi(\eta) \text{ is the distribution of } \mathcal{S} := g(\xi, \mathcal{S}_i : 0 \leq i \leq^* N), \end{aligned} \tag{21}$$

and where the notation $\leq^* N$ means $\leq N$ for $N < \infty$ and $< \infty$ for $N = \infty$. This lends itself to a fixpoint perspective of RDEs, where we find a distribution of $\mathcal{S}$ such that

$$\eta = \Phi(\eta) \iff \mathcal{S} \overset{d}{=} g(\xi, \mathcal{S}_i : 0 \leq i \leq^* N) \quad \text{ on } \mathbb{S}. \tag{22}$$

3.2 A Pólya Urn Model and the Chinese Restaurant Process

Given parameters $\beta_1, \beta_2, \ldots, \beta_n > 0$, define the *Dirichlet distribution*, $\mathrm{Dir}(\beta_1, \beta_2, \ldots, \beta_n)$, as the distribution with density

$$\frac{\Gamma\left(\sum_{i=1}^n \beta_i\right)}{\prod_{i=1}^n \Gamma(\beta_i)} \left(\prod_{i=1}^{n-1} x_i^{\beta_i - 1}\right) \left(1 - \sum_{i=1}^{n-1} x_i\right)^{\beta_n - 1} \tag{23}$$

with respect to the Lebesgue measure on $\{(x_1, x_2, \ldots, x_{n-1}) \in [0,1]^{n-1} : \sum_{i=1}^{n-1} x_i < 1\}$. A *Beta distribution* with parameters $\beta_1, \beta_2 > 0$ has distribution $\mathrm{Dir}(\beta_1, \beta_2)$.

Given parameters $\gamma, \lambda > 0$, the *Gamma distribution*, $\mathrm{Gamma}(\gamma, \lambda)$, has density

$$\frac{\lambda^\gamma}{\Gamma(\gamma)} x^{\gamma - 1} \exp(-\lambda x) \quad \text{ for } x > 0. \tag{24}$$

Given $\beta > 0$ and $\theta > -\beta$, a random variable L valued in $[0, \infty)$ has a *generalised Mittag-Leffler distribution* with parameters (β, θ), denoted by $L \sim \mathrm{ML}(\beta, \theta)$, if it has p-th moment

$$\mathbb{E}[L^p] = \frac{\Gamma(\theta+1)\Gamma(\theta/\beta+1+p)}{\Gamma(\theta/\beta+1)\Gamma(\theta+\beta p+1)}, \quad \text{for } p \geq 1. \tag{25}$$

[58, Display (0.42)] checks that the collection of moments determine this distribution uniquely. Employing Theorem 2.6.2, [3, Lemma 11] shows that α times the distance between two uniformly sampled points on an α-stable tree characterises a size-biased $\mathrm{ML}(\beta, 0)$ distribution. Suppose $L \sim \mathrm{ML}(\beta, 0)$, denote by $\tilde{L}$ the size-biased distribution of L. Then, for $p \geq 1$,

$$\mathbb{E}\left[\tilde{L}^p\right] = \frac{\mathbb{E}\left[L^{p+1}\right]}{\mathbb{E}[L]} = \frac{\Gamma(\beta+1)\Gamma(p+2)}{\Gamma(\beta p+\beta+1)} = \frac{\Gamma(\beta+1)\Gamma(\beta/\beta+1+p)}{\Gamma(\beta/\beta+1)\Gamma(\beta+\beta p+1)}. \tag{26}$$

Therefore, a size-biased $\mathrm{ML}(\beta, 0)$ distribution is $\mathrm{ML}(\beta, \beta)$ distributed. As the α-stable tree remains invariant under uniform re-rooting [41, Theorem 11], this is the distribution of α times the distance between the root and a uniformly sampled point. [36, Section 1.1] gives the distributional identity

$$\frac{1}{2}\mathrm{ML}\left(\frac{1}{2}, \frac{1}{2}\right) \stackrel{d}{=} \sqrt{\mathrm{Gamma}(1,1)}. \tag{27}$$

Note that (27) agrees with (11) earlier, as a standard change of variables argument gives

$$\sqrt{\mathrm{Gamma}(1,1)} \stackrel{d}{=} \mathrm{Rayleigh}\left(1/\sqrt{2}\right).$$

We cite an identity between the Dirichlet and Mittag-Leffler distributions.

Proposition 3.2.1. *[36, Proposition 4.2] For $n \geq 2$, let $\beta \in (0,1)$, $\theta_1, \theta_2, \ldots, \theta_n > 0$ and $\theta := \sum_{i=1}^n \theta_i$. If $L \sim \mathrm{ML}(\beta, \theta)$ and $(Y_1, Y_2, \ldots, Y_n) \sim \mathrm{Dir}(\theta_1/\beta, \theta_2/\beta, \ldots, \theta_n/\beta)$ are independent, then*

$$L \cdot (Y_1, Y_2, \ldots, Y_n) \stackrel{d}{=} \left(X_1^\beta M^{(1)}, X_2^\beta M^{(2)}, \ldots, X_n^\beta M^{(n)}\right)$$

where $(X_1, X_2, \ldots, X_n) \sim \mathrm{Dir}(\theta_1, \theta_2, \ldots, \theta_n)$ and $M^{(i)} \sim \mathrm{ML}(\beta, \theta_i)$, $1 \leq i \leq n$, are independent.

The result above is vital in linking mass and length splits across branches of the α-stable tree [60]. We prove the *aggregation property* of the Dirichlet distribution.

Proposition 3.2.2. *Suppose $\beta_1, \ldots, \beta_n > 0$ and $Y := (Y_1, \ldots, Y_n) \sim \mathrm{Dir}(\beta_1, \ldots, \beta_n)$ and given $1 \leq m \leq n-1$, define $Y' := (\sum_{i=1}^m Y_i, Y_{m+1}, \ldots, Y_n)$. Then, $Y' \sim \mathrm{Dir}(\sum_{i=1}^m \beta_i, \beta_{m+1}, \ldots, \beta_n)$.*

Proof: Using the Gamma distribution representation of the Dirichlet distribution, given $W_i \sim \mathrm{Gamma}(\beta_i, \lambda)$ independent for $i = 1, \ldots, n$ and for some $\lambda > 0$, we may write $Y_i = \frac{W_i}{\sum_{j=1}^n W_j}$. Furthermore, $W' := \sum_{i=1}^m W_i \sim \mathrm{Gamma}(\sum_{i=1}^m \beta_i, \lambda)$ by properties of independent Gamma variables, and is independent of $(W_{m+1}, \ldots, W_n)$. Applying the Gamma distribution representation to $(W', W_{m+1}, \ldots, W_n)$ gives the desired distribution of Y'. □

Dirichlet and Mittag-Leffler distributions arise naturally in a variety of urn models, see [58] and [44] respectively. For our purposes, we restrict attention to the following specification of Pólya's urn model.

Definition 3.2.3. *Given $K \geq 2$ and a fixed vector $\vec{\beta} = (\beta_1, \beta_2, \ldots, \beta_K)$, where $\beta_1, \beta_2, \ldots, \beta_K > 0$, consider a Pólya urn scheme with K colours, initialisation $\vec{\beta}$ and step-size $\alpha > 0$ evolving in discrete time. Represent the K colours by the set $C := \{1, 2, \ldots, K\}$. We say that a random variable X valued in C has distribution $X \sim \mathrm{PMF}(\vec{\beta})$ if $\mathbb{P}(X = i) = \frac{\beta_i}{\sum_{i=1}^K \beta_i}$ for all $i \in C$. Generate a sequence of draws $(X_1, X_2, \ldots,)$ from C according to the following scheme:*

- ***STEP 1:*** *Set $\vec{\beta}_1 := \vec{\beta}$, sample $X_1 \sim \mathrm{PMF}(\vec{\beta}_1)$.*
- ***STEP 2:*** *For $n \geq 1$, set $\vec{\beta}_{n+1} := \vec{\beta}_n + \alpha \vec{e}_{X_n}$, where $\vec{e}_j$ denotes the j-th standard Euclidean basis vector of $\mathbb{R}^K$. Sample $X_{n+1} \mid X_1, \ldots, X_n \sim \mathrm{PMF}(\vec{\beta}_{n+1})$.*

For $j \in C$, denote the number of j-th coloured balls observed after n draws by $D_j^{(n)} := \sum_{i=1}^{n} \mathbf{1}\{X_i = j\}$. Define the vector of frequency of colours observed after n draws as

$$\left(P_1^{(n)}, P_2^{(n)}, \ldots, P_K^{(n)}\right) := \left(\frac{D_1^{(n)}}{n}, \frac{D_2^{(n)}}{n}, \ldots, \frac{D_K^{(n)}}{n}\right).$$

The asymptotic frequencies of colours observed are known to converge to an almost sure limit, due to Blackwell and MacQueen [13]. The original result was proved for a Ferguson distribution defined on a Polish metric space X. For our purpose, set $X = C$, the finite set of K colours. We then have the following result.

Theorem 3.2.4. *[13, Theorem 1] Given a sequence of draws $(X_1, X_2, \ldots)$ from the urn scheme in Definition 3.2.3, the asymptotic frequencies satisfy*

$$\left(P_1^{(n)}, P_2^{(n)}, \ldots, P_K^{(n)}\right) \xrightarrow{a.s.} (P_1, P_2, \ldots, P_K) \qquad \text{as } n \to \infty,$$

where $(P_1, P_2, \ldots, P_K) \sim \mathrm{Dir}\left(\frac{\beta_1}{\alpha}, \frac{\beta_2}{\alpha}, \ldots, \frac{\beta_K}{\alpha}\right)$. Consequently, the proportions of colours also converge as follows,

$$\left(\frac{\beta_1 + \alpha D_1^{(n)}}{\sum_{i=1}^K \beta_i + \alpha n}, \frac{\beta_2 + \alpha D_2^{(n)}}{\sum_{i=1}^K \beta_i + \alpha n}, \ldots, \frac{\beta_K + \alpha D_K^{(n)}}{\sum_{i=1}^K \beta_i + \alpha n}\right) \xrightarrow{a.s.} (P_1, P_2, \ldots, P_K) \qquad \text{as } n \to \infty.$$

A natural extension to the urn scheme introduced in Definition 3.2.3 is the *Chinese Restaurant Process* (CRP) due to Dubins and Pitman [57]. We shall only be interested in a specific unordered two-parameter CRP model.

Definition 3.2.5. *Given $\beta \in [0, 1]$ and $\theta > -\beta$, the two-parameter Chinese Restaurant Process with a (β, θ) seating plan, denoted by $\mathrm{CRP}(\beta, \theta)$, proceeds as follows. Label customers by $n \geq 1$. Seat customer 1 at the first table. For $n \geq 1$, let K_n denote the number of tables occupied after customer n has been seated and let $L_j(n)$ denote the number of customers seated at the j-th table for $j \in \{1, \ldots, K_n\}$. At the next arrival, conditional on $(L_1(n), \ldots, L_j(n))$, customer $n + 1$*

- *sits at the j-th table with probability $(L_j(n) - \beta) / (n + \theta)$ for $j \in \{1, \ldots, K_n\}$,*
- *opens a $(K_n + 1)$-st table with the complementary probability $(\theta + K_n\beta) / (n + \theta)$.*

For each $n \geq 1$, the process at step n induces a partition of the set $\{1, \ldots, n\}$ into blocks, denoted $\Pi_n := (\Pi_{n,1}, \ldots, \Pi_{n,K_n})$, given by the collection of customer labels at each occupied table, and with blocks ordered by least labels. Define $\mathrm{CRP}\,(\beta, \theta) := (\Pi_n : n \geq 1)$ *to be this partition-valued process.*

As with Pólya urn schemes, the CRP also satisfies limit theorems associated with the Dirichlet and Mittag-Leffler distributions.

Theorem 3.2.6. *[58, Theorem 3.2 and Theorem 3.8] In the context of* $\mathrm{CRP}\,(\beta, \theta)$*, the rescaled number of tables occupied at step n satisfies*

$$\frac{K_n}{n^\beta} \xrightarrow{a.s.} K_\infty \qquad \text{as } n \to \infty,$$

where $K_\infty \sim \mathrm{ML}(\beta, \theta)$. Furthermore, the asymptotic proportion of relative table sizes has an almost sure limit, in size-biased order of least labels,

$$\left(\frac{L_1(n)}{n}, \frac{L_2(n)}{n}, \ldots, \frac{L_{K_n}(n)}{n}\right) \xrightarrow{a.s.} \left(W_1, \overline{W}_1 W_2, \overline{W}_1 \overline{W}_2 W_3, \ldots\right) \qquad \text{as } n \to \infty,$$

where W_j are independent with $W_j \sim \mathrm{Beta}\,(1-\beta, \theta + j\beta)$, and $\overline{W}_j := 1 - W_j$ for all $j \geq 1$.

Definition 3.2.7. *The distribution of the vector $(P_1, P_2, P_3, \ldots) := \left(W_1, \overline{W}_1 W_2, \overline{W}_1 \overline{W}_2 W_3, \ldots\right)$ as defined in Theorem 3.2.6 is a Griffiths-Engen-McCloskey distribution with parameters (β, θ), denoted* $\mathrm{GEM}(\beta, \theta)$*. Ordering $(P_i : i \geq 1)$ in decreasing order, we arrive at a Poisson-Dirichlet distribution with parameters (β, θ) given by $(P_i^\downarrow : i \geq 1) := (P_i : i \geq 1)^\downarrow$, denoted* $\mathrm{PD}(\beta, \theta)$*.*

3.3 Marchal's Random Growth Algorithm

Marchal's *random growth algorithm* generalises Rémy's algorithm [62], and follows from Marchal's earlier work on the Lukasiewicz correspondence of random trees to excursions of a simple random walk converging to a Brownian excursion [50]. We adapt the notation employed in Curien and Haas [21] in the following.

Definition 3.3.1. *[51, Marchal's random growth algorithm] Given a parameter $\alpha \in (1, 2]$, the following recursive construction yields a discrete time Markov chain $(\mathbf{T}_\alpha(n))_{n \geq 1}$ valued in the set of leaf-labelled trees, with $\mathbf{T}_\alpha(n)$ having n leaves and a root.*

- ***STEP 1:*** *Initialise $\mathbf{T}_\alpha(1)$ as the unique tree with one edge and two labelled endpoints, A_0 and A_1. Regard A_0 as a root and A_1 as a marked leaf.*

- ***STEP 2:*** *For $n \geq 1$, conditional on $\mathbf{T}_\alpha(n)$, assign a weight of $\alpha - 1$ to any edge of $\mathbf{T}_\alpha(n)$, a weight $d - 1 - \alpha$ to any vertex of degree $d \geq 3$ (a branch point), and no weight to other vertices. Choose an edge or a branch point of $\mathbf{T}_\alpha(n)$ with probability proportional to its weight.*

- ***STEP 3a:*** *If an edge was selected in **STEP 2**, split the chosen edge into two edges at its midpoint by a middle vertex, denoted V_{n+1}. At V_{n+1}, attach a new edge carrying the $(n+1)$-st leaf, denoted A_{n+1}.*

- ***STEP 3b:*** *If a branch point was selected in **STEP 2**, attach a new edge carrying the $(n+1)$-st leaf at the chosen vertex. Denote the new leaf as A_{n+1}.*

- ***STEP 4:*** *Repeat* ***STEP 2*** *with* $n \mapsto n+1$.

Denoting $\tilde{I} := \{k : V_k \text{ is created}\}$, *define the limiting set of vertices at time* ∞ *as*

$$\mathbf{T}_\alpha(\infty) := \bigcup_{n\geq 0}\{A_n\} \cup \bigcup_{n\in\tilde{I}}\{V_n\}.$$

Define the operator $W(\cdot)$ which measures the total weight of a given sub-structure in Marchal's algorithm. Marchal's algorithm immediately yields the following results on weights and tree-shapes.

