

# The stable trees are nested

Nicolas Curien · Bénédicte Haas

Received: 26 July 2012 / Revised: 26 November 2012 / Published online: 16 December 2012  
© Springer-Verlag Berlin Heidelberg 2012

**Abstract** We show that we can construct simultaneously all the stable trees as a nested family. More precisely, if  $1 < \alpha < \alpha' \leq 2$  we prove that hidden inside any  $\alpha$ -stable tree we can find a version of an  $\alpha'$ -stable tree rescaled by an independent Mittag-Leffler type distribution. This tree can be explicitly constructed by a pruning procedure of the underlying stable tree or by a modification of the fragmentation associated with it. Our proofs are based on a recursive construction due to Marchal which is proved to converge almost surely towards a stable tree.

**Keywords** Stable Lévy trees · Pruning · Dissipative self-similar fragmentations · Marchal's algorithm

**Mathematics Subject Classification (2000)** 60J25 · 60J80

## 1 Introduction

Since the beginning of the early 1990's and the introduction of the Brownian continuum random tree by Aldous [2], random trees have been the object of intense research in probability theory. Various classes of continuous random trees have been considered to extend the initial Brownian tree setup, two important such classes are Lévy and fragmentation trees. The Lévy trees have been introduced by Le Gall and Le Jan [20] to describe the genealogical structure of continuous-state branching processes and

---

N. Curien (✉)  
École normale supérieure, Paris, France  
e-mail: nicolas.curien@gmail.com

B. Haas  
Université Paris-Dauphine and École normale supérieure, Paris, France  
e-mail: haas@ceremade.dauphine.fr

furnish all the possible scaling limits of Galton–Watson trees [12]. The fragmentation trees [16] encode the genealogy of self-similar fragmentation processes and are the scaling limits of Markov branching trees [15]. The intersection of these two classes consists of the stable trees of parameter

$$\alpha \in (1, 2].$$

When  $\alpha = 2$ , the 2-stable tree corresponds to Aldous’ Brownian CRT, up to a multiplicative scaling. For general  $\alpha \in (1, 2]$ , the  $\alpha$ -stable trees can either be seen as the scaling limits of conditioned critical Galton–Watson trees with offspring distribution in the domain of attraction of a stable law of index  $\alpha$  [11], or as a  $1/\alpha - 1$  self-similar fragmentation tree which is invariant under uniform re-rooting [18]. The increasing function  $\alpha \mapsto 1 - 1/\alpha$  will repeatedly appear when dealing with stable trees and we shall use the following notation throughout the paper:

$$\bar{\alpha} = 1 - 1/\alpha.$$

We adopt the convention of Duquesne and Le Gall [12] and denote by  $\mathcal{T}_\alpha$  the *standard*  $\alpha$ -stable tree, which is the one describing the genealogical structure of a continuous-state branching process with branching mechanism  $\lambda \mapsto \lambda^\alpha$ . Alternatively, for  $\alpha = 2$ , it can be defined as the scaling limit of a conditioned Galton–Watson tree with offspring distribution  $\eta$  with mean 1 and variance  $\sigma^2 \in (0, \infty)$  in the following sense:

$$\frac{T_n^{\text{GW}}}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{(d)} \frac{\sqrt{2}}{\sigma} \mathcal{T}_2,$$

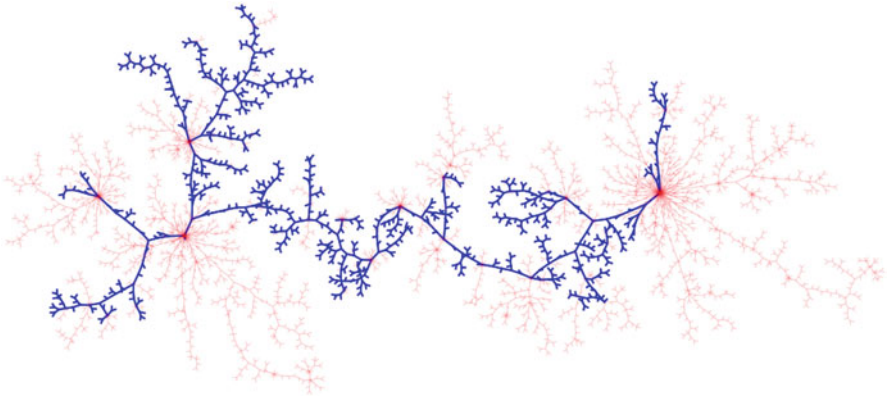
where  $T_n^{\text{GW}}$  is a version of the above mentioned Galton–Watson tree conditioned to have  $n$  vertices. Similarly, for  $\alpha \in (1, 2)$  and for an offspring distribution  $\eta$  with mean 1 and such that  $\eta(k) \sim Ck^{-1-\alpha}$  as  $k \rightarrow \infty$ , we have that:

$$\frac{T_n^{\text{GW}}}{n^{\bar{\alpha}}} \xrightarrow[n \rightarrow \infty]{(d)} \left( \frac{\alpha(\alpha - 1)}{C\Gamma(2 - \alpha)} \right)^{1/\alpha} \mathcal{T}_\alpha.$$

These convergences hold for the Gromov–Hausdorff topology, as recalled in Sect. 2.

**The nested family of stable trees.** Some connections between different stable and Lévy trees have already been observed. In [1] Abraham et al. present a pruning procedure for a large class of Lévy trees that leads to other Lévy trees. However, when applied to stable trees, their procedure gives pruned subtrees that are not stable anymore. In the other direction, Bertoin et al. [9] obtain stable trees by folding randomly the branches of a Brownian tree. These results are a priori unrelated to the ones presented here.

It is well-known [13, 16] that the Hausdorff dimension of the  $\alpha$ -stable tree is almost surely  $1/\bar{\alpha}$ , for all  $\alpha \in (1, 2]$ . Thus, in some sense, the stable trees are decreasing in the parameter  $\alpha$ . The goal of this work is to give a precise geometric statement of the last heuristic and to show that all the stable trees can be seen as a single family



**Fig. 1** A Brownian tree hidden in a stable tree

of (random) nested trees. To do so, rather than considering the standard versions of the stable trees  $\mathcal{T}_\alpha$ , we will consider randomly scaled versions so that it is possible to build on the same probability space a family of nested stable trees. More precisely, we let  $\Gamma_\alpha, \alpha > 0$ , denote a Gamma random variable with shape parameter  $\alpha$  and scale parameter 1, that is with density proportional to  $x^{\alpha-1} e^{-x} \mathbf{1}_{\{x>0\}}$ . We then set

$$J_\alpha \stackrel{(d)}{=} \alpha \cdot (\Gamma_{1+\alpha})^{\bar{\alpha}}.$$

This variable has been designed so that if we rescale a stable tree  $\mathcal{T}_\alpha$  by a independent variable distributed as  $J_\alpha$  (that is we multiply the distances in  $\mathcal{T}_\alpha$  by the factor  $J_\alpha$ ) then the height of a random uniform point in  $J_\alpha \cdot \mathcal{T}_\alpha$  has a distribution that does not depend on  $\alpha$ , namely a Gamma distribution of parameter 2. Our main result is then

**Theorem 1** *There exists a process of rescaled nested stable trees  $(\mathcal{T}_\alpha, 1 < \alpha \leq 2)$  such that:*

- $\mathcal{T}_\alpha \stackrel{(d)}{=} J_\alpha \cdot \mathcal{T}_\alpha$ , where  $J_\alpha$  is independent of  $\mathcal{T}_\alpha, \alpha \in (1, 2]$ ,
- $\mathcal{T}_{\alpha'} \subset \mathcal{T}_\alpha$ , for all  $1 < \alpha \leq \alpha' \leq 2$ .

The existence of this decreasing process of rescaled stable trees is actually a simple corollary of the following proposition, which says that hidden inside any stable tree of parameter  $\alpha$ , there exists a rescaled version of a stable tree of parameter  $\alpha' > \alpha$  (see Fig. 1).

**Proposition 2** *Let  $1 < \alpha < \alpha' \leq 2$ . There exists a closed (random) subtree  $\mathcal{T}_{\alpha,\alpha'}$  of the  $\alpha$ -stable tree  $\mathcal{T}_\alpha$ , such that*

$$\mathcal{T}_{\alpha,\alpha'} \stackrel{(d)}{=} (M_{\alpha,\alpha'})^{\bar{\alpha}'} \cdot \mathcal{T}_{\alpha'},$$

where  $\mathcal{T}_{\alpha'}$  is a standard  $\alpha'$ -stable tree and  $M_{\alpha,\alpha'}$  is an independent variable distributed as  $(\alpha'/\alpha)^{1/\bar{\alpha}'}$  times a generalized Mittag-Leffler distribution with parameters  $(\bar{\alpha}/\bar{\alpha}', \bar{\alpha})$ .

The definition of generalized Mittag-Leffler distributions is recalled in Sect. 3.1. Throughout this work, the boundary case  $\alpha' = 2$ , where a (rescaled) Brownian tree is extracted from an  $\alpha$ -stable tree, deserves a special attention since our constructions and proofs simplify substantially in this case. Although not unique, the subtree  $\mathcal{T}_{\alpha,\alpha'}$  of the last proposition can be explicitly constructed by a pruning procedure of the tree  $\mathcal{T}_\alpha$ .