Proposition 3.3.2. *Regardless of tree-shape, for all* $n \geq 1$, *the total weight of the tree is* $W(\mathbf{T}_\alpha(n)) = n\alpha - 1$.

Proof: Induct on n. The base case $n = 1$ is clear. Suppose $n \geq 1$. By the induction hypothesis, $W(\mathbf{T}_\alpha(n)) = n\alpha - 1$, independently of tree-shape. Consider the next step in Marchal's algorithm. If an edge is selected in $\mathbf{T}_\alpha(n)$, the number of edges increases by 2 and we gain one branch point of degree 3. Thus, $W(\mathbf{T}_\alpha(n+1)) = W(\mathbf{T}_\alpha(n)) + 2(\alpha-1) + (3-1-\alpha) = (n+1)\alpha - 1$. If a branch point of degree $d \geq 3$ is selected, then we gain one edge and increase the weight of that branch point by 1. Thus, $W(\mathbf{T}_\alpha(n+1)) = W(\mathbf{T}_\alpha(n)) + (\alpha-1) + 1 = (n+1)\alpha - 1$. Hence, $W(\mathbf{T}_\alpha(n+1)) = (n+1)\alpha - 1$ holds regardless of tree-shape. The induction proceeds. □

Theorem 3.3.3. *[51, Theorem 1] Suppose* $\mathbf{t}_0(n)$ *is a given leaf-labelled tree with* n *leaves and a root, where* $n \geq 2$, *then the tree-shape of* $\mathbf{T}_\alpha(n)$ *has distribution*

$$\mathbb{P}(\mathbf{T}_\alpha(n) = \mathbf{t}_0(n)) = \frac{\prod_{v\in\mathbf{t}_0(n)} p_{\deg(v)}}{\prod_{i=1}^{n-1}(i\alpha - 1)},$$

where $p_1 = 1$, $p_2 = 0$, *and* $p_k = \left|\prod_{i=1}^{k-2}(\alpha - i)\right|$ *for* $k \geq 3$.

Proof: We induct on n. The base case $n = 2$ is immediate. Suppose the result holds at time $n \geq 2$. To get the distribution of the tree-shape at time $n+1$, we consider two possibilities. Suppose the algorithm selects an edge at time n, conditional on the tree-shape at time n, the probability of this occurring is $(\alpha-1)/(n\alpha-1)$ from Proposition 3.3.2. We get the desired distribution of tree-shapes of $\mathbf{T}_\alpha(n+1)$ by multiplying this with the expression in the induction hypothesis. Otherwise, the algorithm selects a pre-existing branch point of degree $d \geq 3$. Conditional on the tree-shape at time n, this occurs with probability $(d-1-\alpha)/(n\alpha-1)$. This gives the correct expression at time $n+1$. □

This agrees with the law of tree-shapes of the n-th reduced sub-trees of the α-stable tree in Theorem 2.6.2. In the limit, a subtlety of Marchal's algorithm is that, almost surely, no two vertices chosen from $\mathbf{T}_\alpha(\infty)$ are adjacent. Suppose u and v are two vertices incident to edge e at time n_0, then almost surely, we observe (countably) infinitely many branch points added to e as Marchal's algorithm progresses. To see this, for $n > n_0$, define A_n as the event in which Step 2 of Marchal's algorithm selects the edge incident to u which is a subset of e at time $(n+1)$. By Marchal's algorithm, $\{A_n : n > n_0\}$ are independent. From Proposition 3.3.2, regardless of tree-shape,

$$\mathbb{P}(A_n) = \frac{\alpha - 1}{n\alpha - 1}. \tag{28}$$

Since the harmonic series diverges, by the second Borel-Cantelli lemma,

$$\mathbb{P}\left(\text{infinitely many branch points added to } e\right) \geq \mathbb{P}\left(\bigcap_{i\geq n_0+1}\bigcup_{n\geq i} A_n\right) = 1.$$

To ensure the limiting object is an $\mathbb{R}$-tree, we take the natural *completion* of $\mathbf{T}_\alpha(\infty)$ by "filling in-between" the countably many pairwise non-adjacent vertices. Also, between two chosen points $u, v \in \mathbf{T}_\alpha(n)$, the above entails that the graph distance between them tends to infinity as $n \to \infty$. We need to rescale this appropriately by a quantity converging to infinity as $n \to \infty$. By identifying a suitable L^2-bounded martingale and invoking the Martingale Convergence Theorem, Marchal demonstrates the following limiting behaviour.

Theorem 3.3.4. *[51, Theorem 2] For every $u, v \in \mathbf{T}_\alpha(\infty)$, the following limit exists almost surely*

$$d(u,v) = \lim_{n\to\infty} \frac{d_n(u,v)}{n^\beta},$$

where d_n is the graph distance on $\mathbf{T}_\alpha(n)$ and $\beta = 1 - 1/\alpha$. Since the four points condition holds trivially for discrete trees, then in the limit, $\left(\overline{\mathbf{T}_\alpha(\infty)}, d\right)$ is an $\mathbb{R}$-tree by Theorem 2.1.2.

We interpret $\overline{\mathbf{T}_\alpha(\infty)}$ as the completion of $\mathbf{T}_\alpha(\infty)$ described above. Thus, it makes sense to regard the *scaling limit* of Marchal's algorithm as an $\mathbb{R}$-tree. Combining these observations, if $(T_\alpha(n))_{n\geq 1}$ is a realisation of $(\mathbf{T}_\alpha(n))_{n\geq 1}$, then the following limit holds in the sense of convergence of finite-dimensional distributions of reduced sub-trees,

$$\frac{T_\alpha(n)}{\alpha n^\beta} \to \mathcal{T}_\alpha \quad \text{as } n \to \infty. \tag{29}$$

[40, Corollary 24] checks that $\mathcal{T}_\alpha$ may be constructed on the same probability space supporting $(\mathbf{T}_\alpha(n))_{n\geq 1}$ with the convergence in (29) holding in probability in the Gromov-Hausdorff sense. We state an improved result by Curien and Haas.

Theorem 3.3.5. *[21, Theorem 5(iii)] Denote by μ_n the empirical mass measure on the leaves of $T_\alpha(n)$, and by d_n the graph distance on $T_\alpha(n)$, then*

$$\left(T_\alpha(n), \frac{d_n}{\alpha n^\beta}, \mu_n\right) \xrightarrow{a.s.} (\mathcal{T}_\alpha, d_\alpha, \mu_\alpha) \qquad \text{as } n \to \infty,$$

in the Gromov-Hausdorff-Prokhorov topology.

A consequence when $\alpha = 2$ is that no weight is ever given to a vertex of $\mathbf{T}_2(n)$ in Step 2 of Marchal's algorithm, for any $n \geq 1$. In the scaling limit, this coheres with the fact that $\mathcal{T}_2$ is binary almost surely.

At this juncture, it is instructive to introduce further developments in analogous constructions of α-stable trees, and more general trees, based on Marchal's algorithm.

Marchal's algorithm is a special case of Chen, Ford and Winkel's *alpha-gamma model* [18]. The alpha-gamma model allows further discrimination between edges adjacent to a leaf (external edges) and the remaining internal edges.

The distribution of the sequence of tree-shapes obtained from Goldschmidt and Haas' line-breaking construction is equivalent to that obtained by Marchal's algorithm [36, Proposition 3.7]. However, Goldschmidt and Haas' constructions focus on distributions of *edge-lengths* rather than *mass* in an α-stable tree.

Recently, Rembart and Winkel introduced a *two-colour line-breaking construction* [60, Algorithm 1.3] which unifies aspects of the alpha-gamma model, and Goldschmidt and Haas' line-breaking construction. It ascribes a notion of length to the weights at branch points of Goldschmidt and Haas' line-breaking algorithm by growing trees at these branch points. Specifically, [60, Theorem 1.4] demonstrates the relationship between Dirichlet mass splits and Mittag-Leffler variables governing the edge-lengths in an α-stable tree, described by Proposition 3.2.1.

However, little emphasis has been placed on the recursive nature of Marchal's algorithm *per se*. In [21], Curien and Haas exploit this property to demonstrate a pruning procedure to obtain a rescaled α'-stable tree from an α-stable tree, where $1 < \alpha < \alpha' \leq 2$. They identified sub-constructions within Marchal's algorithm with parameter α that evolve as a time-changed Marchal's algorithm with parameter α'. In this vein, we utilise their approach to formulate an RDE where the law of the α-stable tree is a solution.

4 Recursive Construction of α-stable Trees

In this section, fix $\alpha \in (1, 2]$. Unless ambiguity arises, we suppress α hereafter. Observe that, in Marchal's algorithm, $\mathbf{T}(2)$ is deterministic, comprising a "Y"-shape with three leaves A_0, A_1 and A_2 and an internal vertex V_1. Denote the edges by $e_0 := [\![A_0, V_1]\!]$, $e_1 := [\![A_1, V_1]\!]$ and $e_2 := [\![A_2, V_1]\!]$. The following heuristic, implicitly employed in the proof of [21, Proposition 10], outlines the argument in this section.

The independent choice at each step of Marchal's algorithm entails that we have independent sub-constructions of Marchal's algorithm with parameter α evolving along each edge of $\mathbf{T}(2)$. This yields three independent copies of $\mathcal{T}_\alpha$, denoted by τ_0, τ_1 and τ_2, modulo rescaling depending on the eventual proportion of mass distributed to each tree. For $\alpha \in (1, 2)$, the internal vertex V_1 will give rise to a countably infinite and independent collection of rescaled $\mathcal{T}_\alpha$. This may be argued similarly to (28). Fix a portion of V_1 with weight $(2 - \alpha)$, call it V_1^{fix}. At time $(n + 1)$, consider the event A_n that V_1^{fix} is selected in Step 2. These events are independent. By the second Borel-Cantelli lemma, $\mathbb{P}\left(\bigcap_{i\geq 1}\bigcup_{n\geq i} A_n\right) = 1$, as $\sum_{n=1}^{\infty}(2 - \alpha)/(n\alpha - 1) = \infty$. Thus, our claim follows. Denote this infinite collection by $(\tau_i : i \geq 3)$, which is independent of τ_0, τ_1 and τ_2. Concatenate our independent collection $(\tau_i : i \geq 0)$ of rescaled $\mathcal{T}_\alpha$ at V_1 to get a copy of $\mathcal{T}_\alpha$. Denote the collection of scaling factors in the limit by $\xi = (\xi_i : i \geq 0)$ and the concatenation operator by g. We obtain an RDE $\mathcal{T}_\alpha \stackrel{d}{=} g\left(\xi, \tau_i : i \geq 0\right)$ in the form (22). To be rigorous, we need to ascertain:

(i) What is the distribution of limiting scaling factors $\xi = (\xi_i : i \geq 0)$?

(ii) Is the collection $(\tau_i : i \geq 0)$ independent, and is it independent of ξ?

(iii) How do we construct the concatenation operation in a measurable way?

4.1 Limiting Weight Partitions in Marchal's Algorithm

For $i \in \{0,1,2\}$ and $n \geq 0$, define $\tau_i^{(n)}$ as the sub-tree of $\mathbf{T}(n+2)$ cut at V_1 containing the edge e_i. For example, we have $\tau_i^{(0)} = e_i$ for each $i \in \{0,1,2\}$. Let $\mathcal{K}_n$ denote the set of edges incident to V_1 in $\mathbf{T}(n+2)$ excluding $\{e_i : i = 0,1,2\}$, and denote $K_n = |\mathcal{K}_n|$. For $\mathcal{K}_n \neq \emptyset$, $\mathcal{K}_n = \{e_j : j = 3, \dots, K_n + 2\}$, ordered according to least leaf labels. Define $\sigma^{(n)}$ as the remaining component of $\mathbf{T}(n+2)$ cut at V_1 excluding $\bigcup_{i=0}^{2} \tau_i^{(n)}$. If $\mathcal{K}_n = \emptyset$, then $\sigma^{(n)} = \emptyset$. Otherwise, $\sigma^{(n)} = \bigcup_{j=3}^{K_n+2} \tau_j^{(n)}$ is a union of sub-trees $\{\tau_j^{(n)} : j = 3, \dots, K_n + 2\}$ growing along their respective edges in $\mathcal{K}_n$. We illustrate this in Figure 2.

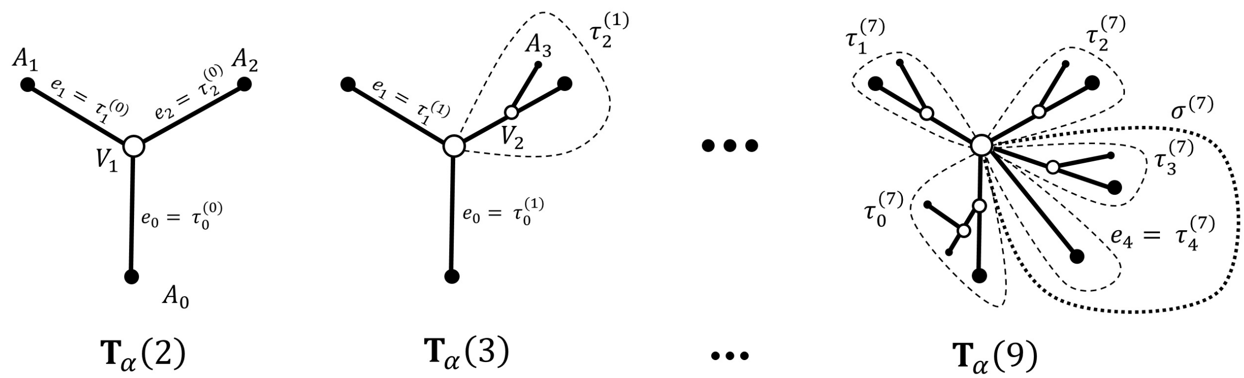

Figure 2: Illustration of Marchal's random growth algorithm and notation employed.

Denote the number of leaves in $\tau_i^{(n)}$ excluding V_1 by $L_i(n)$ for all $i = 0, 1, \dots, K_n + 2$, and define its inverse $L_i^{-1}(n) := \min\{k \geq 0 : L_i(k) = n\}$ as the first time k at which $\tau_i^{(k)}$ has n leaves excluding V_1, with the convention $\min \emptyset = \infty$.

Regard V_1 as a (weightless) root from the perspective of each of $\{\tau_i^{(n)} : i = 1, \dots, K_n + 2\}$ and as a marked leaf of $\tau_0^{(n)}$. For each $i = 1, \dots, K_n + 2$, mark the first leaf created in $\tau_i^{(n)}$ by Marchal's algorithm, that is, the other endpoint of e_i which is not V_1. Recall that A_0 is the root of $\tau_0^{(n)}$.

In the limit, denote by $\tau_i^{(\infty)}$ the limiting set of vertices corresponding to the i-th sub-tree. Denote its associated $\mathbb{R}$-tree by τ_i, obtained by the completion of $\tau_i^{(\infty)}$ in the scaling limit as described in Section 3.3. Likewise, define $\sigma^{(\infty)}$ and σ respectively as the limiting set of vertices and the $\mathbb{R}$-tree associated with $\sigma^{(n)}$.

Recall that $W(\cdot)$ measures the total weight of a given sub-structure. For example, for each $i \in \{0,1,2\}$, $W\left(\tau_i^{(0)}\right) = \alpha - 1$, since e_i is assigned a weight of $\alpha - 1$ by construction and V_1 is regarded as a weightless root (or leaf) from the perspective of $\tau_i^{(0)}$. The following result shows that the weight of a particular sub-tree only depends on the number of the leaves it has, and not its tree-shape.

Proposition 4.1.1. *Regardless of tree-shape, the total weight of the i-th sub-tree is* $W\left(\tau_i^{(n)}\right) = \alpha L_i(n) - 1$ *for* $i = 0, 1, \dots, K_n + 2$ *and* $n \geq 0$.

Proof: Apply the inductive argument in Proposition 3.3.2 to each sub-tree. $\square$

We are now able to give a result on limiting weight partitions in the sub-trees due to Marchal's algorithm. Recall that $\beta = 1 - 1/\alpha$.

Proposition 4.1.2. *We have the following results regarding the limiting distribution of weights.*

(i) For $\alpha = 2$, $K_n = 0$ almost surely. The proportion of weight split has an almost sure limit,

$$\left(\frac{W\left(\tau_0^{(n)}\right)}{2n+3}, \frac{W\left(\tau_1^{(n)}\right)}{2n+3}, \frac{W\left(\tau_2^{(n)}\right)}{2n+3} \right) \xrightarrow{a.s.} (X_0, X_1, X_2) \qquad \text{as } n \to \infty,$$

where $(X_0, X_1, X_2) \sim \mathrm{Dir}\,(1/2, 1/2, 1/2)$.

(ii) For $\alpha \in (1, 2)$, $K_n \to \infty$ as $n \to \infty$ almost surely. The limiting proportion of weight split within the sub-components as $n \to \infty$ is,

$$\left(\frac{W\left(\tau_0^{(n)}\right)}{(n+2)\alpha - 1}, \frac{W\left(\tau_1^{(n)}\right)}{(n+2)\alpha - 1}, \frac{W\left(\tau_2^{(n)}\right)}{(n+2)\alpha - 1}, \frac{W\left(\sigma^{(n)}\right) + W\left(\{V_1\}\right)}{(n+2)\alpha - 1} \right) \xrightarrow{a.s.} (X_0, X_1, X_2, X_3)\,.$$

where $(X_0, X_1, X_2, X_3) \sim \mathrm{Dir}\,(\beta, \beta, \beta, 1 - 2\beta)$. Within σ, denote the eventual proportion of weight distributed to the sub-tree τ_{i+2} by P_i for $i \geq 1$. Then, $(P_i : i \geq 1) \sim \mathrm{GEM}(1-\beta, 1-2\beta)$. In particular, the sub-trees $\{\tau_i : i = 3, 4, \ldots\}$ have their weights partitioned as $\mathrm{PD}(1-\beta, 1-2\beta)$ when ranked in decreasing order by Definition 3.2.7.

Proof: We prove (ii). From Proposition 4.1.1, conditional on an edge or branch point in $\tau_i^{(n)}$ being selected in the next step of Marchal's algorithm, we increase the weight in $\tau_i^{(n)}$ by α. It is easy to check this also holds for $\sigma^{(n)}$ with one weighted copy of V_1 included. Hence,

$$\left(W\left(\tau_0^{(n)}\right), W\left(\tau_1^{(n)}\right), W\left(\tau_2^{(n)}\right), W\left(\sigma^{(n)}\right) + W\left(\{V_1\}\right) \right) \tag{30}$$

evolves precisely as the Pólya urn scheme in Definition 3.2.3 with $K = 4$, initialisation $\vec{\beta} = (\alpha - 1, \alpha - 1, \alpha - 1, 2 - \alpha)$ and step-size α. The almost sure limiting proportion of weight split then holds by Theorem 3.2.4.

Next, we focus on the sub-trees within $\sigma^{(n)}$. The above implies that $W\left(\sigma^{(n)}\right) + W\left(\{V_1\}\right) \to \infty$ as $n \to \infty$ almost surely. So, almost surely, we observe infinitely many leaves being added to $\left(\sigma^{(n)} : n \geq 1\right)$. Consequently, we may condition on the times where a leaf is added to $\left(\sigma^{(n)} : n \geq 1\right)$, say $(q_i : i \geq 1)$, where $1 \leq q_1 < q_2 < \cdots < \infty$ almost surely. Conditional on the preceding event, the first leaf added creates $\tau_3^{(q_1)}$. At each $r \in \{q_n, \ldots, q_{n+1} - 1\}$, we have n leaves (not including V_1) with K_{q_n} sub-trees which unite to form $\sigma^{(r)}$. For $j = 3, \ldots, K_{q_n} + 2$, $\tau_j^{(r)}$ has $L_j(q_n)$ leaves (not including V_1), and so has total weight $\alpha L_j(q_n) - 1$, by Proposition 4.1.1. Thus, as the total weight of V_1 is $2 + K_{q_n} - \alpha$, the total weight of $\sigma^{(r)}$ and $\{V_1\}$ is $\alpha n + (2 - \alpha)$. At the next arrival time q_{n+1}, we add a leaf to $\tau_j^{(q_{n+1}-1)}$ with probability $(\alpha L_j(q_n) - 1) / (\alpha n + 2 - \alpha)$ and we create a new sub-tree with probability $(2 + K_{q_n} - \alpha) / (\alpha n + 2 - \alpha)$. Regarding the leaves (excluding V_1) as customers and each sub-tree as a table, and using Definition 3.2.5, this models a $\mathrm{CRP}(1 - \beta, 1 - 2\beta)$. From Theorem 3.2.6, $K_n \to \infty$ as $n \to \infty$ almost surely. Recall $q_1 < \infty$ almost surely, so we may assume $n \geq q_1$. From Theorem 3.2.6, we can identify the almost sure limiting proportion of leaves split within sub-trees of σ as GEM$(1-\beta, 1-2\beta)$ holding along the increasing subsequence $(q_i : i \geq 1)$. That is,

$$\left(\frac{L_j(q_n)}{n} : j = 3, \ldots, K_{q_n} + 2\right) \xrightarrow{a.s.} (P_i : i \geq 1) \quad \text{ as } n \to \infty,$$

where $(P_i : i \geq 1) \sim \text{GEM}(1-\beta, 1-2\beta)$. Write $L_\sigma(n)$ as the number of leaves in $\sigma^{(n)}$ excluding V_1. Noting that $L_\sigma(n) > 0$ for $n \geq q_1$, we may rephrase the above as

$$\left(\frac{L_j(n)}{L_\sigma(n)} : j = 3, \ldots, K_n + 2\right) \xrightarrow{a.s.} (P_i : i \geq 1) \quad \text{ as } n \to \infty. \tag{31}$$

Using the relation $W\left(\sigma^{(n)}\right) + W\left(\{V_i\}\right) = \alpha L_\sigma(n) + 2 - \alpha$, and the aggregation property of the Dirichlet distribution in Proposition 3.2.2 applied to (30), we get that

$$\frac{L_\sigma(n)}{n} \xrightarrow{a.s.} X_3 \quad \text{ as } n \to \infty, \tag{32}$$

where $X_3 \sim \text{Beta}(1-2\beta, 3\beta)$. By the algebra of almost sure convergence,

$$\left(\frac{L_j(n)}{n} : j = 3, \ldots, K_n + 2\right) \xrightarrow{a.s.} (X_3 P_i : i \geq 1) \quad \text{ as } n \to \infty,$$

Therefore, for all $j = 3, \ldots, K_n + 2$ and $n \geq q_1$, as we have $W\left(\tau_j^{(n)}\right) = \alpha L_j(n) - 1$, the above implies that, jointly in j,

$$\frac{W\left(\tau_j^{(n)}\right)}{(n+2)\alpha - 1} = \frac{\frac{L_j(n)}{n} - \frac{1}{\alpha n}}{\frac{n+2}{n} - \frac{1}{\alpha n}} \xrightarrow{a.s.} X_3 P_{j-2} \quad \text{ as } n \to \infty,$$

where $X_3 \sim \text{Beta}(1-2\beta, 3\beta)$ and $(P_i : i \geq 1) \sim \text{GEM}(1-\beta, 1-2\beta)$. Thus, we have obtained the almost sure limiting weight partition for the sub-trees $\{\tau_j : j \geq 0\}$. The proof of (i) follows noting that $\sigma^{(n)} = \emptyset$ for all $n \geq 1$ almost surely. $\square$

4.2 Independences in Marchal's Algorithm

After ascertaining the limiting weight partitions in the sub-trees of Marchal's algorithm, we now proceed to verify the requisite independences. We first require an original technical lemma.

Lemma 4.2.1. *Let T be an almost surely finite stopping time with respect to a filtration $(\mathcal{F}_n)_{n\geq 1}$. Suppose that X is a non-negative and bounded random variable satisfying, for each $n \geq 1$,*

$$\mathbb{E}\left[X \mid \mathcal{F}_T\right] = \mathbb{E}\left[X \mid \mathcal{F}_m\right] \quad \textit{almost surely, for all } m \geq n \textit{ on } \{T = n\},$$

then $\mathbb{E}\left[X \mid \mathcal{F}_T\right] = \mathbb{E}\left[X \mid \mathcal{F}_\infty\right]$ almost surely.

Proof: Clearly, $\mathbb{E}\left[X \mid \mathcal{F}_T\right]$ is $\mathcal{F}_\infty$-measurable. Define $\mathcal{G} := \bigcup_{n\geq 1} \mathcal{F}_n$ and consider the collection

$$\mathcal{H} := \{A \in \mathcal{F}_\infty : \mathbb{E}\left[X \mathbf{1}_A\right] = \mathbb{E}\left[\mathbb{E}\left[X \mid \mathcal{F}_T\right] \mathbf{1}_A\right]\}.$$

We claim that $\mathcal{G} \subseteq \mathcal{H}$. Suppose that $A \in \mathcal{G}$, then $A \in \mathcal{F}_m$ for some $m \geq 1$. For $1 \leq j \leq m$, by the given hypotheses,

$$\begin{aligned}\mathbb{E}\left[\mathbb{E}\left[X \mid \mathcal{F}_T\right] \mathbf{1}_{A \cap \{T=j\}}\right] &= \mathbb{E}\left[\mathbb{E}\left[X \mid \mathcal{F}_m\right] \mathbf{1}_{A \cap \{T=j\}}\right] \\ &= \mathbb{E}\left[X \mathbf{1}_{A \cap \{T=j\}}\right],\end{aligned}$$

where $A \cap \{T = j\} \in \mathcal{F}_m$ since T is a stopping time. Next, we claim that $A \cap \{T > m\} \in \mathcal{F}_T$. The claim holds, since T is a stopping time, and so

$$A \cap \{T > m\} \cap \{T = n\} = \begin{cases} \emptyset \in \mathcal{F}_n & \text{if } 1 \leq n \leq m, \\ A \cap \{T = n\} \in \mathcal{F}_n & \text{if } n > m. \end{cases}$$

Consolidating our observations above,

$$\begin{aligned}\mathbb{E}\left[\mathbb{E}\left[X \mid \mathcal{F}_T\right] \mathbf{1}_A\right] &= \sum_{j=1}^{m} \mathbb{E}\left[\mathbb{E}\left[X \mid \mathcal{F}_T\right] \mathbf{1}_{A \cap \{T=j\}}\right] + \mathbb{E}\left[\mathbb{E}\left[X \mid \mathcal{F}_T\right] \mathbf{1}_{A \cap \{T>m\}}\right] \\ &= \sum_{j=1}^{m} \mathbb{E}\left[X \mathbf{1}_{A \cap \{T=j\}}\right] + \mathbb{E}\left[X \mathbf{1}_{A \cap \{T>m\}}\right] \\ &= \mathbb{E}\left[X \mathbf{1}_A\right].\end{aligned}$$

We conclude that $\mathcal{G} \subseteq \mathcal{H}$. As X is non-negative and bounded, the Monotone Convergence Theorem applies, and so $\mathcal{H}$ is a monotone class. Lastly, $\mathcal{G}$ is a π-system generating $\mathcal{F}_\infty$, so $\mathcal{H} = \mathcal{F}_\infty$ by the Monotone Class Theorem. Hence, $\mathbb{E}\left[X \mid \mathcal{F}_T\right]$ is a version of $\mathbb{E}\left[X \mid \mathcal{F}_\infty\right]$. $\square$

The following proof is inspired by [21, Lemma 8] in considering transition times at which a leaf is added into a sub-tree. However, we extend their result by considering transitions jointly over multiple sub-trees. Fix $\alpha \in (1, 2)$ so that we are in the infinitary case. It is easy to adapt our results to the BCRT.

Theorem 4.2.2. *For $n \geq 1, i \geq 0$, $\tau_i^{\left(L_i^{-1}(n)\right)} \overset{d}{=} \mathbf{T}(n)$. That is, at transition times in which a leaf is added into the i-th sub-tree, it evolves as Marchal's algorithm with parameter $\alpha \in (1, 2)$ with initial edge e_i. The sigma-field generated by*

$$\left(\tau_i^{\left(L_i^{-1}(n)\right)} : n \geq 1\right)_{i \geq 0}$$

is independent of the sigma-field generated by $(L_i(n) : n \geq 1, i \geq 0)$. Consequently, $(\tau_i : i \geq 0)$ are independent, and is independent of $(L_i(n) : n \geq 1, i \geq 0)$.

Proof: From Proposition 4.1.2, we have $K_n \to \infty$ as $n \to \infty$ almost surely. In particular, almost surely, for all $i \geq 0$ and $n \geq 1$, $L_i^{-1}(n) < \infty$. We assume this holds henceforth. To prove our desired result, it suffices to show the independence of the sigma-field generated by $(L_i(n) : n \geq 1, i \geq 0)$ and the sigma-field generated by $\left(\tau_i^{\left(L_i^{-1}(n)\right)} : n \geq 1\right)_{0 \leq i \leq m+2}$, where $m \geq 0$ is arbitrary but fixed.

To prove the first claim, consider a given time $n \geq L_{m+2}^{-1}(1)$. Conditional on a leaf being added to the i-th sub-tree for $0 \leq i \leq m+2$, we have the dynamic of Marchal's algorithm with parameter

α by the weight-leaf relation in Proposition 4.1.1. Likewise, the transition in the other components, not including the i-th sub-trees for $0 \le i \le m+2$, follows the correct conditional distributions of Marchal's algorithm. This proves the distributional identity $\tau_i^{\left(L_i^{-1}(n)\right)} \overset{d}{=} \mathbf{T}(n)$ at transition times in the i-th sub-tree for $0 \le i \le m+2$.