**Pruning procedure.** Let  $1 < \alpha < \alpha' \leq 2$ . Conditionally on the stable tree  $\mathcal{T}_\alpha$ , let  $X_i, i \geq 0$  be i.i.d. random leaves sampled according to its uniform mass measure. Then write  $t_n$  for the subtree spanned by  $X_0, X_1, \dots, X_n$  in  $\mathcal{T}_\alpha$ . For  $d, d' \in \{0, 2, 3, 4, 5, \dots\}$  with  $d \geq d'$  and  $d \geq 2$ , we introduce the following probabilities

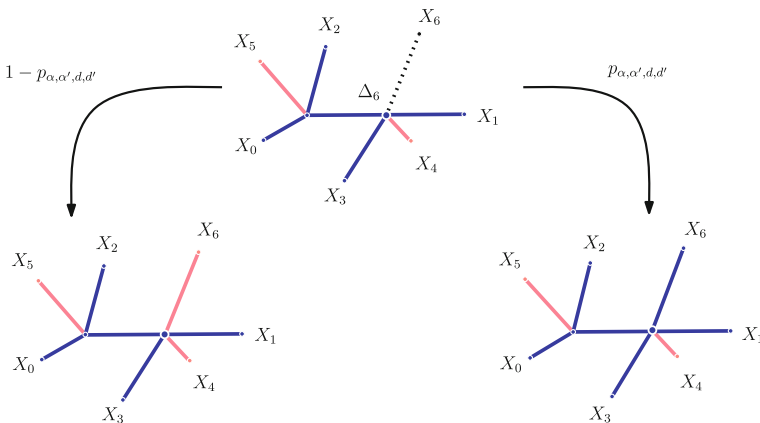
$$p_{\alpha,\alpha',d,d'} = \begin{cases} 0 & \text{if } d' = 0, \\ 1 & \text{if } d' = 2, \\ \frac{(d'-1-\alpha')(\alpha-1)}{(d-1-\alpha)(\alpha'-1)} & \text{otherwise,} \end{cases} \in [0, 1], \tag{1}$$

from which we construct inductively a sequence of subtrees  $\tau_n \subset t_n \subset \mathcal{T}_\alpha$  as follows:

- $\tau_1 = t_1 = [[X_0, X_1]]$  is the line segment linking  $X_0$  to  $X_1$  in  $\mathcal{T}_\alpha$ ,
- at step  $i \geq 2$ , write  $t_i = t_{i-1} \cup [[X_i, \Delta_i]]$ , where  $\Delta_i \in t_{i-1}$  and  $[[X_i, \Delta_i]]$  is the shortest path in  $t_i$  connecting  $X_i$  to  $t_{i-1}$ . Then let  $d_i$  denote the degree (multiplicity) of  $\Delta_i$  in  $t_{i-1}$  and  $d'_i$  its degree in  $\tau_{i-1}$ , with the convention that  $d'_i = 0$  if  $\Delta_i \notin \tau_{i-1}$ . Finally set (see Fig. 2)

$$\begin{aligned} \tau_i &= \tau_{i-1} \cup [[X_i, \Delta_i]] \text{ with probability } p_{\alpha,\alpha',d_i,d'_i}, \\ \tau_i &= \tau_{i-1} \text{ otherwise.} \end{aligned}$$

The sequence  $(\tau_n)$  is clearly increasing in  $\mathcal{T}_\alpha$  and we denote by  $\text{Prun}_{\alpha,\alpha'}(\mathcal{T}_\alpha; (X_i)_{i \geq 0})$  the closure of this increasing union.



**Fig. 2** Illustration of the pruning construction. Here  $d' = 3, d = 4$  and the trees  $\tau_5$  and  $\tau_6$  are in blue (color figure online)

Notice that when  $\alpha' = 2$  we have  $p_{\alpha,2,d,3} = 0$  and thus no branch point of degree larger or equal to 4 is created in  $\tau_n$ . In other words,  $(\tau_n)$  is a sequence a binary trees. In this case, the pruning procedure boils down to adding those points  $X_{n+1}$  whose attachment to  $\tau_n$  preserves the binary structure of the tree and thus  $\text{Prun}_{\alpha,2}(\mathcal{T}_\alpha; (X_i)_{i \geq 0})$  can be seen as a deterministic function of  $\mathcal{T}_\alpha$  and of the leaves  $X_i, i \geq 0$ .

**Theorem 3** *The statement of Proposition 2 holds with*

$$\mathfrak{T}_{\alpha,\alpha'} = \text{Prun}_{\alpha,\alpha'}(\mathcal{T}_\alpha; (X_i)_{i \geq 0}).$$

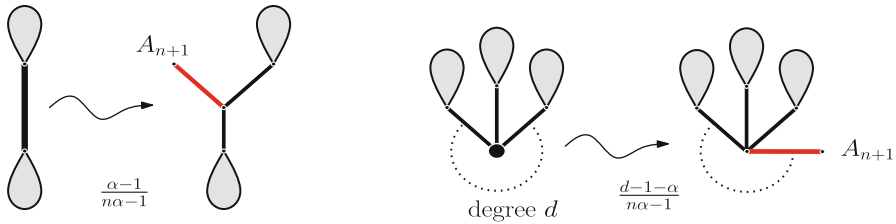
Moreover, since  $\alpha' \mapsto p_{\alpha,\alpha',d,d'}$  is decreasing, we can couple the realizations of the pruned subtrees such that  $\alpha' \in (\alpha, 2] \mapsto \text{Prun}_{\alpha,\alpha'}(\mathcal{T}_\alpha; (X_i)_{i \geq 0})$  is decreasing inside  $\mathcal{T}_\alpha$ . The pruning operation can also be viewed from a fragmentation point of view which sheds new light on the construction of  $\mathfrak{T}_{\alpha,\alpha'}$ .

**Modification of the stable fragmentation.** Fragmentation theory is a mathematical attempt to model the behavior of particles that undergo a splitting process, see [7] for an account on this field. Any  $\alpha$ -stable tree  $\mathcal{T}_\alpha$  can be associated with a self-similar fragmentation process that describes the masses of the subtrees of  $\mathcal{T}_\alpha$  above a certain height. In other words, an  $\alpha$ -stable tree can be seen as the genealogical tree (in the sense of [16]) of a pure-jump self-similar fragmentation process with index  $-\bar{\alpha}$ , whose dislocation measure  $\nu_\alpha$  has been identified by Bertoin [6] in the Brownian case  $\alpha = 2$  and by Miermont [22] in the general case (see Sect. 5 for an expression of  $\nu_\alpha$ ). All these dislocation measures are conservative, which means that the mass is kept at each dislocation. In Sect. 5 we introduce for  $1 < \alpha < \alpha' \leq 2$  a modification of the dislocation measure  $\nu_\alpha$  by (roughly speaking) keeping randomly some of the fragments created by  $\nu_\alpha$ . When  $\alpha' = 2$ , this simply consists in keeping only two fragments, chosen proportionally to their sizes. In the other cases the procedure is more complex and depends on the probabilities  $p_{\alpha,\alpha',d,d'}$  introduced in (1). Obviously, the resulting dislocation measure  $\nu_{\alpha,\alpha'}$  is now *dissipative* (some mass is lost at each dislocation). It is still possible to associate with it a random tree coding its genealogy (see [26]) and we have:

**Proposition 4** *The tree  $\mathfrak{T}_{\alpha,\alpha'}$  of Proposition 2 can be seen as the genealogical tree of a pure-jump self-similar fragmentation of index  $-\bar{\alpha}$  and dislocation measure  $\nu_{\alpha,\alpha'}$ .*

Thus, in addition to being the genealogy of a canonical conservative fragmentation process, the stable tree of parameter  $\alpha'$  is also, up to a random scaling, the genealogical tree of a dissipative fragmentation process with dislocation measure  $\nu_{\alpha,\alpha'}$  and auto-similarity index  $-\bar{\alpha}$  for all  $\alpha \in (1, \alpha')$ . Besides, as such a dissipative fragmentation tree,  $\mathfrak{T}_{\alpha,\alpha'}$  naturally carries a Malthusian measure whose total mass is distributed, up to a deterministic scaling, as the random variable  $M_{\alpha,\alpha'}$  appearing in the random scaling in Proposition 2. In Theorem 15 of Sect. 5, we reinforce Proposition 4 by showing that this dissipative fragmentation tree endowed with its Malthusian measure is distributed as an  $\alpha'$ -stable tree with its natural uniform mass measure, up to a random scaling depending on  $M_{\alpha,\alpha'}$ .