Let $M > 1$ be arbitrary, but fixed, and denote the natural filtration of $(L_i(n) : i \ge 0)_{n \ge 1}$ by $(\mathcal{F}_n)_{n\ge1}$. Note that for any fixed $n \ge 1$, $(L_i(n) : i \ge 0)$ is almost surely a vector with finitely many non-trivial entries. Define $T := \max_{i=0,\dots,m+2} L_i^{-1}(M)$, which is a stopping time relative to $(\mathcal{F}_n)_{n\ge1}$. By assumption, T is almost surely finite. Conditional on $\mathcal{F}_T$ (which is the same as conditioning on relative weights on sub-trees till time T), we have factorisation of tree-shape probabilities in Theorem 3.3.3 into tree-shape probabilities for the respective sub-trees cut at V_1. In particular, given $\mathcal{F}_T$, the tree-shapes $\left(\tau_i^{\left(L_i^{-1}(n)\right)} : 1 \le n \le M\right)_{0\le i\le m+2}$ are independent. Furthermore, on the event $\{T = t\}$, conditioning on the sigma-field generated at a later time $k \ge t$ does not affect the tree-shapes under consideration. Hence, the hypotheses in Lemma 4.2.1 are fulfilled. Let $\mathbf{t}_i^{(n)}$ be some given leaf-labelled tree with n leaves and a root. Then,

$$
\begin{aligned}
&\mathbb{P}\left(\tau_i^{\left(L_i^{-1}(n)\right)} = \mathbf{t}_i^{(n)} : 1 \le n \le M, 0 \le i \le m+2 \;\middle|\; \mathcal{F}_\infty\right) \\
&= \mathbb{P}\left(\tau_i^{\left(L_i^{-1}(n)\right)} = \mathbf{t}_i^{(n)} : 1 \le n \le M, 0 \le i \le m+2 \;\middle|\; \mathcal{F}_T\right) \\
&= \mathbb{P}\left(\tau_i^{\left(L_i^{-1}(M)\right)} = \mathbf{t}_i^{(M)} : 0 \le i \le m+2 \;\middle|\; \mathcal{F}_T\right) && (33) \\
&= \prod_{i=0}^{m+2} \mathbb{P}\left(\tau_i^{\left(L_i^{-1}(M)\right)} = \mathbf{t}_i^{(M)} \;\middle|\; \mathcal{F}_T\right) && (34) \\
&= \prod_{i=0}^{m+2} \mathbb{P}\left(\mathbf{T}(M) = \mathbf{t}_i^{(M)} \;\middle|\; \mathcal{F}_T\right) && (35) \\
&= \prod_{i=0}^{m+2} \mathbb{P}\left(\mathbf{T}(M) = \mathbf{t}_i^{(M)}\right), && (36)
\end{aligned}
$$

where (33) holds since $\tau_i^{\left(L_i^{-1}(M)\right)}$ determines $\tau_i^{\left(L_i^{-1}(n)\right)}$ for all $1 \le n \le M$, and where (34) holds by Theorem 3.3.3, and (36) follows since there is no dependence on $\mathcal{F}_\infty$ in evaluating (35) and we are conditioning over an almost surely finite number of discrete random variables. Furthermore, since the final expression does not depend on $\mathcal{F}_\infty$, then the sigma-field of $\left(\tau_i^{\left(L_i^{-1}(n)\right)} : 1 \le n \le M\right)_{0\le i\le m+2}$ is independent of $\mathcal{F}_\infty$. Letting $M \to \infty$, and recalling $m \ge 0$ is arbitrary, $\left(\tau_i^{\left(L_i^{-1}(n)\right)} : n \ge 1\right)_{i\ge0}$ and $(L_i(n) : n \ge 1, i \ge 0)$ are independent.

As $(\tau_i : i \ge 0)$ is measurably constructed from $\left(\tau_i^{\left(L_i^{-1}(n)\right)} : n \ge 1\right)_{i\ge0}$, it is independent of $(L_i(n) : n \ge 1, i \ge 0)$.

Dropping the conditioning in (35), we get that $\left(\tau_i^{\left(L_i^{-1}(n)\right)} : 1 \le n \le M\right)_{0\le i\le m+2}$ are independent. Thus, in the limit as $M \to \infty$, $\left(\tau_i^{\left(L_i^{-1}(n)\right)} : n \geq 1\right)_{0 \leq i \leq m+2}$ are independent. As τ_i is measurably constructed from $\left(\tau_i^{\left(L_i^{-1}(n)\right)} : n \geq 1\right)$ for each $0 \leq i \leq m+2$, $(\tau_i : 0 \leq i \leq m+2)$ are independent. Let $m \to \infty$ to conclude that $(\tau_i : i \geq 0)$ are independent. □

Theorem 4.2.3. *For $\alpha \in (1, 2)$, the random variables (X_0, X_1, X_2, X_3) and $(P_i : i \geq 1)$ as defined in Proposition 4.1.2 are independent. In particular, this fully specifies their joint distribution.*

Proof: Recall $L_\sigma(n)$ denotes the number of leaves (excluding V_1) in $\sigma^{(n)}$, and define $L_\sigma^{-1}(n) := \min\{k \geq 0 : L_\sigma(k) = n\}$. We claim that the sigma-field generated by $\left(L_j\left(L_\sigma^{-1}(n)\right) : n \geq 1\right)_{j \geq 3}$ is independent of the sigma-field generated by $\left(L_0^{-1}(n), L_1^{-1}(n), L_2^{-1}(n), L_\sigma^{-1}(n) : n \geq 1\right)$. Denote the natural filtration of $\left(L_0^{-1}(n), L_1^{-1}(n), L_2^{-1}(n), L_\sigma^{-1}(n)\right)_{n \geq 1}$ by $(\mathcal{F}_n)_{n \geq 1}$. It suffices to prove that the sigma-field generated by $\left(L_j\left(L_\sigma^{-1}(n)\right) : n \geq 1\right)_{3 \leq j \leq m}$ is independent of $\mathcal{F}_\infty$, where $m \geq 3$ is arbitrary but fixed. By Proposition 4.1.2, almost surely, $L_\sigma^{-1}(n) < \infty$ for all $n \geq 1$. Let $M \geq 1$ be arbitrary but fixed, then $T := L_\sigma^{-1}(M)$ is an almost surely finite stopping time relative to $(\mathcal{F}_n)_{n \geq 1}$. Consider the random variable $\left(L_j\left(L_\sigma^{-1}(n)\right) : 1 \leq n \leq M\right)_{3 \leq j \leq m}$. On the event $\{T = t\}$, conditioning on the sigma-field generated at a later time $k \geq t$ does not affect conditional expectations. Hence, by Lemma 4.2.1, for all non-negative integers $l_j(n)$,

$$
\begin{aligned}
&\mathbb{P}\left(L_j\left(L_\sigma^{-1}(n)\right) = l_j(n) : 1 \leq n \leq M, 3 \leq j \leq m \;\middle|\; \mathcal{F}_\infty\right) \\
&= \mathbb{P}\left(L_j\left(L_\sigma^{-1}(n)\right) = l_j(n) : 1 \leq n \leq M, 3 \leq j \leq m \;\middle|\; \mathcal{F}_T\right) \\
&= \mathbb{P}\left(L_j\left(L_\sigma^{-1}(n)\right) = l_j(n) : 1 \leq n \leq M, 3 \leq j \leq m\right).
\end{aligned}
$$

The last equality follows, since conditional on a leaf being added to $\sigma^{(n)}$ at time $(n+1)$, the process of adding leaves to each sub-tree within $\sigma^{(n)}$ is modelled by a CRP $(1-\beta, 1-2\beta)$ by Proposition 4.1.2, and does not depend on the times at which the leaf is added. Since we are conditioning over an almost surely finite number of discrete random variables, we may drop conditioning on $\mathcal{F}_T$. This implies that the sigma-field generated by $\left(L_j\left(L_\sigma^{-1}(n)\right) : 1 \leq n \leq M\right)_{3 \leq j \leq m}$ is independent of $\mathcal{F}_\infty$ for all $M \geq 1$. Let $M \to \infty$, so the sigma-field generated by $\left(L_j\left(L_\sigma^{-1}(n)\right) : n \geq 1\right)_{3 \leq j \leq m}$ is independent of $\mathcal{F}_\infty$, as desired. Since we may rewrite equation (32) to get

$$
\frac{n}{L_\sigma^{-1}(n)} \xrightarrow{a.s.} X_3 \quad \text{as } n \to \infty,
$$

then X_3 is $\mathcal{F}_\infty$-measurable. Likewise, X_0, X_1 and X_2 are $\mathcal{F}_\infty$-measurable. From equation (31),

$$
\frac{L_{i+2}\left(L_\sigma^{-1}(n)\right)}{n} \xrightarrow{a.s.} P_i \quad \text{as } n \to \infty,
$$

for all $i \geq 1$. So, $(P_i : i \geq 1)$ is measurable with respect to the sigma-field of $\left(L_j\left(L_\sigma^{-1}(n)\right) : n \geq 1\right)_{j \geq 3}$. The desired result follows. □

Combining the results above, we finally obtain the main result regarding the self-similarity of Marchal's algorithm. This proves the self-similarity property of α-stable trees when decomposed at the first branch point.

Theorem 4.2.4. *In either of the following cases, the limiting trees are independent, and are independent of their scaling factors. For each sub-tree $\tau_i^{(n)}$, $d_i^{(n)}$ denotes the graph distance and $\mu_i^{(n)}$ is the empirical mass measure on its leaves.*

(i) If $\alpha = 2$, for each $i \in \{0,1,2\}$, we have the convergence

$$\left(\tau_i^{(n)}, \frac{d_i^{(n)}}{2n^{1/2}}, \mu_i^{(n)}\right) \xrightarrow{a.s.} \left(\mathcal{T}_2, \xi_i^{1/2} d_2, \xi_i \mu_2\right) \qquad \text{as } n \to \infty,$$

in the Gromov-Hausdorff-Prokhorov topology, where $(\xi_0, \xi_1, \xi_2) \sim \mathrm{Dir}\,(1/2, 1/2, 1/2)$.

(ii) If $\alpha \in (1,2)$, for each $i \geq 0$, we have the convergence

$$\left(\tau_i^{(n)}, \frac{d_i^{(n)}}{\alpha n^{\beta}}, \mu_i^{(n)}\right) \xrightarrow{a.s.} \left(\mathcal{T}_\alpha, \xi_i^{\beta} d_\alpha, \xi_i \mu_\alpha\right) \qquad \text{as } n \to \infty,$$

in the Gromov-Hausdorff-Prokhorov topology, where $\xi_i = X_i$ for $i \in \{0,1,2\}$ and $\xi_{j+2} = X_3 P_j$ for $j \geq 1$, with $(X_0, X_1, X_2, X_3) \sim \mathrm{Dir}\,(\beta, \beta, \beta, 1-2\beta)$ and $(P_i, i \geq 1)^{\downarrow} \sim \mathrm{PD}(1-\beta, 1-2\beta)$ independent.

Proof: The almost sure convergence in the rescaled sub-trees arises by applying Theorem 3.3.5 and Proposition 4.1.2. The independence between the limiting sub-trees comes immediately from Theorem 4.2.2. The arguments in Propositions 4.1.1 and 4.1.2 show that the limiting proportion of weights is measurably constructed from $(L_i(n) : n \geq 1, i \geq 0)$. Hence, by Theorem 4.2.2, the limiting sub-trees are independent of their scaling factors. The joint distribution of $(\xi_i : i \geq 0)$ when $\alpha \in (1,2)$ is fully specified by Theorem 4.2.3. □

The results of Theorem 4.2.4 agree with similar decompositions of the BCRT at a branch point in Aldous [8, Theorem 2], Albenque and Goldschmidt [4, Section 1.4], and Croydon and Hambly [19, Lemma 6], where the branch point is uniquely determined by a uniformly chosen point according to the mass measure within each of the three sub-trees. We point out that Albenque and Goldschmidt deal with an unrooted BCRT, while Croydon and Hambly's construction uses a doubly-marked rooted BCRT. Our construction thus far does not require a notion of a mass measure, but rather a marked point in each sub-tree.

4.3 Measurability of the Concatenation Operation

After verifing that the sub-trees $\{\tau_i : i \geq 0\}$ are rescaled versions of $\mathcal{T}_\alpha$ in the limit with the required independences, the next step is to show that the concatenation operation induced by Marchal's algorithm is well-defined and measurable. For the subsequent discussion, we adapt the general setup and terminology from [61].

Denote $\Xi := \{(x_0, x_1, x_2, x_3 p_j : j \geq 1) : \sum_{i=0}^{3} x_i = 1, x_i \geq 0, \sum_{i=1}^{\infty} p_i \leq 1, p_1 \geq p_2 \geq \cdots \geq 0\}$. For notational convenience, we write

$$\xi_i = \begin{cases} x_i & \text{if } i \in \{0,1,2\}, \\ x_3 p_{i-2} & \text{otherwise.} \end{cases}$$

Recall that $\mathbb{T}_m$ is the set of GH^m-equivalence classes of marked compact rooted $\mathbb{R}$-trees. As with the setup in (20), define $\Xi^* := \Xi \times \mathbb{T}_m^\infty$. Note that in our case, $N = \inf\{i \geq 0 : \xi_i = 0\}$ with the convention $\inf \emptyset = \infty$, so that N is a function of $\xi = (\xi_i : i \geq 0) \in \Xi$. Furthermore, in the case of α-stable trees, recall that $N = 2$ or $N = \infty$ almost surely. Hence, we drop dependence on N in our notation. Fix $\beta \in \left(0, \frac{1}{2}\right]$, and equip Ξ^* with the following metric

$$d_\beta\left(\left(\xi, \tau_i : i \geq 0\right), \left(\xi', \tau_i' : i \geq 0\right)\right) := \sup_{i \geq 0} \left(|\xi_i^\beta - \xi_i'^\beta| \vee d_{GH}^m\left(\tau_i, \tau_i'\right) \vee d_{GH}^m\left(\xi_i^\beta \tau_i, \xi_i'^\beta \tau_i'\right)\right), \tag{37}$$

where $\xi = (\xi_i : i \geq 0) \in \Xi, \xi' = (\xi_i' : i \geq 0) \in \Xi$, and $(\tau_i, d_i, \rho_i, x_i), (\tau_i', d', \rho_i', x_i')$ are representatives of GH^m-equivalence classes in $\mathbb{T}_m$. However, as d_{GH}^m only depends on GH^m-equivalence classes, our metric d_β also only depends on GH^m-equivalence classes. Hence, we may define d_β on Ξ^* and denote by τ any representative of the GH^m-equivalence class of (τ, d, ρ, x).

Proposition 4.3.1. *(Ξ^*, d_β) is a Polish metric space.*

Proof: We adapt [61, Proposition 3.1] to work with the marked Gromov-Hausdorff metric. Recall from Theorem 2.2.1 that $(\mathbb{T}_m, d_{GH}^m)$ is Polish. Fix $\epsilon > 0$ in the following. By standard results about the product topology, if (A_i, d_i) are Polish metric spaces for $i \geq 0$, then $(A_0 \times A_1 \times A_2 \times \ldots, d_\infty)$ is a Polish metric space, where d_∞ is the canonical supremum metric over the metrics d_i. Thus, it suffices to prove that $\left([0,1] \times \mathbb{T}_m, \tilde{d}_\beta\right)$ is a Polish metric space, where

$$\tilde{d}_\beta\left(\left(\xi, \tau\right), \left(\xi', \tau'\right)\right) := |\xi^\beta - \xi'^\beta| \vee d_{GH}^m\left(\tau, \tau'\right) \vee d_{GH}^m\left(\xi^\beta \tau, \xi'^\beta \tau'\right),$$

given $\xi, \xi' \in [0,1]$ and $\tau, \tau' \in \mathbb{T}_m$. This defines a metric as $([0,1], |\cdot|)$ and $(\mathbb{T}_m, d_{GH}^m)$ are metric spaces, with the maximum operator preserving the triangle inequality. To verify completeness, take a Cauchy sequence $\left((\xi_n, \tau_n)\right)_{n \geq 1}$ in $[0,1] \times \mathbb{T}_m$. Then, there exists $N_0(\epsilon) \geq 1$ such that for all $m, n \geq N_0(\epsilon)$, $\tilde{d}_\beta\left(\left(\xi_m, \tau_m\right), \left(\xi_n, \tau_n\right)\right) < \epsilon$. This implies that $(\xi_n)_{n \geq 1}$ and $(\tau_n)_{n \geq 1}$ are Cauchy sequences in $[0,1]$ and $\mathbb{T}_m$ respectively. By completeness of $([0,1], |\cdot|)$ and $(\mathbb{T}_m, d_{GH}^m)$, $\left((\xi_n, \tau_n)\right)_{n \geq 1}$ has a limit in $[0,1] \times \mathbb{T}_m$, so $\left([0,1] \times \mathbb{T}_m, \tilde{d}_\beta\right)$ is complete. By separability of $([0,1], |\cdot|)$ and $(\mathbb{T}_m, d_{GH}^m)$, there exist countable dense subsets $\mathbb{Q} \cap [0,1] \subseteq [0,1]$ and $\mathbb{T}_m' \subseteq \mathbb{T}_m$ such that given any $\xi \in [0,1]$ and $\tau \in \mathbb{T}_m$ with $\mathrm{ht}(\tau) > 0$, there exist $\xi' \in \mathbb{Q} \cap (0,1]$ and $\tau' \in \mathbb{T}_m'$ satisfying

$$|\xi^\beta - \xi'^\beta| < \min\{\epsilon, \epsilon / \left(2\mathrm{ht}(\tau)\right)\} \quad \text{and} \quad d_{GH}^m\left(\tau, \tau'\right) < \min\{\epsilon, \epsilon / 2\xi'^\beta\}.$$

Define $D := (\mathbb{Q} \cap [0,1]) \times \mathbb{T}_m'$, which is countable as a Cartesian product of countable sets. By considering any embedding in a common metric space and rescaling the metric, we deduce that

$$d_{GH}^m\left(\xi^\beta \tau, \xi'^\beta \tau\right) \leq |\xi^\beta - \xi'^\beta| \ \mathrm{ht}(\tau).$$

Hence, by the triangle inequality applied to d_{GH}^m,

$$\begin{aligned} d_{GH}^m\left(\xi^\beta \tau, \xi'^\beta \tau'\right) &\leq d_{GH}^m\left(\xi^\beta \tau, \xi'^\beta \tau\right) + d_{GH}^m\left(\xi'^\beta \tau, \xi'^\beta \tau'\right) \\ &\leq |\xi^\beta - \xi'^\beta| \ \mathrm{ht}(\tau) + \xi'^\beta d_{GH}^m\left(\tau, \tau'\right) \\ &< \epsilon. \end{aligned}$$

Therefore, D is a countable dense set in $[0,1]\times\mathbb{T}_m$, and so $\left([0,1]\times\mathbb{T}_m,\tilde{d}_\beta\right)$ is separable. $\square$

We now formally define our concatenation operator. Given $\xi\in\Xi$ and let (τ_i,d_i,ρ_i,x_i) be representatives of GH^m-equivalence classes in $\mathbb{T}_m$ for $i\geq 0$. Define the *concatenated tree* (τ',d',ρ',x') by the following:

(i) Let $\tilde{\tau}':=\coprod_{i\geq 0}\xi_i^\beta\tau_i$ be the the disjoint union of rescaled trees. Let $\sim_c$ be the equivalence relation on $\tilde{\tau}'$ in which $\rho_i\sim_c x_0$ for all $i\geq 1$. Define $\tau':=\tilde{\tau}'/\sim_c$. Write ψ_c for the canonical projection from $\tilde{\tau}'$ onto τ'.

(ii) Define a metric $\tilde{d}'$ on $\tilde{\tau}'$ such that

$$\tilde{d}'(v,w)=\begin{cases}\xi_i^\beta d_i(v,w) & \text{if } v,w\in\tau_i, i\geq 0,\\ \xi_0^\beta d_0(v,x_0)+\xi_j^\beta d_j(\rho_j,w) & \text{if } v\in\tau_0 \text{ and } w\in\tau_j, j\neq 0,\\ \xi_i^\beta d_i(v,\rho_i)+\xi_0^\beta d_0(x_0,w) & \text{if } v\in\tau_i \text{ and } w\in\tau_0, i\neq 0,\\ \xi_i^\beta d_i(v,\rho_i)+\xi_j^\beta d_j(\rho_j,w) & \text{if } v\in\tau_i \text{ and } w\in\tau_j, i,j\neq 0.\end{cases}\tag{38}$$

Let d' be the metric induced on τ' under ψ_c by $\tilde{d}'$.

(iii) Retain $x'=\psi_c(x_1)$ as our marked point in τ' and set $\rho'=\psi_c(\rho_0)$ as the root of τ'.

(iv) We illustrate this construction in Figure 3.

By virtue of this construction, the GH^m-equivalence class of (τ',d',ρ',x') only depends on the GH^m-equivalence classes of (τ_i,d_i,ρ_i,x_i) for $i\geq 0$. Thus, it makes sense to define $C_\beta\subseteq\Xi^*$ as the set of elements $\kappa=(\xi_i,\tau_i:i\geq 0)\in\Xi^*$ such that the concatenated tree (τ',d',ρ',x') formed by any equivalence class representatives of $((\tau_i,d_i,\rho_i,x_i):i\geq 0)$ is compact. Equip $\mathbb{T}_m$ and Ξ^* with their respective Borel sigma-algebras, $\mathcal{B}(\mathbb{T}_m)$ and $\mathcal{B}(\Xi^*)$. The *concatenation operator* $g_\beta\colon\Xi^*\to\mathbb{T}_m$ is,

$$g_\beta(\kappa)=\begin{cases}(\tau',d',\rho',x') & \text{if } \kappa\in C_\beta,\\ (\{x'\},0,x',x') & \text{otherwise,}\end{cases}\tag{39}$$

where $(\{x'\},0,x',x')$ denotes the equivalence class of a trivial one-point rooted tree marked at its root. We need the following technical lemma to prove the measurability of g_β.

Lemma 4.3.2. *Let $(A,\mathcal{B}(A)),(B,\mathcal{B}(B))$ be separable metric spaces equipped with their Borel sigma-algebras, and let $A_0\in\mathcal{B}(A)$ and $b_0\in B$. If $f\colon A\to B$ is such that $f(A\setminus A_0)=\{b_0\}$ and $f_0:=f\restriction_{A_0}$ is continuous, then f is $\mathcal{B}(A)$-measurable.*

Proof: Our proof parallels [61, Lemma 3.3], though furnished with more details. First, assume $A_0=A$, so that f is continuous on A. Consider the collection

$$\mathcal{G}:=\{B_0\in\mathcal{B}(B):f^{-1}(B_0)\in\mathcal{B}(A)\}.$$

The open sets of B form a π-system that is contained in $\mathcal{G}$ by continuity of f. As set operations commute with pre-images of functions, we conclude that $\mathcal{G}=\mathcal{B}(B)$ by the Monotone Class Theorem. So, f is $\mathcal{B}(A)$-measurable. Suppose now that $A_0\neq A$. The subset topology on A_0 is separable and

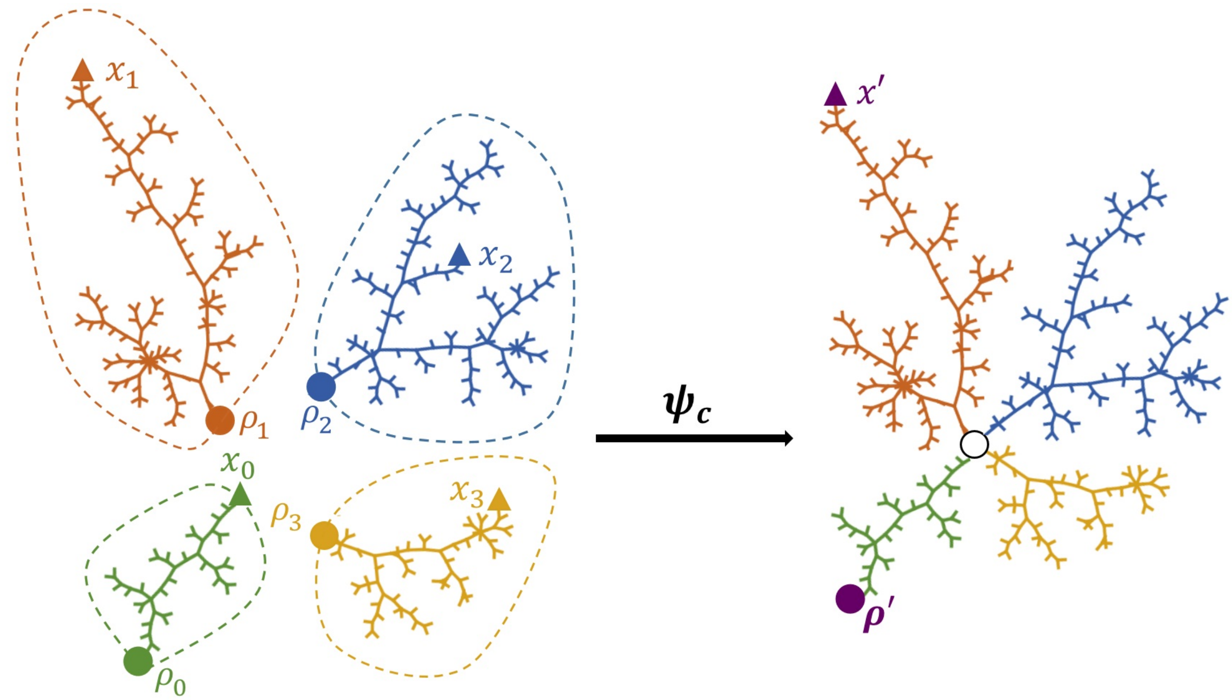

Figure 3: Construction of concatenated tree from 4 marked trees, rescaling is not shown. Adapted from images by Igor Kortchemski.

induced by the restriction the metric of A to A_0. In particular, f_0 is $\mathcal{B}(A)$-measurable. Given any $B_0 \in \mathcal{B}(B)$, by properties of the pre-image, we have

$$f^{-1}(B_0) = \begin{cases} \left(f^{-1}(b_0) \setminus f_0^{-1}(b_0)\right) \cup f_0^{-1}(B_0) & \text{if } b_0 \in B_0, \\ f_0^{-1}(B_0) & \text{otherwise.} \end{cases}$$

As f_0 is $\mathcal{B}(A)$-measurable, $f_0^{-1}(B_0) \in \mathcal{B}(A)$ and $f^{-1}(b_0) \setminus f_0^{-1}(b_0) = A \setminus A_0 \in \mathcal{B}(A)$. So, given any $B_0 \in \mathcal{B}(B)$, $f^{-1}(B_0) \in \mathcal{B}(A)$ and we are done. $\square$

Proposition 4.3.3. *The map $g_\beta \colon \Xi^* \to \mathbb{T}_m$ is $\mathcal{B}(\Xi^*)$-measurable.*

Proof: We adapt [61, Proposition 3.2] to handle the marked Gromov-Hausdorff metric using an argument via correspondences. Apply Lemma 4.3.2 to $A = \Xi^*$, $B = \mathbb{T}_m$, $f = g_\beta$, $A_0 = C_\beta$, and $b_0 = (\{x'\}, 0, x', x')$. The topological requirements are met using Theorem 2.2.1 and Proposition 4.3.1. It suffices to prove that $g_\beta \restriction_{C_\beta}$ is continuous and $C_\beta \in \mathcal{B}(\Xi^*)$.

Suppose $(\kappa^{(n)} : n \geq 1)$ is a sequence in (C_β, d_β) converging to some limit $\kappa \in C_\beta$. Denote $\kappa^{(n)} = \left(\xi_i^{(n)}, \tau_i^{(n)} : i \geq 0\right)$ for $n \geq 1$ and $\kappa = (\xi_i, \tau_i : i \geq 0)$. Write $\mathcal{T}^{(n)} := g_\beta(\kappa^{(n)})$ and $\mathcal{T} := g_\beta(\kappa)$.

Recall that $\mathcal{C}^m\left(\mathcal{T}^{(n)}, \mathcal{T}\right)$ denotes the set of a correspondences between $\mathcal{T}^{(n)}$ and $\mathcal{T}$ with $\left(\rho^{(n)}, \rho\right)$ and $\left(x^{(n)}, x\right)$ in correspondence. Denote $\mathcal{C}_i^{(n)}$ as the set of correspondences between $\tau_i^{(n)}$ and τ_i with $\left(\rho_i^{(n)}, \rho_i\right)$ and $\left(x_i^{(n)}, x_i\right)$ in correspondence. Then, write $\mathcal{C}^{(n)}$ to be set of correspondences formed by combining the correspondences $\mathcal{C}_i^{(n)}$ for $i \geq 0$ under ψ_c, so that $\mathcal{C}^{(n)} \subseteq \mathcal{C}^m\left(\mathcal{T}^{(n)}, \mathcal{T}\right)$. For notational simplicity, write

$$\hat{\rho}_i^{(n)} = \begin{cases} x_0^{(n)} & \text{if } i = 0, \\ \rho_i^{(n)} & \text{otherwise,} \end{cases}$$

with a similar definition for $\hat{\rho}_i$. Using the formula in (6), we have

$$\begin{aligned}
&d_{GH}^m\left(\left(\mathcal{T}^{(n)}, (d')^{(n)}, \rho^{(n)}, x^{(n)}\right), \left(\mathcal{T}, d', \rho, x\right)\right) \\
&= \frac{1}{2} \inf_{\mathcal{R} \in \mathcal{C}^m\left(\mathcal{T}^{(n)}, \mathcal{T}\right)} \sup_{\left(u^{(n)}, u\right), \left(v^{(n)}, v\right) \in \mathcal{R}} \left| (d')^{(n)}\left(u^{(n)}, v^{(n)}\right) - d'(u,v) \right| \\
&\leq \frac{1}{2} \inf_{\mathcal{R} \in \mathcal{C}^{(n)}} \sup_{\left(u^{(n)}, u\right), \left(v^{(n)}, v\right) \in \mathcal{R}} \left| (d')^{(n)}\left(u^{(n)}, v^{(n)}\right) - d'(u,v) \right| \\
&\leq \frac{1}{2} \sup_{i,j \geq 0} \inf_{\substack{\mathcal{R}_i \in \mathcal{C}_i^{(n)} \\ \mathcal{R}_j \in \mathcal{C}_j^{(n)}}} \sup_{\substack{\left(u^{(n)}, u\right) \in \mathcal{R}_i \\ \left(v^{(n)}, v\right) \in \mathcal{R}_j}} \left| (\tilde{d}')^{(n)}\left(u^{(n)}, v^{(n)}\right) - \tilde{d}'(u,v) \right| && (40) \\
&\leq \frac{1}{2} \sup_{i,j \geq 0} \inf_{\substack{\mathcal{R}_i \in \mathcal{C}_i^{(n)} \\ \mathcal{R}_j \in \mathcal{C}_j^{(n)}}} \sup_{\substack{\left(u^{(n)}, u\right) \in \mathcal{R}_i \\ \left(v^{(n)}, v\right) \in \mathcal{R}_j}} \Bigg| \left(\xi_i^{(n)}\right)^\beta d_i^{(n)}\left(u^{(n)}, \hat{\rho}_i^{(n)}\right) + \left(\xi_j^{(n)}\right)^\beta d_j^{(n)}\left(\hat{\rho}_j^{(n)}, v^{(n)}\right) \\
&\qquad\qquad - \xi_i^\beta d_i(u, \hat{\rho}_i) - \xi_j^\beta d_j(\hat{\rho}_j, v) \Bigg| && (41) \\
&\leq \sup_{i \geq 0} \inf_{\mathcal{R}_i \in \mathcal{C}_i^{(n)}} \sup_{\left(u^{(n)}, u\right) \in \mathcal{R}_i} \left| \left(\xi_i^{(n)}\right)^\beta d_i^{(n)}\left(u^{(n)}, \hat{\rho}_i^{(n)}\right) - \xi_i^\beta d_i(u, \hat{\rho}_i) \right| && (42) \\
&\leq \sup_{i \geq 0} \inf_{\mathcal{R}_i \in \mathcal{C}_i^{(n)}} \sup_{\left(u^{(n)}, u\right), \left(v^{(n)}, v\right) \in \mathcal{R}_i} \left| \left(\xi_i^{(n)}\right)^\beta d_i^{(n)}\left(u^{(n)}, v^{(n)}\right) - \xi_i^\beta d_i(u, v) \right| && (43) \\
&= 2 \sup_{i \geq 0} d_{GH}^m\left(\left(\xi_i^{(n)}\right)^\beta \tau_i^{(n)}, \xi_i^\beta \tau_i\right) \leq 2 d_\beta\left(\kappa^{(n)}, \kappa\right),
\end{aligned}$$

with (40) and (41) holding by the metric defined in (38), and with (42) holding by the triangle inequality, and with (43) holding as $\left(\hat{\rho}_i^{(n)}, \hat{\rho}_i\right)$ are in correspondence by definition of $\mathcal{C}_i^{(n)}$. Hence, $g_\beta \restriction_{C_\beta}$ is continuous.