**Strategy and organization of the paper.** The construction of the subtree  $\mathfrak{T}_{\alpha,\alpha'}$  presented in Proposition 2 relies on a discrete approach. In [21], Marchal intro-



**Fig. 3** Illustration of the recursive rules to construct  $\mathbf{T}_\alpha(n + 1)$  from  $\mathbf{T}_\alpha(n)$

duced a Markov chain  $(\mathbf{T}_\alpha(n), n \geq 1)$  made of increasing labeled trees which, once re-normalized, converge towards the stable tree of parameter  $\alpha$ . Let us present quickly this construction: we start with  $\mathbf{T}_\alpha(1)$ , the tree with a single edge, then inductively at step  $n \geq 2$  we associate a weight  $\alpha - 1$  to every edges of the tree  $\mathbf{T}_\alpha(n)$  and a weight  $d - 1 - \alpha$  to each vertex with degree  $d \geq 3$ . An edge or a vertex of  $\mathbf{T}_\alpha(n)$  is chosen at random accordingly to these weights and a new edge is attached either directly to the vertex or in the “middle” of the chosen edge, see Fig. 3 and Sect. 2.3 for more details. In the case  $\alpha = 2$  this construction is the famous algorithm for growing binary trees due to Rémy [25]. We prove in Theorem 5 that, once re-normalized, the trees  $\mathbf{T}_\alpha(n)$  converge towards a stable tree

$$\frac{\mathbf{T}_\alpha(n)}{n^{\bar{\alpha}}} \xrightarrow[n \rightarrow \infty]{\text{a.s.}} \alpha \mathcal{T}_\alpha,$$

in the Gromov–Hausdorff sense. This result thus strengthens the convergence in probability already obtained in [17, Corollary 24].

The key observation that triggered this work is that it is possible to identify within Marchal’s construction of the  $\mathbf{T}_\alpha(n)$ ’s a “sub-Markov chain” of trees  $\mathbf{T}_{\alpha,\alpha'}(n) \subset \mathbf{T}_\alpha(n)$  whose growth mechanism is identical to Marchal’s algorithm but with parameter  $\alpha' > \alpha$ , see Sect. 3. The scaling limit of the sequence  $(\mathbf{T}_{\alpha,\alpha'}(n), n \geq 1)$  then furnishes the tree  $\mathfrak{T}_{\alpha,\alpha'}$  of Proposition 2.

The paper is organized as follows. Section 2 contains the background on discrete and continuous trees as well as the presentation of Marchal’s construction and its almost sure convergence towards the stable tree (Theorem 5). We then move in Sect. 3 to the observation that sub-constructions lie inside Marchal’s algorithm and deduce Proposition 2 and Theorem 1. Section 4 is devoted to the pruning construction of  $\mathfrak{T}_{\alpha,\alpha'}$  (Theorem 3). Finally, the last section explores the fragmentation approach and its consequences.

## 2 Background on stable trees

In this section, we present the recursive construction of Marchal and prove (Theorem 5) that it converges *almost surely*, after rescaling, towards a stable tree. Before embarking into that topic, we start by introducing some background on trees and the Gromov–Hausdorff and Gromov–Hausdorff–Prokhorov distances.

### 2.1 Discrete and continuous trees

**Discrete trees.** A discrete tree  $\tau$  is a finite connected graph without cycle, considered up to graph isomorphisms: it is not embedded in any space and its vertices are unlabeled. If  $x$  and  $y$  are two vertices of a tree  $\tau$ , we denote by  $[[x, y]]$  the discrete geodesic in  $\tau$  between  $x$  and  $y$ . If  $y_1, y_2, \dots, y_k$  are (distinct) vertices of  $\tau$  we write

$$\text{Span}(\tau; y_1, \dots, y_k) = \bigcup_{1 \leq i, j \leq k} [[y_i, y_j]],$$

for the discrete tree spanned by these vertices. The degree of a vertex is the number of edges adjacent to it, for instance a leaf is a vertex of degree one. A tree is binary if the degrees of its vertices are in  $\{1, 2, 3\}$ .

A labeled discrete tree  $\tau = (\tau; x_0, x_1, \dots, x_n)$  is a pair formed by a discrete tree  $\tau$  given with an exhaustive enumeration of its *leaves*. If  $\tau$  is a labeled tree we call the shape of  $\tau$  the tree  $\tau$  obtained after forgetting the labeling of the tree. We will systematically use bold letters for labeled trees and standard ones for their associated shapes.

**Continuous trees.** We briefly recall here some facts on  $\mathbb{R}$ -trees and refer to [14, 19] for an overview of this topic. An  $\mathbb{R}$ -tree is a metric space  $(\mathcal{T}, d)$  such that for every  $x, y \in \mathcal{T}$ ,

- there is an isometry  $\varphi_{x,y} : [0, d(x, y)] \rightarrow \mathcal{T}$  such that  $\varphi_{x,y}(0) = x$  and  $\varphi_{x,y}(d(x, y)) = y$
- for every continuous, injective function  $c : [0, 1] \rightarrow \mathcal{T}$  with  $c(0) = x, c(1) = y$ , one has  $c([0, 1]) = \varphi_{x,y}([0, d(x, y)])$ .

We identify two  $\mathbb{R}$ -trees when they are isometric, and still use the notation  $(\mathcal{T}, d)$  to represent an isometry class. Note that a discrete tree may be seen as a  $\mathbb{R}$ -tree by “replacing” its edges by segments. Unless specified, it is implicit in this paper that these line segments are all of length 1. We use much of the notation we introduced in the context of discrete trees when it is non ambiguously extended to  $\mathbb{R}$ -trees. In particular if  $(\mathcal{T}, d)$  is an  $\mathbb{R}$ -tree and if  $a, b \in \mathcal{T}$  we denote by  $[[a, b]]$  the geodesic line segment between  $a$  and  $b$  in  $\mathcal{T}$ . Also  $\text{Span}(\mathcal{T}; y_1, \dots, y_k) = \cup_{1 \leq i, j \leq k} [[y_i, y_j]]$  still denotes the subtree spanned by  $y_1, \dots, y_k \in \mathcal{T}$ . The degree (or multiplicity) of a point  $x \in \mathcal{T}$  is the number of connected components of  $\mathcal{T} \setminus \{x\}$ . A leaf is a point of degree 1, a branch point has degree larger than or equal to 3, and an  $\mathbb{R}$ -tree is said to be binary if the degrees of its points are in  $\{1, 2, 3\}$ .

**Gromov–Hausdorff–Prokhorov topology.** The reader interested by the Gromov–Hausdorff and Gromov–Hausdorff–Prokhorov topologies should consult [10, 14, 23] for details and proofs. Let  $k \in \{0, 1, 2, \dots\}$ . A compact metric space  $(E, d)$  is  $k$ -pointed if it is given with  $k$  points  $x_1, \dots, x_k \in E$  (when  $k = 0$  this is just a compact metric space). An isometry between two  $k$ -pointed compact metric spaces is an isometry between the spaces that maps the  $k$  distinguished points to each other. The set of isometry classes of  $k$ -pointed compact metric spaces is endowed with the classical ( $k$ -pointed) Gromov–Hausdorff topology, that makes it Polish. This distance can be

defined as

$$d_{GH}((E, d; x_1, \dots, x_k), (E', d'; x'_1, \dots, x'_k)) = \inf \left\{ d_H(\phi(E), \phi(E')) \vee \max_{1 \leq i \leq k} \delta(\phi(x_i), \phi'(x'_i)) \right\},$$

where the infimum is taken over all choices of metric spaces  $(F, \delta)$  and isometric embeddings  $\phi : E \rightarrow F, \phi' : E' \rightarrow F$ , and where  $d_H$  denotes the Hausdorff distance in  $F$ . The class of compact  $\mathbb{R}$ -trees forms the accumulation points of the class of rescaled discrete trees for the Gromov–Hausdorff topology and is thus closed. In the remainder of the paper, we use the notation  $a \cdot E$  for the rescaled metric space  $(E, ad)$  for any  $a > 0$ .

The Gromov–Hausdorff topology can be enriched in order to take into account measured spaces. A compact metric space  $(E, d)$  endowed with a Borel probability measure  $\mu$  is called a measured compact metric space. We can extend the Gromov–Hausdorff topology to isometry classes (defined in the obvious way) of measured compact metric spaces by putting

$$d_{GHP}((E, d, \mu), (E', d', \mu')) = \inf \left\{ d_H(\phi(E), \phi'(E')) \vee d_P(\phi_*\mu, \phi'_*\mu') \right\},$$

where again  $\phi, \phi'$  are isometries from  $E, E'$  into a common space  $(F, \delta)$ ,  $\phi_*\mu, \phi'_*\mu'$  are the push-forward of  $\mu, \mu'$  by  $\phi, \phi'$ , and  $d_P$  is the Prokhorov distance between Borel probability measures on  $F$ :

$$d_P(m, m') = \inf \{ \varepsilon > 0 : m(C) \leq m'(C^\varepsilon) + \varepsilon \text{ for every } C \subset F \text{ closed} \},$$

where  $C^\varepsilon = \{x \in F : \inf_{y \in C} d(x, y) < \varepsilon\}$  is the  $\varepsilon$ -neighborhood of  $C$ . The function  $d_{GHP}$  is a distance on the set of isometry classes of measured compact metric spaces that makes it Polish.