Next, we prove that $C_\beta \in \mathcal{B}(\Xi^*)$. By standard topological arguments, the concatenated tree is compact if and only if $\lim_{i \to \infty} \text{ht}\left(\xi_i^\beta \tau_i\right) = 0$. Thus, we have the following representation of C_β,

$$C_\beta = \bigcap_{N \geq 1} \bigcup_{I \geq 0} \bigcap_{i \geq I} \{\kappa = (\xi_i, \tau_i : i \geq 0) \in \Xi^* : \xi_i^\beta \text{ht}(\tau_i) < 1/N\}.$$

Since $\text{ht}\colon \mathbb{T}_m \to [0, \infty)$ is a continuous function, $C_\beta \in \mathcal{B}(\Xi^*)$. □

We now deduce our main theorem.

Theorem 4.3.4. *The law of the marked (rooted) α-stable tree satisfies the RDE*

$$\mathcal{T}_\alpha \stackrel{d}{=} g_\beta\left(\xi, \mathcal{T}_i : i \geq 0\right) \quad \text{on } \mathbb{T}_m, \tag{44}$$

where $\mathcal{T}_i \overset{i.i.d}{\sim} \mathcal{T}_\alpha$ *for all* $i \geq 0$ *and* $\xi = (X_0, X_1, X_2, X_3 P_j : j \geq 1) \in \Xi$ *such that:*

- *If* $\alpha = 2$*, then* $\xi_{j+2} = X_3 P_j = 0$ *almost surely for all* $j \geq 1$ *and* $(X_0, X_1, X_2) \sim \mathrm{Dir}\,(1/2, 1/2, 1/2)$*.*
- *If* $\alpha \in (1,2)$*, then* $(X_0, X_1, X_2, X_3) \sim \mathrm{Dir}\,(\beta, \beta, \beta, 1-2\beta)$ *and* $(P_i : i \geq 1) \sim \mathrm{PD}(1-\beta, 1-2\beta)$*, with* (X_0, X_1, X_2, X_3) *and* $(P_i : i \geq 1)$ *independent.*

Recall that $\mathcal{P}(\mathbb{T}_m)$ *denotes the set of Borel probability measures on* $\mathbb{T}_m$*. From* (22)*, RDE* (44) *induces a mapping* $\Phi_\beta \colon \mathcal{P}(\mathbb{T}_m) \to \mathcal{P}(\mathbb{T}_m)$*, to which the law of the marked* α*-stable tree satisfies* $\mathcal{T}_\alpha = \Phi_\beta(\mathcal{T}_\alpha)$*.*

Proof: Recall that for the sub-trees involved in the recursive application of Marchal's algorithm, we regarded V_1 as a root and marked the first leaf in the i-th sub-tree for each $i \geq 1$. We regarded A_0 as the root for the overall tree, and V_1 as a marked leaf for the 0-th sub-tree. Thus, our construction using Marchal's algorithm agrees with the concatenation operator g_β acting on the sub-trees. Theorem 4.2.4 gives the required independences and the distribution of $\xi = (\xi_i : i \geq 0)$. Proposition 4.3.3 ascertains the measurability of g_β. □

However, this is not the only distributional fixpoint. Observe that if the metrics d_i of (the representatives of) $(\tau_i, d_i, \rho_i, x_i) \in \mathbb{T}_m$ were multiplied by some constant $c > 0$, then the concatenated tree will also have its metric d' multiplied by c. Furthermore, if the original concatenated tree were a marked compact rooted $\mathbb{R}$-tree, then so would the concatenated tree with metric multiplied by c. Thus, since the law of $(\mathcal{T}_\alpha, d_\alpha, \rho_\alpha, x_\alpha)$ is a fixpoint of RDE (44), then so is $(\mathcal{T}_\alpha, c d_\alpha, \rho_\alpha, x_\alpha)$.

We adapt the following from Broutin and Sulzbach [14, Proposition 3(i)], which generalises an example for the BCRT in [4]. It gives us another distributional fixpoint for RDE (44). While their construction uses the Gromov-Hausdorff-Prokhorov topology with a finite concatenation operation, we verify that it can be extended to the marked Gromov-Hausdorff topology when concatenating countably infinitely many sub-trees.

Proposition 4.3.5. *There exists a solution to RDE* (44) *which does not have finite* $1/\beta$*-th height moment.*

Proof: For $i \geq 0$, let $H^{(i)}$ be the (i-th copy of the) height process associated with the α-stable tree, and let $\mathcal{T}_i$ be the i-th α-stable tree coded by $H^{(i)}$. Fix $\eta_0 \in (0,1)$, so it corresponds to some point x_0 on $\mathcal{T}_0$ under the canonical projection $\tau_{H^{(0)}}$. Recalling the scaling property in (14), rescaling $H^{(i)}$ by factor $\xi_i > 0$ along the x-axis to get a continuous excursion of length ξ_i results in the mass measure of $\mathcal{T}_i$ being rescaled by ξ_i and the metric being rescaled by ξ_i^β. We then define a continuous normalised excursion $\tilde{H}$ such that the disjoint intervals $S_0 := [0, \xi_0 - \xi_0\eta_0] \cup (1 - \xi_0\eta_0, 1]$ correspond to $H^{(0)}$ rescaled by ξ_0, and $S_i := \left(\sum_{j=0}^{i-1} \xi_j - \xi_0\eta_0, \sum_{j=0}^{i} \xi_j - \xi_0\eta_0\right]$ correspond to the rescaled $H^{(i)}$ which are now excursions from level $H^{(0)}_{1-\eta_0}$ for $i \geq 1$. Observe that for the coded trees $\mathcal{T}_i$, this process is analogous to the concatenation operator g_β acting on them. Denote the tree coded by $\tilde{H}$ as $\mathcal{T}_\alpha$, which has the law of the α-stable tree by Theorem 4.3.4.

For each $i \geq 0$, run an independent Poisson point process (PPP) on $[0,1] \times [0,\infty)$ with intensity $ds_i \otimes y_i^{-(1/\beta+1)} dy_i$. For each point (s_i, y_i) of the PPP, graft a massless branch of length y_i to the point on $\mathcal{T}_i$ corresponding to s_i under the canonical projection $\tau_{H^{(i)}}$. Almost surely, only finitely many of these branches have length longer than any given $\epsilon > 0$, since $\int_\epsilon^\infty y_i^{-(1/\beta+1)} dy_i = \beta\epsilon^{-1/\beta} < \infty$. Hence, this process of decorating $\mathcal{T}_i$ with massless branches yields an almost surely compact space.

Furthermore, by construction, the GH^m-equivalence class of the decorated space only depends on the GH^m-equivalence classes associated with $\mathcal{T}_i$. We now show that the process of concatenation and decoration by branches commute.

Define $R_i := S_i \times [0,\infty)$ for $i \geq 0$. We still have a point process on $[0,1] \times [0,\infty)$ after concatenation formed by the induced PPP within each R_i. Given disjoint Borel measurable sets $A^{(j)} \subseteq [0,1] \times [0,\infty)$ for $j = 1,\dots,n$, define $A_i^{(j)} := A^{(j)} \cap R_i$ which are Borel measurable and disjoint in R_i. By assumption, for each i, the count functions $\left(N\left(A_i^{(j)}\right) : j = 1,\dots,n\right)$ are independent. For fixed j, $N\left(A^{(j)}\right)$ is measurably constructed from $\left(N\left(A_i^{(j)}\right) : i \geq 0\right)$, so $\left(N\left(A^{(j)}\right) : j = 1,\dots,n\right)$ are independent. Given $A \subseteq [0,1] \times [0,\infty)$ Borel measurable, define $A_i := A \cap R_i$. We obtain R_i in $(\tilde{s},\tilde{y})$-coordinates from (s_i,y_i)-coordinates by the invertible transformation,

$$\tilde{s} = \xi_0 s_0 \mathbf{1}\{s_0 \leq 1-\eta_0\} + (\xi_0 s_0 + (1-\xi_0))\,\mathbf{1}\{s_0 > 1-\eta_0\} \qquad \text{and} \qquad \tilde{y} = \xi_i^{\beta} y_i \qquad \text{for } i = 0,$$

$$\tilde{s} = \sum_{j=0}^{i-1} \xi_j - \xi_0\eta_0 + \xi_i s_i \qquad \text{and} \qquad \tilde{y} = \xi_i^{\beta} y_i \qquad \text{for } i \geq 1.$$

Using the Mapping Theorem in [45], the induced measure under rescaling is still non-atomic, so we have a PPP in R_i with intensity given by change of variables,

$$\tilde{\lambda}(d\tilde{s},d\tilde{y}) = \frac{1}{\xi_i} d\tilde{s} \otimes \left(\frac{\tilde{y}}{\xi_i^{\beta}}\right)^{-(1/\beta+1)} \frac{1}{\xi_i^{\beta}} d\tilde{y} = d\tilde{s} \otimes \tilde{y}^{-(1/\beta+1)} d\tilde{y}.$$

As A_i are disjoint, the Monotone Convergence Theorem gives that the count $N(A) = \sum_{i=0}^{\infty} N(A_i)$. Let $\mathrm{Po}(\lambda)$ denote a Poisson distribution with mean λ, then by the Countable Additivity Theorem in [45] and independence of $(N(A_i) : i \geq 0)$ by construction,

$$N(A) \sim \mathrm{Po}\left(\sum_{i=0}^{\infty} \iint_{A_i} d\tilde{s} \otimes \tilde{y}^{-(1/\beta+1)} d\tilde{y}\right) = \mathrm{Po}\left(\iint_{A} d\tilde{s} \otimes \tilde{y}^{-(1/\beta+1)} d\tilde{y}\right),$$

where equality follows by the Monotone Convergence Theorem. So, the concatenation operation induces a PPP of the desired intensity. Denote the decoration as inducing a mapping $\Psi\colon \mathcal{P}(\mathbb{T}_m) \to \mathcal{P}(\mathbb{T}_m)$, we conclude that Φ_β and Ψ commute. This implies that since $\mathcal{T}_\alpha$ is a fixpoint distribution of Φ_β, then so is $\Psi(\mathcal{T}_\alpha)$, the law of the decorated version of $\mathcal{T}_\alpha$. Denote a representative of the decorated $\mathcal{T}_\alpha$ by $\mathcal{T}$, then the height of $\mathcal{T}$ is at least the length of the longest segment, so for any $h > 0$,

$$\mathbb{P}(\mathrm{ht}(\mathcal{T}) \geq h) \geq \mathbb{P}\left(\mathrm{Po}\left(\int_h^{\infty} y^{-(1/\beta+1)} dy\right) \geq 1\right) = 1 - \exp\left(-\beta h^{-1/\beta}\right). \tag{45}$$

The above implies that

$$\mathbb{P}\left(\mathrm{ht}(\mathcal{T}) \geq h^{\beta}\right) \geq 1 - \exp\left(-\beta h^{-1}\right). \tag{46}$$

However, $1 - \exp\left(-\beta h^{-1}\right)$ is not integrable on $(0,\infty)$, so it follows that $\mathbb{E}\left[\mathrm{ht}(\mathcal{T})^{1/\beta}\right] = \infty$. □

One might find the above unsatisfactory as $\mathcal{T}$ fails to be a continuum tree, since $\Psi(\mathcal{T}_\alpha)$ assigns zero mass to a branch when decomposed at a decorated branch point. We refer the reader to [14, Proposition 3(ii)] for a more involved counterexample which is a continuum tree when $\alpha = 2$.

Furthermore, an α-stable tree has the desirable property of having finite height moments of all orders. We cite the following result which relies on the fact that α-stable trees may be constructed via self-similar fragmentation processes with negative index [39].

Proposition 4.3.6. *[38, Proposition 14] There exists a constant $C_\alpha > 0$, such that for $t > 0$ sufficiently large, $\mathbb{P}(\mathrm{ht}(\mathcal{T}_\alpha) > t) \leq \exp(-C_\alpha t)$. In particular, given $p \geq 1$, writing $\mathbb{E}[\mathrm{ht}(\mathcal{T}_\alpha)^p] = \int_0^\infty \mathbb{P}\left(\mathrm{ht}(\mathcal{T}_\alpha) > t^{1/p}\right) dt$, the α-stable tree has finite p-th height moment.*

Sharper bounds do exist, the above was recently refined to $\mathbb{P}(\mathrm{ht}(\mathcal{T}_\alpha) > t) \leq \exp(-\tilde{C}_\alpha t^\alpha)$ for sufficiently large $t > 0$ [32, Theorem 1.5]. This begs the question of whether the α-stable tree is the unique fixpoint to RDE (44) with finite height moments of all orders, up to rescaling in its metric. This motivates scrutiny of Albenque and Goldschmidt's [4], and Broutin and Sulzbach's [14], recursive construction of the BCRT, since they establish uniqueness and attractiveness results up to a stronger Gromov-Hausdorff-Prokhorov topology.

5 Recursive Construction of the BCRT in the Literature

We study the arguments used in [4] and [14] to discern their applicability in extending uniqueness and attractiveness results to the general α-stable tree as a distributional fixpoint. We will reformulate results cited from [4] and [14] to cohere with our definition of a BCRT being coded by $\sqrt{2}$ times a normalised Brownian excursion.

5.1 Description of Constructions

We present Albenque and Goldschmidt's construction. Suppose M is a law of a CRT, for $i \in \{0, 1, 2\}$,

(i) Sample $(\mathcal{T}_i, d_i, \mu_i) \overset{i.i.d}{\sim} M$ and $X_i \sim \mu_i$ in $\mathcal{T}$ independently.

(ii) Sample $\Delta = (\Delta_0, \Delta_1, \Delta_2) \sim \mathrm{Dir}(1/2, 1/2, 1/2)$ independently.

(iii) Rescale the trees $\left(\mathcal{T}_i, \Delta_i^{1/2} d_i, \Delta_i \mu_i\right)$.

(iv) Identify the points X_0, X_1 and X_2 in the rescaled trees. Concatenate the rescaled trees at those points to form a larger tree $(\mathcal{T}^o, d, \mu)$ with a marked branch point, where μ is the measure induced on $\mathcal{T}^o$ by the measures $\Delta_i \mu_i$.

(v) Forget the marked branch point to obtain $(\mathcal{T}, d, \mu)$. Define $\Phi_{(\mathrm{a})}(M)$ as the distribution of $(\mathcal{T}, d, \mu)$.

We then obtain an RDE,

$$M = \Phi_{(\mathrm{a})}(M). \tag{47}$$

Broutin and Sulzbach's constructions work on a higher generality. We provide their construction specific to the BCRT case, and refer the reader to [14, Section 2.3] for more information. They work over GHP-equivalence classes of compact rooted measured metric spaces, not necessarily $\mathbb{R}$-trees, call this space $\mathbb{K}_w$. Their construction grafts trees onto a *structural tree* Γ, where Γ is a rooted plane tree with a specified tree-shape, ordered by depth-first lexicographical order. For the the BCRT, we take Γ to be the tree on vertex set $\{0,1,2\}$ with vertices 1 and 2 being children of 0. Given compact rooted measured metric spaces $(\mathcal{K}_i, d_i, \rho_i, \mu_i) \in \mathbb{K}_w$ for $i \in \{0,1,2\}$,

(i) Sample $\eta_i \sim \mu_i$ in $\mathcal{K}_i$ independently.