Let us give another convenient way to express Gromov–Hausdorff distances. A correspondence between two  $k$ -pointed compact metric spaces  $(E, d; x_1, \dots, x_k)$  and  $(E', d'; x'_1, \dots, x'_k)$  is a subset  $\mathcal{R}$  of  $E \times E'$  such that, for every  $y_1 \in E$ , there exists at least one point  $y_2 \in E'$  such that  $(y_1, y_2) \in \mathcal{R}$  and conversely, for every  $z_2 \in E'$ , there exists at least one point  $z_1 \in E$  such that  $(z_1, z_2) \in \mathcal{R}$ . Also  $(x_1, x'_1), \dots, (x_k, x'_k) \in \mathcal{R}$ . The distortion of the correspondence  $\mathcal{R}$  is defined by

$$\text{dis}(\mathcal{R}) = \sup \{ |d(x_1, y_1) - d'(x_2, y_2)| : (x_1, x_2), (y_1, y_2) \in \mathcal{R} \}.$$

The Gromov–Hausdorff distance between  $(E, d; x_1, \dots, x_k)$  and  $(E', d'; x'_1, \dots, x'_k)$  can be expressed as half the infimum of the distortions of correspondences between  $E$  and  $E'$ . A similar (but more involved) definition of  $d_{GHP}$  via correspondences can be found in [23].



### 2.2 The stable trees

We now give some additional background on stable trees and refer to [12, 13] for a complete account. Stable trees are random variables  $\mathcal{T}_\alpha$  for  $\alpha \in (1, 2]$  taking values in the set of isometry classes of compact  $\mathbb{R}$ -trees that can be defined in various manners. For example, the tree  $\mathcal{T}_\alpha$  can be defined by its finite-dimensional marginals (see [12, Theorem 3.3.3]) or by its contour process. In particular, the 2-stable tree can be identified with  $\sqrt{2}$  times the Brownian tree  $\mathcal{T}_{Br}$  coded by a standard normalized excursion  $(e_t : 0 \leq t \leq 1)$ . Note also that the Brownian tree introduced by Aldous in [3] is equal to twice  $\mathcal{T}_{GHP}$  so that we have in distribution

$$\mathcal{T}_2 = \sqrt{2}\mathcal{T}_{GHP} = \frac{\mathcal{T}_{Aldous}}{\sqrt{2}}.$$

We mention that contrary to the Brownian tree where all the branch points are of degree 3, in the stable case  $1 < \alpha < 2$ , they are of infinite multiplicity ([13]).

Let us make precise the definition of  $\mathcal{T}_\alpha$  via scaling limits of Galton–Watson trees and introduce the mass measure on  $\mathcal{T}_\alpha$ . For  $\alpha \in (1, 2)$ , consider an (unordered) Galton–Watson tree with offspring  $\eta$  with mean 1 and such that  $\eta(k) \sim Ck^{-1-\alpha}$  when  $k \rightarrow \infty$ . Let  $T_n^{GW}$  denote a version of this tree conditioned to have  $n$  vertices, and equip it with the uniform probability measure on its vertices, which is denoted by  $\mu_n$ . Then by Duquesne [11] (recall that  $\bar{\alpha} = 1 - 1/\alpha$ ),

$$\left( \frac{T_n^{GW}}{n^{\bar{\alpha}}}, \mu_n \right) \xrightarrow[n \rightarrow \infty]{(d)} \left( \left( \frac{\alpha(\alpha - 1)}{C\Gamma(2 - \alpha)} \right)^{1/\alpha} \mathcal{T}_\alpha, \mu_\alpha \right),$$

for the Gromov–Hausdorff–Prokhorov topology, where  $\mathcal{T}_\alpha$  is a stable tree of index  $\alpha$  equipped with a probability measure  $\mu_\alpha$  which can be interpreted as the uniform mass measure on  $\mathcal{T}_\alpha$ . A similar result holds for  $\alpha = 2$ . It is well-known [12] that for all  $\alpha \in (1, 2]$ , the measure  $\mu_\alpha$  is actually fully supported by the set of leaves of  $\mathcal{T}_\alpha$  and that it can be measurably constructed from  $\mathcal{T}_\alpha$ . It thus makes sense to speak about an i.i.d. sample  $X_n, n \geq 0$  of leaves according to  $\mu_\alpha$  conditionally on  $\mathcal{T}_\alpha$ . We shall repeatedly use the following easy fact:  $\{X_n, n \geq 0\}$  is dense in  $\mathcal{T}_\alpha$  a.s.

### 2.3 Marchal’s recursive construction

Let  $\alpha \in (1, 2]$ . In [21], Marchal proposed a recursive construction to build random finite trees that converge in the scaling limit towards the  $\alpha$ -stable tree. The construction is in fact a Markov chain  $(\mathbf{T}_\alpha(n))_{n \geq 1}$  with values in the set of labeled trees, such that  $\mathbf{T}_\alpha(n)$  has  $n + 1$  leaves and is defined as follows: we start with the tree  $\mathbf{T}_\alpha(1)$  which is the only tree with one edge and two labeled vertices denoted by  $A_0$  and  $A_1$ . We then build recursively  $\mathbf{T}_\alpha(n + 1)$  from  $\mathbf{T}_\alpha(n)$  by adding a new edge. More precisely, given  $\mathbf{T}_\alpha(n)$ , assign a weight  $\alpha - 1$  to any edge of  $\mathbf{T}_\alpha(n)$  and a weight  $d - 1 - \alpha$  to each vertex of degree  $d \geq 3$  (all other vertices have zero weight). The total weight

$W(\mathbf{T}_\alpha(n))$  of the tree  $\mathbf{T}_\alpha(n)$  is easily checked to be

$$W(\mathbf{T}_\alpha(n)) = n\alpha - 1, \tag{2}$$

and is in particular independent of the shape of the tree. We then choose an edge or a vertex of  $\mathbf{T}_\alpha(n)$  proportionally to its weight:

- if we picked an edge then we split it into two edges with a middle vertex on which we attach a new edge carrying the  $n + 1$ th leaf denoted by  $A_{n+1}$ .
- if a vertex has been selected then we attach a new edge carrying  $A_{n+1}$  to it.

In [21], Marchal exactly computed the distribution of the tree  $\mathbf{T}_\alpha(n)$ . More precisely, he proved by induction on  $n$  that if  $\mathbf{t}_0$  is a fixed labeled tree with  $n + 1$  leaves, then

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

with  $p_1 = 1$ ,  $p_2 = 0$  and  $p_k = |(\alpha - 1)(\alpha - 2) \dots (\alpha - k + 2)|$  for  $k \geq 3$ .

From this formula, it is easy to deduce (see [21]) that  $\mathbf{T}_\alpha(n)$  has the same distribution as the labeled tree obtained from the finite  $n + 1$ -dimensional marginal of the standard  $\alpha$ -stable tree, see Theorem 3.3.3 in [12]. With this identification in hand, it is shown in [17, Corollary 24] that there exists a stable tree  $\mathcal{T}_\alpha$  built on the same probability space that supports  $\mathbf{T}_\alpha(n)$ ,  $n \geq 1$ , and equipped with its uniform mass measure  $\mu_\alpha$ , such that

$$\left( \frac{\mathbf{T}_\alpha(n)}{\alpha n^\alpha}, \frac{\sum_{i=0}^n \delta_{A_i}}{n + 1} \right) \xrightarrow[n \rightarrow \infty]{(\mathbb{P})} (\mathcal{T}_\alpha, \mu_\alpha)$$

for the Gromov–Hausdorff–Prokhorov topology. The goal of the next section is to improve this convergence into an almost-sure one.

### 2.4 The convergence theorem

Let  $\mathcal{T}_\alpha$  be an  $\alpha$ -stable tree and conditionally on  $\mathcal{T}_\alpha$ , let  $(X_i, i \geq 0)$  be a sample of i.i.d. leaves distributed according to the mass distribution  $\mu_\alpha$  on  $\mathcal{T}_\alpha$ . We consider the finite  $n + 1$ th-dimensional trees  $\text{Span}(\mathcal{T}_\alpha; X_0, X_1, \dots, X_n)$  for  $n \geq 1$ . These objects are by definition continuous trees but can equivalently be seen as discrete labeled trees

$$\boldsymbol{\tau}_\alpha(n) = (\tau_\alpha(n); x_0, x_1, \dots, x_n)$$

carrying positive lengths  $(\ell_e)$  on their edges, see Fig. 4.

**Theorem 5** (i) *We have the joint identity in distribution*

$$(\mathbf{T}_\alpha(n), n \geq 1) \stackrel{(d)}{=} (\boldsymbol{\tau}_\alpha(n), n \geq 1).$$

(ii) For every  $k \geq 0$ , we have in the sense of  $k$ -pointed Gromov–Hausdorff distance

$$\left( \frac{\tau_\alpha(n)}{\alpha n^{\bar{\alpha}}}; x_0, x_1, \dots, x_k \right) \xrightarrow[n \rightarrow \infty]{a.s.} (\mathcal{T}_\alpha; X_0, X_1, \dots, X_k). \tag{3}$$

(iii) If  $\mu_n = \frac{1}{n+1} \sum_{i=0}^n \delta_{x_i}$  denotes the uniform mass measure on the leaves of  $\tau_\alpha(n)$  then we have the following almost sure convergence for the Gromov–Hausdorff–Prokhorov topology

$$\left( \frac{\tau_\alpha(n)}{\alpha n^{\bar{\alpha}}}, \mu_n \right) \xrightarrow[n \rightarrow \infty]{a.s.} (\mathcal{T}_\alpha, \mu_\alpha).$$

Consequently,  $\mathcal{T}_\alpha(n)/\alpha n^{\bar{\alpha}}$  endowed with the uniform mass measure on its leaves indeed converges almost surely for the Gromov–Hausdorff–Prokhorov topology towards an  $\alpha$ -stable tree endowed with its uniform mass measure.

*Proof* To simplify the notation during the proof we fix  $\alpha \in (1, 2]$  once for all and drop the indices  $\alpha$  by writing  $\tau_n = (\tau_n; x_0, \dots, x_n)$  instead of  $\tau_\alpha(n) = (\tau_\alpha(n); x_0, \dots, x_n)$ . To get (i), we use the fact discussed at the end of the previous section, and proved by Marchal in [21], that  $\mathbf{T}_\alpha(n) = \tau_n$  in distribution in terms of labeled trees, for all  $n \geq 1$ . Since for  $k \leq n$  the labeled trees  $\tau_k$  and  $\mathbf{T}_\alpha(k)$  are deduced from  $\tau_n$  resp.  $\mathbf{T}_\alpha(n)$  by the same deterministic procedure, we deduce that  $(\mathbf{T}_\alpha(k))_{1 \leq k \leq n}$  has the same law as  $(\tau_k)_{1 \leq k \leq n}$ . Since this holds for all  $n \geq 1$  we indeed get (i).

Let us turn to the second point and remark first that for every  $k \geq 1$  we have the almost sure convergence in the sense of pointed Gromov–Hausdorff metric

$$\left( \text{Span}(\mathcal{T}_\alpha; (X_i)_{0 \leq i \leq n}); X_0, X_1, \dots, X_k \right) \xrightarrow[n \rightarrow \infty]{a.s.} \left( \mathcal{T}_\alpha; X_0, X_1, \dots, X_k \right). \tag{4}$$

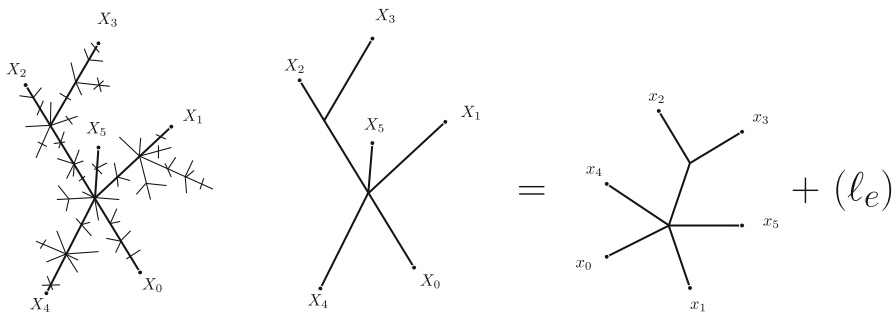
This is easy to obtain and follows from the fact that the sequence of leaves  $(X_i)_{i \geq 0}$  is almost surely dense in the compact tree  $\mathcal{T}_\alpha$ . The convergence (3) will be established by comparing the labeled tree  $\tau_n$  to its continuous analogue  $\text{Span}(\mathcal{T}_\alpha; (X_i)_{0 \leq i \leq n})$  obtained by dressing it up with edge lengths. More precisely, we recall that the continuous tree  $\text{Span}(\mathcal{T}_\alpha; (X_i)_{0 \leq i \leq n})$  can be seen as the discrete labeled tree  $\tau_n$  where each edge  $e$  has been replaced by a Euclidean segment of length  $\ell_e^{(n)}$ , see Fig. 4 below.

For notation convenience we write

- $d_{\text{gr}}^{(n)}$  for the graph metric in  $\tau_n$  (with edge lengths 1)
- $d_\ell^{(n)}$  for the metric on the vertices of  $\tau_n$  associated to the edge-lengths  $(\ell_e^{(n)})$ .

In particular the two metrics live on the set of vertices of  $\tau_n$ . Also, for  $0 \leq i, j \leq n$  the  $d_\ell^{(n)}$  distance between  $x_i$  and  $x_j$  in  $\tau_n$  is equal to the distance between  $X_i$  and  $X_j$  in  $\mathcal{T}_\alpha$ . Our goal is to show that

$$\sup_{a, b \in \tau_n} \left| \frac{d_{\text{gr}}^{(n)}(a, b)}{n^{\bar{\alpha}}} - \alpha d_\ell^{(n)}(a, b) \right| \xrightarrow[n \rightarrow \infty]{a.s.} 0. \tag{5}$$



**Fig. 4** The trees  $\mathcal{T}_\alpha$ ,  $\text{Span}(\mathcal{T}_\alpha; X_0, \dots, X_5)$  and  $\tau_\alpha(5)$  with its edge-lengths

Let us first show how to use the last display to complete the proof of the theorem. For that we use the definition of Gromov–Hausdorff distance via correspondences. Let  $\delta$  denote the metric on  $\mathcal{T}_\alpha$  and define a correspondence  $\mathcal{R}_n$  between the trees  $(\tau_n, n^{-\tilde{\alpha}} d_{\text{gr}}^{(n)})$  and  $(\text{Span}(\mathcal{T}_\alpha; X_0, \dots, X_n), \alpha\delta)$  by declaring that  $a \in \tau_n$  and  $x \in \text{Span}(\mathcal{T}_\alpha; X_0, \dots, X_n)$  are in correspondence if  $x$  belongs to an “edge” of the continuous tree that is combinatorially adjacent to the associate vertex  $a$  in  $\tau_n$ . Note that  $x_i$  is in correspondence with  $X_i$  for all  $0 \leq i \leq n$ . The distortion of this correspondence is bounded by

$$\text{dist}(\mathcal{R}_n) \leq \sup_{a,b \in \tau_n} \left| \frac{d_{\text{gr}}^{(n)}(a,b)}{n^{\tilde{\alpha}}} - \alpha d_\ell^{(n)}(a,b) \right| + 2\alpha \sup \ell_e^{(n)}.$$

By (4) and the fact that branch-points are dense in  $\mathcal{T}_\alpha$  we deduce that  $\sup_e \ell_e^{(n)} \rightarrow 0$  as  $n \rightarrow \infty$ . This fact and (5) thus show that the distortion of  $\mathcal{R}_n$  vanishes as  $n \rightarrow \infty$ . We finally use (4) to get (3).

It thus remains to prove (5). Fix  $\varepsilon > 0$ . For any  $n \geq 0$ , we denote by  $\mathcal{L}_n = \sum \ell_e^{(n)}$  the total length of the tree  $\text{Span}(\mathcal{T}_\alpha; X_0, \dots, X_n)$  and by  $|\tau_n|$  the number of edges of the discrete tree  $\tau_n$ . In order to simplify notation we set

$$\chi_n = \frac{|\tau_n|}{\mathcal{L}_n}.$$

It follows from [12, Theorem 3.3.3] that conditionally on  $\tau_n$  and on  $\mathcal{L}_n$ , the edge lengths  $\ell_e^{(n)}$  are distributed uniformly on  $\mathbb{R}_+^{|\tau_n|}$  subject to the condition  $\sum \ell_e^{(n)} = \mathcal{L}_n$ . In particular, conditionally on  $\tau_n$  and on  $\mathcal{L}_n$ , if  $a$  and  $b$  are two vertices of  $\tau_n$  then the distribution of the random variable  $d_\ell^{(n)}(a,b)$  is given by first dividing uniformly the interval  $[0, \mathcal{L}_n]$  into  $|\tau_n|$  pieces (by throwing  $|\tau_n| - 1$  i.i.d. uniform variables over  $[0, \mathcal{L}_n]$ ) and summing the first  $d_{\text{gr}}^{(n)}(a,b)$ -th ones. We can then apply the following lemma to deduce

$$\mathbb{P} \left( \left| d_\ell^{(n)}(a,b) - \frac{d_{\text{gr}}^{(n)}(a,b)}{\chi_n} \right| \geq \varepsilon d_\ell^{(n)}(a,b) \mid \tau_n, \mathcal{L}_n \right) \leq c_\varepsilon^{-1} \exp \left( -c_\varepsilon d_{\text{gr}}^{(n)}(a,b) \right), \quad (6)$$

$$\mathbb{P}\left(d_\ell^{(n)}(a, b) \geq \varepsilon \mid \tau_n, \mathcal{L}_n\right) \leq c_\varepsilon^{-1} \exp\left(-c_\varepsilon \frac{\chi_n}{d_{\text{gr}}^{(n)}(a, b)}\right). \tag{7}$$