(ii) Let $\mathcal{K}^o$ be the disjoint union of $\mathcal{K}_i$ for $i \in \{0,1,2\}$. Identify ρ_1 and ρ_2 with η_0. Define $\mathcal{K}$ as the resulting space after identification. Let ρ be identified with ρ_0 in $\mathcal{K}$.

(iii) Sample $r = (r_0, r_1, r_2) \sim \mathrm{Dir}(1/2, 1/2, 1/2)$. Define a metric d and measure on $\mathcal{K}$ compatible with $r_i^{1/2} d_i$ and $r_i \mu_i$ when restricted to the identification of $\mathcal{K}_i$ in $\mathcal{K}$.

We then have an RDE,

$$\mathcal{K} \overset{d}{=} g_{(\mathrm{b})}\left(r, \mathcal{K}_0, \mathcal{K}_1, \mathcal{K}_2\right) \qquad \text{on } \mathbb{K}_w, \tag{48}$$

The construction then induces a mapping $\Phi_{(\mathrm{b})} \colon \mathcal{P}(\mathbb{K}_w) \to \mathcal{P}(\mathbb{K}_w)$.

5.2 Uniqueness and Attractiveness of the BCRT as a Fixpoint

We compare the main uniqueness results obtained via each construction.

Theorem 5.2.1. *[4, Theorem 1.6] The law of the BCRT solves* (47)*. If M solves* (47)*, then, there exists $\gamma > 0$ such that if $(\mathcal{T}, d, \mu) \sim M$, then $(\mathcal{T}, \gamma d, \mu)$ has the law of the BCRT.*

Theorem 5.2.2. *[14, Theorem 1] If $0 < c < \infty$, then $(\mathcal{T}_2, cd_2, \rho_2, \mu_2)$ is the unique $\mathbb{R}$-tree satisfying* (48) *and $\mathbb{E}\left[d(\rho, \zeta)\right] = c\sqrt{\pi}/2$, where $\zeta \sim \mu_2$.*

Uniqueness of solutions for the respective RDEs only holds up to a scalar multiple in the metric. Closer scrutiny of the proofs for these results yields deeper connections. Albenque and Goldschmidt, conditional on $(\mathcal{T}, d, \mu) \sim M$ and $\Delta \sim \mathrm{Dir}\left(1/2, 1/2, 1/2\right)$, sample two uniform points according to μ and denote by $P = (P_0, P_1, P_2)$ the number of points picked in each sub-tree. Thus, P follows a multinomial distribution with $n = 2$ and parameters $(\Delta_0, \Delta_1, \Delta_2)$. Denote by D the distance between two uniformly sampled points under μ, then

$$D \overset{d}{=} \Delta_0^{1/2} \mathbf{1}_{\{P_0 > 0\}} D_0 + \Delta_1^{1/2} \mathbf{1}_{\{P_1 > 0\}} D_1 + \Delta_2^{1/2} \mathbf{1}_{\{P_2 > 0\}} D_2 \qquad \text{on } \mathbb{R}^+, \tag{49}$$

where $D_i \overset{i.i.d}{\sim} D$ for $i \in \{0,1,2\}$, independently of P and Δ. Broutin and Sulzbach's inductive proof of Theorem 5.2.2 reduces to showing

$$Y \overset{d}{=} r_0^{1/2} Y_0 + r_1^{1/2} \mathbf{1}_{\{J(\zeta)=1\}} Y_1 + r_2^{1/2} \mathbf{1}_{\{J(\zeta)=2\}} Y_2 \qquad \text{on } \mathbb{R}^+, \tag{50}$$

where $Y := d(\rho, \zeta)$, $\zeta \sim \mu$, $r \sim \mathrm{Dir}(1/2, 1/2, 1/2)$, and Y_0, Y_1 and Y_2 are independent copies of Y which are independent of r and ζ. We denote $J(\zeta) = i$ if ζ belongs in the sub-component associated with $\mathcal{K}_i$ in $\mathcal{K}$. The difference between (49) and (50) rests on the fact that the former depends on the relative positions between two sampled points, while the latter depends on which sub-component the sampled point falls with respect to the root. Both rely on the log-smoothing transform introduced by Durrett and Liggett [33, Theorems 1 and 2(b)] to prove uniqueness. As Durrett and Liggett's theorems yield uniqueness up to a multiplicative constant, this explains the parallel between Theorems 5.2.1 and 5.2.2. Apart from identifying a unique solution, both constructions yield that the BCRT fixpoint is attractive.

Theorem 5.2.3. *[4, Theorem 1.7] Denote the law of the unrooted BCRT by $\mathbb{M}$. Given $(\mathcal{T}, d, \mu) \sim M$, sample $V_1, V_2 \sim \mu$ independently. Suppose that $\mathbb{E}[d(V_1, V_2)] = \sqrt{\pi}/2$, denoting $M_n = \Phi^n_{(a)} M$, then $M_n \to \mathbb{M}$ as $n \to \infty$ with respect to weak convergence of measures in the Gromov-Prokhorov topology.*

Theorem 5.2.4. *[14, Theorem 2] Define $\mathbb{T}^*_w \subseteq \mathbb{T}_w$ by $\mathbb{T}^*_w := \{(\mathcal{T}, d, \rho, \mu) \in \mathbb{T}_w : \mathrm{supp}(\mu) = \mathcal{T}\}$. In particular, $\mathbb{T}^*_w$ contains all continuum trees, which are leaf-dense and have their measure supported on leaves. Suppose that $\nu \in \mathcal{P}(\mathbb{T}^*_w)$, with $\mathbb{E}_\nu[d(\rho, \zeta)] = c\sqrt{\pi}/2$ and $\mathbb{E}_\nu\left[\mathrm{ht}(\mathcal{T})^3\right] < \infty$. Let $\nu_n = \Phi^n_{(b)}\nu$, then $\nu_n \to \mathcal{T}_2$ as $n \to \infty$ with respect to the weak convergence of measures in the Gromov-Hausdorff-Prokhorov topology.*

By virtue of the BCRT being invariant under uniform re-rooting [41, Theorem 11], if Broutin and Sulzbach's construction were restricted to continuum trees and modified to use a uniform pick of two points, then Theorem 5.2.4 (with $c = 1$) would extend Theorem 5.2.3, contingent on finite third height moments.

To prove attractiveness, Albenque and Goldschmidt use a coupling argument based on reduced sub-trees. Denote by $S_n(m)$ and $R(m)$ the m-th reduced sub-trees of of $\mathcal{U}_n \sim M_n$ (as in Theorem 5.2.3) and the BCRT respectively. Convergence in the Gromov-Prokhorov topology is equivalent to showing $S_n(m) \xrightarrow{d} R(m)$ as $n \to \infty$ for every $m \geq 2$. The coupling argument simplifies this to (49), that is, showing the convergence for $m = 2$.

In justifying a stronger Gromov-Hausdorff-Prokhorov convergence, Broutin and Sulzbach rely on the machinery of the *distance matrix* associated with $\mathcal{K} \in \mathbb{K}_w$, defined as $\mathfrak{D}_{\mathcal{K}} := (d(\zeta_i, \zeta_j))_{i,j \geq 0}$, where $\zeta_0 := \rho$ and $\zeta_i \overset{i.i.d}{\sim} \mu$ for $i \geq 1$. Let $\mathcal{D}^m_{\mathcal{K}}$ denote the m-th sub-matrix $(d(\zeta_i, \zeta_j))_{0 \leq i,j \leq m}$. As the class of distance matrices is convergence-determining in the Gromov-Hausdorff-Prokhorov topology [23, Theorem 5], it suffices to prove that $\mathfrak{D}^m_{\mathcal{S}_n} \xrightarrow{d, L^1} \mathfrak{D}^m_{\mathcal{T}_2}$ where $\mathcal{S}_n \sim \nu_n$, as defined in Theorem 5.2.4, for all $m \geq 1$. An inductive argument reduces to proving this for $m = 1$.

We observe that convergence of the 1-st reduced sub-matrix $\mathfrak{D}^1_{\mathcal{S}_n}$ only depends on the convergence in distribution in the height of a uniformly chosen point in $\mathcal{S}_n$. This is a direct analogue of how Albenque and Goldschmidt's argument reduces to the convergence in distribution between two uniformly sampled points in $\mathcal{U}_n$. The distance matrix encodes the distribution of tree-shapes of the reduced sub-trees by the distances between each pair of chosen points in the reduced sub-tree, thereby linking both results.

5.3 Relation to Results Obtained

For the BCRT, our recursive construction using one identified point in each rooted sub-tree is exactly Broutin and Sulzbach's construction, agreeing up to the rescaled measures in each sub-tree from Theorem 4.2.4. In this setting, we can conclude that the law of the (marked) BCRT is the unique and attractive fixpoint of our RDE (44) up to scalar multiplication in the metric and under finite third height moments. This even holds in a stronger (marked) Gromov-Hausdorff-Prokhorov topology. We might then be tempted to mimic (50) for a general α-stable tree. Suppose that measures could be incorporated into our concatenation operation g_β, we would get an RDE

$$Y \overset{d}{=} \xi_0^\beta Y_0 + \sum_{j=1}^{\infty} \xi_j^\beta \mathbf{1}_{\{J(\zeta)=j\}} Y_j \quad \text{on } \mathbb{R}^+, \tag{51}$$

where $Y := d(\rho, \zeta)$, $\zeta \sim \mu_\alpha$, and $(Y_i : i \geq 0) \overset{i.i.d}{\sim} Y$, independent of ξ and ζ. Denote $J(\zeta) = i$ if ζ belongs in the sub-tree associated with τ_i in $\mathcal{T}$. Recall that $\xi = (X_0, X_1, X_2, X_3 P_i : i \geq 1)$ such that $(X_0, X_1, X_2, X_3) \sim \mathrm{Dir}(\beta, \beta, \beta, 1-2\beta)$ and $(P_i : i \geq 1)^{\downarrow} \sim \mathrm{PD}(1-\beta, 1-2\beta)$ are independent. Aldous and Bandyopadhyay [9, Theorem 16] showed that standard arguments to verify uniqueness of fixpoints are untenable on RDEs of the form

$$X \overset{d}{=} \sum_{i=0}^{N} \eta_i X_i \quad \text{on } \mathbb{R}^+, \tag{52}$$

where $\mathbb{E}\left[\sum_{i=0}^{N} \eta_i\right] = 1$, under technical assumptions on the moments of N and $\eta = (\eta_i : i \geq 0)$. When N is almost surely bounded, as in (49) and (50), uniqueness up to multiplicative constants is handled by Durrett and Liggett's results. An analogous result holds for the α-stable tree utilising results from Iksanov [43].

Proposition 5.3.1. *In the setting of RDE* (51),

$$\mathbb{E}\left[\xi_0^{\beta} + \sum_{j=1}^{\infty} \xi_j^{\beta} \mathbf{1}_{\{J(\zeta)=j\}}\right] = 1. \tag{53}$$

Among solutions of (51) *with finite mean, the distribution of the distance between the root and a uniformly sampled point on $\mathcal{T}_\alpha$ is the unique solution of* (51) *up to a multiplicative constant.*

Proof: From the aggregation property in Proposition 3.2.2,

$$X_3 \sim \mathrm{Dir}(1-2\beta, 3\beta) \quad \text{and} \quad X_j \sim \mathrm{Dir}(\beta, 1) \quad \text{for } j = 0, 1, 2.$$

We calculate

$$\begin{aligned}
\mathbb{E}\left[X_0^{\beta}\right] &= \frac{\Gamma(2\beta)\Gamma(\beta+1)}{\Gamma(2\beta+1)\Gamma(\beta)} = \frac{1}{2}, \\
\mathbb{E}\left[X_j^{\beta+1}\right] &= \frac{\Gamma(2\beta+1)\Gamma(\beta+1)}{\Gamma(2\beta+2)\Gamma(\beta)} = \frac{\beta}{2\beta+1}, \quad \text{for } j = 1, 2, \\
\mathbb{E}\left[X_3^{\beta+1}\right] &= \frac{\Gamma(2-\beta)\Gamma(\beta+1)}{\Gamma(2\beta+2)\Gamma(1-2\beta)}.
\end{aligned}$$

From Theorem 3.2.6 and recalling $(P_i : i \geq 1)^{\downarrow} \sim \mathrm{PD}(1-\beta, 1-2\beta)$, we get that

$$\mathbb{E}\left[P_1^{\beta+1}\right] = \frac{\Gamma(2-2\beta)\Gamma(2\beta+1)}{\Gamma(3-\beta)\Gamma(\beta)},$$

and that, for $j \geq 2$,

$$\begin{aligned}
\mathbb{E}\left[P_j^{\beta+1}\right] &= \frac{\Gamma(2-2\beta)\Gamma(2\beta+1)}{\Gamma(3-\beta)\Gamma(\beta)} = \prod_{k=1}^{j-1} \mathbb{E}\left[\overline{W}_k^{\beta+1}\right] \times \mathbb{E}\left[W_j^{\beta+1}\right] \\
&= \prod_{k=1}^{j-1} \frac{\Gamma((k+2)-(k+1)\beta)\Gamma((k+1)-(k+1)\beta)}{\Gamma((k+1)-(k+2)\beta)\Gamma((k+2)-k\beta)} \times \frac{\Gamma((j+1)-(j+1)\beta)\Gamma(1+2\beta)}{\Gamma((j+2)-j\beta)\Gamma(\beta)} \\
&= \frac{\Gamma(2-2\beta)\Gamma(2\beta+1)}{\Gamma(3-\beta)\Gamma(\beta)} \times \prod_{k=2}^{j} \frac{k-(k+1)\beta}{(k+1)-k\beta}.
\end{aligned}$$

Writing $\delta = \beta/(1-\beta) \in (0,1)$, we are able to evaluate the following series

$$\begin{aligned}
\sum_{j=1}^{\infty}\prod_{k=1}^{j}\frac{k-(k+1)\beta}{(k+1)-k\beta} &= \sum_{j=1}^{\infty}\prod_{k=1}^{j}\frac{k-\delta}{k+\delta+1} \\
&= -1+\sum_{j=0}^{\infty}\frac{(1-\delta)_j(1)_j}{(\delta+2)_j\ j!} \qquad (54)\\
&= -1+\lim_{z\to 1}F_{2,1}(1-\delta,1,\delta+2;z) \qquad (55)\\
&= -1+\frac{\Gamma(\delta+2)\Gamma(2\delta)}{\Gamma(\delta+1)\Gamma(2\delta+1)} \qquad (56)\\
&= -1+\frac{1+\delta}{2\delta} = \frac{1-2\beta}{2\beta},
\end{aligned}$$

where we used the rising Pochhammer symbol in (54), which is defined as

$$(\delta)_j = \begin{cases} 1 & \text{if } j=0,\\ \delta(\delta+1)\dots(\delta+n-1) & \text{if } j=n\geq 1,\end{cases}$$

and the hypergeometric function in (55) defined as

$$F_{2,1}(a,b,c;z) := \sum_{j=0}^{\infty}\frac{(a)_j(b)_j}{(c)_j\ j!}z^j \qquad \text{for } |z|<1,$$

and is defined at $z=1$ if $c-b-a>0$ with limit

$$\frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-b)\Gamma(c-a)},$$

from [37, Display 9.122.1]. Since $c-b-a=2\delta>0$ in our case, the equality in (56) follows. Consolidating our results to establish (53),

$$\begin{aligned}
\mathbb{E}\left[\xi_0^\beta+\sum_{j=1}^{\infty}\xi_j^\beta\mathbf{1}_{\{J(\zeta)=j\}}\right] &= \frac{1}{2}+\sum_{j=1}^{\infty}\mathbb{E}\left[\xi_j^\beta\mathbf{1}_{\{J(\zeta)=j\}}\right] \qquad (57)\\
= \frac{1}{2}+\sum_{j=1}^{\infty}\mathbb{E}\left[\xi_j^\beta\mathbb{P}\left(J(\zeta)=j\mid\xi_j\right)\right] &= \frac{1}{2}+\sum_{j=1}^{\infty}\mathbb{E}\left[\xi_j^{\beta+1}\right] \qquad (58)\\
= \frac{1}{2}+\sum_{j=1}^{2}\mathbb{E}\left[X_j^{\beta+1}\right]+\mathbb{E}\left[X_3^{\beta+1}\right]\sum_{j=1}^{\infty}\mathbb{E}\left[P_j^{\beta+1}\right] & \qquad (59)\\
= \frac{1}{2}+\frac{2\beta}{2\beta+1}+\frac{\Gamma(2-\beta)\Gamma(\beta+1)\Gamma(2-2\beta)\Gamma(2\beta+1)}{\Gamma(2\beta+2)\Gamma(1-2\beta)\Gamma(3-\beta)\Gamma(\beta)}\times\left(1+\sum_{j=2}^{\infty}\prod_{k=2}^{j}\frac{k-(k+1)\beta}{(k+1)-k\beta}\right)&\\
= \frac{1}{2}+\frac{2\beta}{2\beta+1}+\frac{\beta}{2\beta+1}\times\sum_{j=1}^{\infty}\prod_{k=1}^{j}\frac{k-(k+1)\beta}{(k+1)-k\beta}&\\
= \frac{1}{2}+\frac{2\beta}{2\beta+1}+\frac{\beta}{2\beta+1}\times\frac{1-2\beta}{2\beta} = 1,&
\end{aligned}$$

where (57) invokes the Monotone Convergence Theorem for series, (58) follows from the Law of Total Expectation recalling that $\mathbb{P}\left(J(\zeta) = j \mid \xi_j\right) = \xi_j$ for $j \geq 1$ under the rescaled measures, and (59) uses the independence of (X_0, X_1, X_2, X_3) and $(P_i : i \geq 1)$. The calculations involving the hypergeometric function may be bypassed by noting that P_1 is distributed as a size-biased pick from $(P_i : i \geq 1)$ [58]. We did not pursue this here as we did not introduce the CRP in such depth.

Let Y be the random variable denoting the distance between the root and a uniformly sampled point on $\mathcal{T}_\alpha$. From [3, Lemma 11] and (26),

$$Y \sim \frac{1}{\alpha}\mathrm{ML}(\beta, \beta).$$

Using (25), Y has finite mean,

$$\mathbb{E}\left[Y\right] = \frac{2\Gamma(\beta+1)}{\alpha\Gamma(2\beta+1)}.$$

By construction, Y a distributional fixpoint of RDE (51). This implies constant multiples of Y are also fixpoints of RDE (51). Thus, a 1-elementary fixpoint exists in the setting of Iksanov [43]. The hypotheses for [43, Proposition 3] are met using [43, Theorem 2], since a 1-elementary fixpoint exists and (53) holds. [43, Proposition 3] then gives that constant multiples of Y are the only fixpoints of RDE (51) with finite mean. □

We believe that the coupling argument of [4] is instrumental in proving the uniqueness and attractiveness of the α-stable tree as a fixpoint. The statement of Proposition 5.3.1 parallels [4, Proposition 2.1]. Issues on labelling and determination of gluing-points in [4] are likewise plausibly argued for the infinitary case. As the distribution of an α-stable tree is determined by the distributions of its reduced sub-trees [29, Theorem 2.2.1], an analogue of [4, Theorem 1.6] then holds for α-stable trees. This will give uniqueness, up to scalar multiplication in the metric, for the law of the α-stable tree.

Recall that Y denotes the distance between the root and a uniformly sampled point on the α-stable tree, which is a fixpoint of RDE (51) with finite mean. From Proposition 4.3.6,

$$\mathbb{E}\left[Y^p\right] \leq \mathbb{E}\left[\mathrm{ht}(\mathcal{T}_\alpha)^p\right] < \infty \quad \text{for all } p \geq 1. \tag{60}$$

[43, Proposition 4] and Proposition 5.3.1 yield that the conditions for [43, Proposition 6] hold. Restrict attention to distributions on $\mathbb{R}^+$ which have mean equal to

$$\mathbb{E}\left[Y\right] = \frac{2\Gamma(\beta+1)}{\alpha\Gamma(2\beta+1)},$$

and which have finite $(1+\epsilon)$ moment for some $\epsilon > 0$. This is no assumption at all if we seek fixpoints of RDE (44) with finite height moments of all orders. By applying Banach's Contraction Mapping Theorem in the context of Aldous and Bandyopadhyay [9, Lemma 5] to the result of [43, Proposition 6], Y is the unique and attractive fixpoint in the space of such distributions. Following the proof of [4, Theorem 1.7] will then give an attractiveness result for the law of the α-stable tree in the Gromov-Prokhorov topology.

However, an outstanding issue is how to revise our concatenation operator to handle unrooted and unmarked trees, or to adapt the operation in [4] for rooted and marked trees. Furthermore, extending this to the Gromov-Hausdorff-Prokhorov topology entails deeper analysis of the distance matrix results in [14], or arguments utilised by Rembart and Winkel [61, Theorem 1.5]. Alternatively, we could verify the technical hypotheses in [7, Corollary 19] to deduce Hausdorff convergence on the supports of weakly convergent measures.

6 Questions and Extensions

We summarise further possibilities and directions arising from this dissertation.

- We conjecture that the law of the marked α-stable tree is the unique solution with finite height moments of all orders to RDE (44), modulo scalar multiplication in the metric, in the marked Gromov-Hausdorff topology. We further conjecture that this also holds in the marked Gromov-Hausdorff-Prokhorov topology.
- Due to word limitations, we did not present how our results relate to the RDE for an α-stable tree presented by Rembart and Winkel [61, Theorem 1.5]. We could analyse whether our decomposition at the first branch point connects to spinal partitions [41], or if resulting mass splits observed link up with bead-splitting processes [59].
- Explore recursive decompositions of various statistics involving α-stable trees. For example, an RDE was formulated for the diameter of the BCRT [19, Display (14)]. Can this be extended for an α-stable tree?

References

[1] R. Abraham and J.-F. Delmas. β-coalescents and stable Galton-Watson trees. arXiv:1303.6882, 2013.

[2] 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(14):1–21, 2013.

[3] L. Addario-Berry, D. Dieuleveut, and C. Goldschmidt. Inverting the cut-tree transform. arXiv:1606.04825, 2016.

[4] M. Albenque and C. Goldschmidt. The Brownian continuum random tree as the unique solution to a fixed point equation. *Electron. Commun. Probab*, 20(61):1–14, 2015.

[5] D. Aldous. The continuum random tree I. *The Annals of Probability*, pages 1–28, 1991.

[6] D. Aldous. The continuum random tree II: an overview. *Stochastic Analysis*, 167:23–70, 1991.

[7] D. Aldous. The continuum random tree III. *The Annals of Probability*, pages 248–289, 1993.

[8] D. Aldous. Recursive self-similarity for random trees, random triangulations and Brownian excursion. *The Annals of Probability*, pages 527–545, 1994.

[9] D. Aldous and A. Bandyopadhyay. A survey of max-type recursive distributional equations. *Annals of Applied Probability*, pages 1047–1110, 2005.

[10] A. Bandyopadhyay. Bivariate uniqueness in the logistic recursive distributional equation. arXiv:math/0401389, 2003.

[11] J. Berestycki, N. Berestycki, and J. Schweinsberg. Beta-coalescents and continuous stable random trees. *The Annals of Probability*, pages 1835–1887, 2007.

[12] J. Bertoin. Self-similar fragmentations. *Annales de l'IHP Probabilités et Statistiques*, 38(3):319–340, 2002.

[13] D. Blackwell and J. B. MacQueen. Ferguson distributions via Pólya urn schemes. *The Annals of Statistics*, pages 353–355, 1973.

[14] N. Broutin and H. Sulzbach. Self-similar real trees defined as fixed-points and their geometric properties. arXiv:1610.05331, 2016.

[15] P. Buneman. A note on the metric properties of trees. *Journal of Combinatorial Theory, Series B*, 17(1):48–50, 1974.

[16] D. Burago, Y. Burago, and S. Ivanov. *A course in metric geometry*, volume 33. American Mathematical Society Providence, 2001.

[17] L. Chaumont and J. C. P. Millan. On the genealogy on conditioned stable Lévy forest. arXiv:0706.2605, 2007.

[18] B. Chen, D. Ford, and M. Winkel. A new family of Markov branching trees: the alpha-gamma model. *Electron. J. Probab*, 14(15):400–430, 2009.

[19] D. Croydon and B. Hambly. Self-similarity and spectral asymptotics for the continuum random tree. *Stochastic Processes and Their Applications*, 118(5):730–754, 2008.

[20] D. Croydon and B. Hambly. Spectral asymptotics for stable trees. *Electron. J. Probab*, 15(57):1772–1801, 2010.

[21] N. Curien and B. Haas. The stable trees are nested. *Probability Theory and Related Fields*, 157(3-4):847–883, 2013.

[22] N. Curien and J.-F. Le Gall. Random recursive triangulations of the disk via fragmentation theory. *The Annals of Probability*, pages 2224–2270, 2011.

[23] A. Depperschmidt, A. Greven, and P. Pfaffelhuber. Marked metric measure spaces. *Electron. Commun. Probab*, 16(17):174–188, 2011.

[24] D. Dieuleveut. The vertex-cut-tree of Galton-Watson trees converging to a stable tree. *The Annals of Applied Probability*, 25(4):2215–2262, 2015.

[25] B. Duplantier, J. Miller, and S. Sheffield. Liouville quantum gravity as a mating of trees. arXiv:1409.7055, 2014.

[26] T. Duquesne. A limit theorem for the contour process of conditioned Galton-Watson trees. *The Annals of Probability*, pages 996–1027, 2003.

[27] T. Duquesne. The coding of compact real trees by real valued functions. arXiv:math.PR/0604106, 2006.

[28] T. Duquesne. The exact packing measure of Lévy trees. *Stochastic Processes and Their Applications*, 122(3):968–1002, 2012.

[29] T. Duquesne and J.-F. Le Gall. *Random trees, Lévy processes and spatial branching processes*, volume 281. Société Mathématique de France, 2002.

[30] T. Duquesne and J.-F. Le Gall. The Hausdorff measure of stable trees. arXiv:math/0509690, 2005.

[31] T. Duquesne and J.-F. Le Gall. Probabilistic and fractal aspects of Lévy trees. *Probability Theory and Related Fields*, 131(4):553–603, 2005.

[32] T. Duquesne and M. Wang. Decomposition of Lévy trees along their diameter. arXiv:1503.05069, 2015.

[33] R. Durrett and T. M. Liggett. Fixed points of the smoothing transformation. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 64(3):275–301, 1983.

[34] S. N. Evans. *Probability and real trees: École d'Été de Probabilités de Saint-Flour XXXV-2005.* 2008.

[35] J. Felsenstein. *Inferring phylogenies*, volume 2. Sinauer Associates Sunderland, 2004.

[36] C. Goldschmidt and B. Haas. A line-breaking construction of the stable trees. *Electron. J. Probab*, 20(16):1–24, 2015.

[37] I. S. Gradshteyn and I. M. Ryzhik. *Table of integrals, series, and products, corrected and enlarged edition.* Academic Press, 2014.

[38] B. Haas. Loss of mass in deterministic and random fragmentations. *Stochastic Processes and Their Applications*, 106(2):245–277, 2003.

[39] B. Haas and G. Miermont. The genealogy of self-similar fragmentations with negative index as a continuum random tree. *Electron. J. Probab*, 9:57–97, 2004.

[40] B. Haas, G. Miermont, J. Pitman, and M. Winkel. Continuum tree asymptotics of discrete fragmentations and applications to phylogenetic models. *The Annals of Probability*, pages 1790–1837, 2008.

[41] B. Haas, J. Pitman, and M. Winkel. Spinal partitions and invariance under re-rooting of continuum random trees. *The Annals of Probability*, pages 1381–1411, 2009.

[42] B. Hambly and T. Lyons. Uniqueness for the signature of a path of bounded variation and continuous analogues for the free group. arXiv:math.CA/0507536, 2006.

[43] A. M. Iksanov. Elementary fixed points of the BRW smoothing transforms with infinite number of summands. *Stochastic Processes and Their Applications*, 114(1):27–50, 2004.

[44] S. Janson. Limit theorems for triangular urn schemes. *Probability Theory and Related Fields*, 134(3):417–452, 2006.

[45] J. F. C. Kingman. *Poisson Processes.* Clarendon Press, 1993.

[46] A. Kyprianou. *Introductory lectures on fluctuations of Lévy processes with applications.* Springer Science & Business Media, 2006.

[47] J.-F. Le Gall. Random trees and applications. *Probab. Surv.*, 2:245–311, 2005.

[48] J.-F. Le Gall. Random real trees. In *Annales de la Faculte des sciences de Toulouse: Mathematiques*, volume 15, pages 35–62, 2006.

[49] J.-F. Le Gall and Y. Le Jan. Branching processes in Lévy processes: the exploration process. *The Annals of Probability*, pages 213–252, 1998.

[50] P. Marchal. Constructing a sequence of random walks strongly converging to Brownian motion. *Discrete Mathematics and Theoretical Computer Science*, pages 181–190, 2003.

[51] P. Marchal. A note on the fragmentation of a stable tree. In *Fifth Colloquium on Mathematics and Computer Science*, pages 489–500. Discrete Mathematics and Theoretical Computer Science, 2008.

[52] G. Miermont. Self-similar fragmentations derived from the stable tree I. *Probability Theory and Related Fields*, 127(3):423–454, 2003.

[53] G. Miermont. Self-similar fragmentations derived from the stable tree II: splitting at nodes. *Probability Theory and Related Fields*, 131(3):341–375, 2005.

[54] G. Miermont. Tessellations of random maps of arbitrary genus. arXiv:0712.3688, 2007.

[55] G. Miermont. The Brownian map is the scaling limit of uniform random plane quadrangulations. *Acta Mathematica*, 210(2):319–401, 2013.

[56] Y. Peres. An invitation to sample paths of Brownian motion. *Lecture notes (unpublished)*, 2001.

[57] J. Pitman. Some developments of the Blackwell-MacQueen urn scheme. *Lecture Notes-Monograph Series*, pages 245–267, 1996.

[58] J. Pitman. *Combinatorial Stochastic Processes: Ecole D'Eté de Probabilités de Saint-Flour XXXII-2002*. Springer, 2006.

[59] J. Pitman and M. Winkel. Regenerative tree growth: Markovian embedding of fragmenters, bifurcators, and bead splitting processes. *The Annals of Probability*, pages 2611–2646, 2015.

[60] F. Rembart and M. Winkel. A binary embedding of the stable line-breaking construction. arXiv:1611.02333, 2016.

[61] F. Rembart and M. Winkel. Recursive construction of continuum random trees. arXiv:1607.05323, 2016.

[62] J.-L. Rémy. Un procédé itératif de dénombrement d'arbres binaires et son application à leur génération aléatoire. *RAIRO, Informatique Théorique*, 19(2):179–195, 1985.

[63] U. Roesler and L. Rüschendorf. The contraction method for recursive algorithms. *Algorithmica*, 29(1-2):3–33, 2001.