**Lemma 6** *Let  $\varepsilon > 0$ . There exists a constant  $c_\varepsilon > 0$  such that for all  $n \geq 1$  and for all  $L > 0$ , if  $0 = U_0 < U_1 < \dots < U_{n-1} < U_n = L$  is a uniform splitting of the interval  $[0, L]$  then for any  $0 \leq k \leq n$  we have*

$$\begin{aligned} \mathbb{P}\left(\left|U_k - \frac{kL}{n}\right| \geq \varepsilon U_k\right) &\leq c_\varepsilon^{-1} \exp(-c_\varepsilon k), \\ \mathbb{P}(U_k \geq \varepsilon) &\leq c_\varepsilon^{-1} \exp\left(-c_\varepsilon \frac{n}{Lk}\right). \end{aligned}$$

*Proof* By standard properties of the uniform splitting of the unit interval we have  $L^{-1}U_k = T_k/T_n$  in distribution where the  $T_i$ 's are the arrival times of a standard Poisson process. In particular for all  $\varepsilon > 0$ , there exists  $c > 0$  such that

$$\mathbb{P}(|T_i - i| \geq \varepsilon i) \leq \exp(-ci)$$

for all  $i \geq 0$  which easily implies the first inequality of the lemma. For the second inequality, one has, for all  $k \geq 1$ ,

$$\begin{aligned} \mathbb{P}(U_k \geq \varepsilon) &= \mathbb{P}\left(T_k \geq \frac{\varepsilon T_n}{L}\right) \\ &\leq \mathbb{P}\left(T_k \geq \frac{\varepsilon(1-\varepsilon)n}{L}\right) + \mathbb{P}(T_n < (1-\varepsilon)n) \\ &\leq \mathbb{P}\left(\mathcal{P}(\varepsilon(1-\varepsilon)nL^{-1}) \leq k\right) + \exp(-cn), \end{aligned}$$

where  $\mathcal{P}(a)$  denotes a Poisson random variable of parameter  $a > 0$ . Next, since the function  $a \in (0, \infty) \mapsto \mathcal{P}(a)$  is stochastically increasing, and since, for  $q \in \mathbb{N}$ ,  $\mathcal{P}(qk)$  is distributed as the sum of  $q$  independent Poisson random variables of parameter  $k$ , one has,

$$\begin{aligned} \mathbb{P}\left(\mathcal{P}(\varepsilon(1-\varepsilon)nL^{-1}) \leq k\right) &\leq \mathbb{P}\left(\mathcal{P}\left(\lfloor \varepsilon(1-\varepsilon)nL^{-1}k^{-1} \rfloor k\right) \leq k\right) \\ &\leq \mathbb{P}(\mathcal{P}(k) \leq k)^{\lfloor \frac{\varepsilon(1-\varepsilon)n}{Lk} \rfloor}. \end{aligned}$$

Putting the pieces together and using  $\mathbb{P}(\mathcal{P}(k) \leq k) \rightarrow 1/2$  as  $k \rightarrow \infty$  we get the second inequality. □

**Lemma 7** *We have the almost sure convergence*

$$\frac{\chi_n}{n^\alpha} \xrightarrow[n \rightarrow \infty]{a.s.} \alpha.$$

*Proof* It is possible to prove the last lemma by analysing separately the processes  $(|\tau_n|)_{n \geq 1}$  and  $(\mathcal{L}_n)_{n \geq 1}$  and show that  $|\tau_n| \sim \alpha n$  and  $\mathcal{L}_n \sim n^{1/\alpha}$  a.s. as  $n \rightarrow \infty$ . However we bypass these calculations by using a result of [17, Sect. 5.4]: we have

$$\frac{d_{\text{gr}}^{(n)}(x_0, x_1)}{n^{\bar{\alpha}}} \xrightarrow[n \rightarrow \infty]{\text{a.s.}} \alpha \delta(X_0, X_1). \tag{8}$$

Recall that for any  $i, j \geq 0$  and any  $n \geq \max(i, j)$  we have  $d_\ell^{(n)}(x_i, x_j) = \delta(X_i, X_j)$ . Let  $\varepsilon > 0$  and set

$$A_n = \left\{ \left| \delta(X_0, X_1) - \frac{d_{\text{gr}}^{(n)}(x_0, x_1)}{\chi_n} \right| \geq \varepsilon \delta(X_0, X_1) \right\}.$$

Our goal is to show that the probability that  $A_n$  is realized for infinitely many  $n$ 's is 0, thus proving that  $d_{\text{gr}}^{(n)}(x_0, x_1)/\chi_n$  almost surely converges towards  $\delta(X_0, X_1)$ . Identifying the limit with (8) then proves the statement of the lemma. To do this, let  $\beta \in (0, \bar{\alpha})$  and consider the events  $B_n = \{d_{\text{gr}}^{(n)}(x_0, x_1) \geq n^\beta\}$ . By (8) (and the fact that  $\delta(X_0, X_1) > 0$  a.s.) we get that  $\mathbb{P}(\cup_{k \geq 1} \cap_{n \geq k} B_n) = 1$ . Hence,

$$\mathbb{P}(A_n \text{ i.o.}) = \mathbb{P}(A_n \cap B_n \text{ i.o.}),$$

where *i.o.* means ‘‘is realized for infinitely many  $n$ 's’’. Using (6), we get that  $\mathbb{P}(A_n \cap B_n)$  is less than  $c_\varepsilon^{-1} \exp(-c_\varepsilon n^\beta)$  hence by Borel–Cantelli  $\mathbb{P}(A_n \cap B_n \text{ i.o.}) = 0$ . This completes the proof of the lemma.

Coming back to the proof of Theorem 5(ii), for any  $n \geq 1$ , conditionally on  $\tau_n$  and on  $\mathcal{L}_n$ , let us evaluate the probability that two vertices  $a, b \in \tau_n$  are such that

$$\left| d_\ell^{(n)}(a, b) - \frac{d_{\text{gr}}^{(n)}(a, b)}{\chi_n} \right| \geq \varepsilon (\text{Diam} \vee 1) + \sqrt{\chi_n^{-1}}, \tag{*}$$

where  $\text{Diam}$  is the diameter of  $\mathcal{T}_\alpha$ , that is the maximal distance between any pair of points in  $\mathcal{T}_\alpha$ . We remark that  $\text{Diam}$  bounds from above all the quantities  $d_\ell^{(n)}(a, b)$ . We then split the cases according to the graph distance between  $a$  and  $b$ :

- if  $d_{\text{gr}}^{(n)}(a, b) \geq \sqrt{\chi_n}$  then the conditional probability of (\*) is bounded by  $c_\varepsilon^{-1} \exp(-c_\varepsilon \sqrt{\chi_n})$  by (6),
- if  $d_{\text{gr}}^{(n)}(a, b) < \sqrt{\chi_n}$  then the conditional probability of (\*) is bounded by  $c_\varepsilon^{-1} \exp(-c_\varepsilon \sqrt{\chi_n})$  by (7).

We then use a conditional version of Borel–Cantelli’s lemma exactly as in the proof of Lemma 7. Specifically, let  $C_n = \{\chi_n \geq n^{\bar{\alpha}}\}$  so that Lemma 7 implies  $\mathbb{P}(\cup_{k \geq 1} \cap_{n \geq k} C_n) = 1$ . Let also  $D_n = \{(*) \text{ holds for some } a, b \in \tau_n\}$ . We thus have  $\mathbb{P}(D_n \text{ i.o.}) = \mathbb{P}(C_n \cap D_n \text{ i.o.})$ . Since there are deterministically less than  $2n$  vertices in  $\tau_n$ , the preceding discussion shows that

$$\mathbb{P}(C_n \cap D_n) \leq \mathbb{P}(D_n \mid C_n) \leq \mathbb{E} \left[ c_\varepsilon^{-1} 4n^2 \exp(-c_\varepsilon \sqrt{\chi_n}) \mid C_n \right] \leq c_\varepsilon^{-1} 4n^2 \exp(-c_\varepsilon n^{\bar{\alpha}/2}).$$

Hence by the standard Borel–Cantelli Lemma we have  $\mathbb{P}(C_n \cap D_n \text{ i.o.}) = 0$  and thus  $\mathbb{P}(D_n \text{ i.o.}) = 0$  as well. Consequently,

$$\sup_{a,b \in \tau_n} \left| \frac{d_{\text{gr}}^{(n)}(a,b)}{\chi_n} - d_\ell^{(n)}(a,b) \right| \xrightarrow[n \rightarrow \infty]{\text{a.s.}} 0.$$

Finally we use Lemma 7 again to replace  $\chi_n$  by  $\alpha n^{\tilde{\alpha}}$  in the last display and get (5).

The last point of the theorem is obtained in the same spirit. Since conditionally on  $\mathcal{T}_\alpha$  the  $X_i$ 's are i.i.d. according to  $\mu_\alpha$  it follows that  $\mu_\alpha^{(n)} = \frac{1}{n+1} \sum_{i=0}^n \delta_{X_i}$  is the empirical measure associated to  $\mu_\alpha$  and then

$$(\mathcal{T}_\alpha, \mu_\alpha^{(n)}) \xrightarrow[n \rightarrow \infty]{\text{a.s.}} (\mathcal{T}_\alpha, \mu_\alpha),$$

in the Gromov–Hausdorff–Prokhorov sense. Furthermore, since  $\text{Span}(\mathcal{T}_\alpha; X_0, \dots, X_n)$  is a subtree of  $\mathcal{T}_\alpha$  that converges in the Gromov–Hausdorff sense towards  $\mathcal{T}_\alpha$  the last display holds with  $(\mathcal{T}_\alpha, \mu_\alpha^{(n)})$  replaced by  $(\text{Span}(\mathcal{T}_\alpha; X_0, \dots, X_n), \mu_\alpha^{(n)})$ . Recall now that  $\text{Span}(\mathcal{T}_\alpha; X_0, \dots, X_n)$  can be seen as the tree  $\tau_n$  where each edge has been replaced by a Euclidean segment of length  $\ell_e^{(n)}$ . This identification transports the measure  $\mu_\alpha^{(n)}$  onto the discrete atomic measure  $\mu_n$ . Thus it suffices to prove that

$$d_{\text{GHP}} \left( \left( \tau_n, \frac{d_{\text{gr}}^{(n)}}{\alpha n^{\tilde{\alpha}}}, \mu_n \right), \left( \tau_n, d_\ell^{(n)}, \mu_n \right) \right) \xrightarrow[n \rightarrow \infty]{} 0.$$

But this again follows from (5). Indeed, the obvious correspondence  $\mathcal{R} = \{(x, x) : x \in \tau_n\}$  between  $(\tau_n, \frac{d_{\text{gr}}^{(n)}}{\alpha n^{\tilde{\alpha}}})$  and  $(\tau_n, d_\ell^{(n)})$  has a vanishing distortion as  $n \rightarrow \infty$  by (5). This means (see [10]) that for any  $\varepsilon > 0$  and for all  $n$  large enough we can isometrically embed these two trees into a common metric space  $(F_n, \delta_n)$  such that two elements in correspondence (that is the same vertex of  $\tau_n$  in the two embeddings of the trees) are at  $\delta_n$ -distance less than  $\varepsilon$  from each other. The image measures  $\nu_n$  and  $\nu'_n$  of  $\mu_n$  by these embeddings then obviously satisfy  $\nu_n(C) \leq \nu'_n(C^\varepsilon)$  and  $\nu'_n(C) \leq \nu_n(C^\varepsilon)$  for all Borel  $C \subset F_n$ . This suffices to prove the claim and to finish the proof of the theorem. □

### 3 The discrete approach

This section is mainly devoted to the construction of the  $\alpha'$ -stable subtrees  $\mathfrak{T}_{\alpha,\alpha'}$  of  $\mathcal{T}_\alpha$ , for  $\alpha' > \alpha$ , as stated in Proposition 2, and then to the proof of Theorem 1. In order to establish the key couplings of Marchal's constructions that will be used for the proofs, we introduce once and for all a sequence  $U_i, i \geq 2$  of i.i.d. random variables uniformly distributed on  $(0, 1)$  that are independent of all other variables introduced so far and later on.

### 3.1 Coupling different Marchal’s constructions

Fix  $\alpha \in (1, 2)$ . We will couple Marchal’s constructions for different values of  $\alpha' \in (\alpha, 2]$  in such a way that the  $\alpha'$ -constructions are nested “sub-constructions” of the  $\alpha$ -one. To see this, fix  $\alpha' \in (\alpha, 2]$  and recall the definition of the probabilities  $p_{\alpha,\alpha',d,d'}$  in (1), as well as the definition of Marchal’s Markov chain  $(\mathbf{T}_\alpha(n), n \geq 1)$  in Sect. 2.3. The idea is to enrich the construction of this Markov chain with two colors, blue and red, such that the only edge and the two leaves of  $\mathbf{T}_\alpha(1)$  are blue and the colors evolve recursively according to the following rules, see Fig. 5. At step  $i \geq 2$ :

- if an edge of color  $\mathbf{C} \in \{\text{red,blue}\}$  is selected, then we split this edge into two  $\mathbf{C}$  edges divided by a  $\mathbf{C}$  vertex and attach a  $\mathbf{C}$  edge-leaf to it;
- if a vertex is selected, then the new edge-leaf attached on it is blue if

$$U_i \leq p_{\alpha,\alpha',d_i,d'_i}$$

and red otherwise, where  $d_i$  is the degree of the vertex in the full tree  $\mathbf{T}_\alpha(i - 1)$  and  $d'_i$  is its degree in the blue subtree of  $\mathbf{T}_\alpha(i - 1)$ , with the convention that  $d'_i = 0$  if this vertex is red.

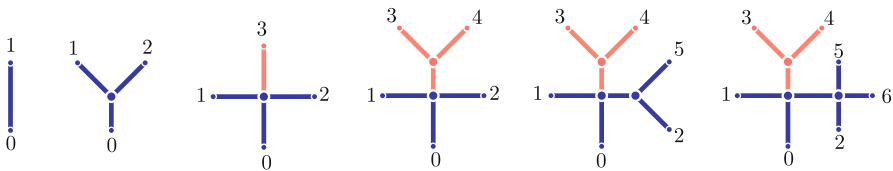


Fig. 5 Illustration of the construction of the blue tree (color figure online)

We denote by  $\mathbf{T}_{\alpha,\alpha'}(n)$  the subtree of  $\mathbf{T}_\alpha(n)$  spanned by the blue edges and leaves. Remark that when  $\alpha' = 2$ , the coloring rules are particularly simple (and deterministic): each edge-leaf attached on a red edge or on a vertex (blue or red) is red, whereas an edge-leaf attached on a blue edge is blue. In this case the blue subtree is therefore binary. Remark also that the function  $\alpha' \in (\alpha, 2] \mapsto p_{\alpha,\alpha',d,d'}$  is decreasing, hence for every  $n \geq 1$  the function  $\alpha' \in (\alpha, 2] \mapsto T_{\alpha,\alpha'}(n)$  is decreasing for the inclusion inside  $T_\alpha(n)$ .

The labeling of  $\mathbf{T}_{\alpha,\alpha'}(n)$  is given by sorting the leaves of  $\mathbf{T}_{\alpha,\alpha'}(n)$  according to their labels in  $\mathbf{T}_\alpha(n)$ . Also, we let

$$L_{\alpha,\alpha'}(n) = \text{number of leaves of } \mathbf{T}_{\alpha,\alpha'}(n) \text{ minus } 1$$

and  $L_{\alpha,\alpha'}^{-1}(n) = \inf\{k \geq 1 : L_{\alpha,\alpha'}(k) = n\}$  its inverse. Here is the main observation:

**Lemma 8** *The subchain  $\mathbf{T}_{\alpha,\alpha'}(\cdot)$  evolves as a time-changed Marchal’s construction with parameter  $\alpha'$ . More precisely, we have the following identity in distribution*



$$\left( \mathbf{T}_{\alpha,\alpha'} \left( L_{\alpha,\alpha'}^{-1}(n) \right) \right)_{n \geq 1} \stackrel{(d)}{=} (\mathbf{T}_{\alpha'}(n))_{n \geq 1}.$$

Futhermore,  $(\mathbf{T}_{\alpha,\alpha'}(L_{\alpha,\alpha'}^{-1}(n)))_{n \geq 1}$  is independent of  $(L_{\alpha,\alpha'}(n))_{n \geq 1}$ .

*Proof* We examine the transition probabilities of the chain  $(\mathbf{T}_\alpha(n), \mathbf{T}_{\alpha,\alpha'}(n))_{n \geq 1}$ . Recall that in Marchal’s construction, a weight  $\alpha - 1$  is assigned to each edge and a weight  $d - 1 - \alpha$  to each vertex of degree  $d \geq 3$ . Without changing the dynamic, we multiply all these weights by  $(\alpha' - 1)/(\alpha - 1)$ , so that each edge has now a weight  $\alpha' - 1$  and each vertex a weight  $(d - 1 - \alpha)(\alpha' - 1)/(\alpha - 1)$ . By (2), the total weight of the tree  $\mathbf{T}_\alpha(n)$  is now

$$\overline{W}(\mathbf{T}_\alpha(n)) = \frac{\alpha' - 1}{\alpha - 1} (n\alpha - 1).$$

Hence, the new edge-leaf added to  $\mathbf{T}_\alpha(n)$  to get  $\mathbf{T}_\alpha(n + 1)$  is blue if it has been

- grafted on an edge of  $\mathbf{T}_{\alpha,\alpha'}(n)$ , which occurs with probability  $(\alpha' - 1)/\overline{W}(\mathbf{T}_\alpha(n))$  for each edge of  $\mathbf{T}_{\alpha,\alpha'}(n)$
- or grafted on a vertex of  $\mathbf{T}_{\alpha,\alpha'}(n)$  and then colored in blue, which occurs with probability

$$\frac{(d - 1 - \alpha)(\alpha' - 1)(\alpha - 1)^{-1} \times p_{\alpha,\alpha',d,d'}}{\overline{W}(\mathbf{T}_\alpha(n))} = \frac{d' - 1 - \alpha'}{\overline{W}(\mathbf{T}_\alpha(n))}$$

when  $d'$  is the degree of the vertex in  $\mathbf{T}_{\alpha,\alpha'}(n)$  and  $d$  its full degree in  $\mathbf{T}_\alpha(n)$ .

Observe that these probabilities are proportional to the weights assigned in Marchal’s construction with parameter  $\alpha'$ . Consequently their sum, which is the total probability that the added edge-leaf is blue, is equal to

$$\frac{L_{\alpha,\alpha'}(n)\alpha' - 1}{\overline{W}(\mathbf{T}_\alpha(n))},$$

where we recall that  $L_{\alpha,\alpha'}(n) + 1$  is the number of leaves of  $\mathbf{T}_{\alpha,\alpha'}(n)$ . Moreover, conditionally on the fact that the added edge-leaf is blue, it is grafted on an edge or vertex of  $\mathbf{T}_{\alpha,\alpha'}(n)$  with the dynamic of Marchal’s algorithm with parameter  $\alpha'$ . In other words, the transition probabilities of the chain  $(\mathbf{T}_\alpha(n), \mathbf{T}_{\alpha,\alpha'}(n))_{n \geq 1}$  can be described as follows. First,  $L_{\alpha,\alpha'}(n), n \geq 1$  is a Markov chain with transition probabilities

$$\begin{aligned} \mathbb{P} \left( L_{\alpha,\alpha'}(n + 1) = L_{\alpha,\alpha'}(n) + 1 \mid \mathcal{F}_n \right) &= 1 - \mathbb{P} \left( L_{\alpha,\alpha'}(n + 1) = L_{\alpha,\alpha'}(n) \mid \mathcal{F}_n \right) \\ &= \frac{L_{\alpha,\alpha'}(n)\alpha' - 1}{\overline{W}(\mathbf{T}_\alpha(n))}, \end{aligned} \tag{9}$$

where  $\mathcal{F}_n$  denotes the sigma-field generated by  $(\mathbf{T}_\alpha(k), \mathbf{T}_{\alpha,\alpha'}(k), k \leq n)$ . Second, conditionally on  $L_{\alpha,\alpha'}(n + 1) = L_{\alpha,\alpha'}(n) + 1$ , the transition probabilities on  $\mathbf{T}_{\alpha,\alpha'}(n)$  are those of Marchal’s construction with parameter  $\alpha'$ , whereas conditionally on

$L_{\alpha,\alpha'}(n + 1) = L_{\alpha,\alpha'}(n)$ , the transition on the red component  $\mathbf{T}_\alpha(n) \setminus \mathbf{T}_{\alpha,\alpha'}(n)$  follows the appropriate conditional distribution. The statements of the lemma follow from these considerations.

The chain  $L_{\alpha,\alpha'}$  can easily be studied using (9). Let us introduce its limit distribution (after re-normalization). A *generalized Mittag-Leffler* random variable of parameters  $(\beta, \theta)$  with  $\beta \in (0, 1)$  and  $\theta > -\beta$  is a random variable denoted by  $ML_{\beta,\theta}$  which is characterized by its positive moments

$$\mathbb{E} \left[ ML_{\beta,\theta}^p \right] = \frac{\Gamma(\theta + 1)\Gamma(\theta/\beta + p + 1)}{\Gamma(\theta/\beta + 1)\Gamma(\theta + p\beta + 1)}, \quad p \geq 0. \tag{10}$$

This random variable is actually a biased version of a power of a stable distribution of index  $\beta$ , see [24, Sect. 3] for background. For later use, remark that if  $0 < a < b < 1$  and if  $ML_{a,a}$ ,  $ML_{b,b}$  and  $ML_{a/b,a}$  are independent generalized Mittag-Leffler with parameters  $(a, a)$ ,  $(b, b)$  and  $(a/b, a)$  respectively, then we have the following identity in distribution

$$ML_{a,a} = (ML_{a/b,a})^b \cdot ML_{b,b}, \tag{11}$$

this can be easily checked using moments (10).

**Lemma 9** *There exists a generalized Mittag-Leffler r.v.  $ML_{\bar{\alpha}/\bar{\alpha}',\bar{\alpha}}$  such that*

$$\frac{L_{\alpha,\alpha'}(n)}{n^{\bar{\alpha}/\bar{\alpha}'}} \xrightarrow[n \rightarrow \infty]{a.s.} ML_{\bar{\alpha}/\bar{\alpha}',\bar{\alpha}}.$$

*Proof* It is possible to analyse the chain  $(L_{\alpha,\alpha'}(n))_{n \geq 1}$  “by hand” using the transition probabilities (9). However we bypass any calculation by using a connection with Pitman’s Chinese restaurant process, see [24, Sect. 3]. Indeed, it is straightforward to check from (9) that the sequence  $(L_{\alpha,\alpha'}(n + 1) - 1, n \geq 1)$  is distributed as the sequence of the number of tables in a  $(\bar{\alpha}/\bar{\alpha}', \bar{\alpha})$ -Chinese restaurant process, see [24, Sect. 3.2.3]. The result then follows from [24, Theorem 3.8].  $\square$

### 3.2 Proofs of Proposition 2 and Theorem 1

Let  $1 < \alpha < \alpha' \leq 2$  and consider an  $\alpha$ -stable tree  $\mathcal{T}_\alpha$  and its uniform mass measure  $\mu_\alpha$ . By Theorem 5, there exists a version of Marchal’s Markov chain  $(\mathbf{T}_\alpha(n), n \geq 1)$  such that

$$\left( \frac{T_\alpha(n)}{\alpha n^{\bar{\alpha}}}, \frac{\sum_{i=0}^n \delta_{A_i}}{n + 1} \right) \xrightarrow[n \rightarrow \infty]{a.s.} (\mathcal{T}_\alpha, \mu_\alpha), \tag{12}$$

in the Gromov–Hausdorff–Prokhorov sense. We will use the blue subchain  $(\mathbf{T}_{\alpha,\alpha'}(n), n \geq 1)$  constructed from  $(\mathbf{T}_\alpha(n), n \geq 1)$  in the previous section to build the closed subtree  $\mathfrak{T}_{\alpha,\alpha'}$  of Proposition 2.

*Proof of Proposition 2* Recall Lemma 8 and note from Lemma 9 that  $L_{\alpha,\alpha'}(n) \rightarrow \infty$  a.s. as  $n \rightarrow \infty$ . Then apply Theorem 5 to obtain the almost sure convergence

$$\frac{T_{\alpha,\alpha'}(n)}{L_{\alpha,\alpha'}(n)^{\bar{\alpha}'}} \xrightarrow[n \rightarrow \infty]{\text{a.s.}} \alpha' \mathcal{T}_{\alpha'},$$

in the Gromov–Hausdorff sense, where  $\mathcal{T}_{\alpha'}$  denotes a stable tree of index  $\alpha'$  (which is not independent of  $\mathcal{T}_{\alpha}$ !). Furthermore, still by Lemma 8, the tree  $\mathcal{T}_{\alpha'}$  is independent of the sequence  $(L_{\alpha,\alpha'}(n), n \geq 1)$ , hence also of its limit after rescaling. Hence

$$\begin{aligned} \frac{T_{\alpha,\alpha'}(n)}{\alpha n^{\bar{\alpha}}} &= \frac{T_{\alpha,\alpha'}(n)}{\alpha L_{\alpha,\alpha'}(n)^{\bar{\alpha}'}} \times \left( \frac{L_{\alpha,\alpha'}(n)}{n^{\bar{\alpha}/\bar{\alpha}'}} \right)^{\bar{\alpha}'} \\ &\xrightarrow[n \rightarrow \infty]{\text{a.s.}} \frac{\alpha'}{\alpha} (\text{ML}_{\bar{\alpha}/\bar{\alpha}',\bar{\alpha}})^{\bar{\alpha}'} \mathcal{T}_{\alpha'} =: \mathfrak{T}_{\alpha,\alpha'}, \end{aligned}$$

where  $\text{ML}_{\bar{\alpha}/\bar{\alpha}',\bar{\alpha}}$  denotes the Mittag-Leffler random variable of Lemma 9, which is independent of  $\mathcal{T}_{\alpha'}$ . The variable  $M_{\alpha,\alpha'}$  of Proposition 2 is thus equal to  $(\alpha'/\alpha)^{1/\bar{\alpha}'}\text{ML}_{\bar{\alpha}/\bar{\alpha}',\bar{\alpha}}$  and the tree  $\mathfrak{T}_{\alpha,\alpha'}$  can be realized as a closed subtree of  $\mathcal{T}_{\alpha}$ .  $\square$

In the constructions of  $\mathbf{T}_{\alpha}(n)$  and  $\mathbf{T}_{\alpha,\alpha'}(n)$ , the two leaves  $A_0$  and  $A_1$  are kept in both trees and provide in the limit, by Theorem 5, two independent uniform points in  $\mathcal{T}_{\alpha}$  and in  $\mathfrak{T}_{\alpha,\alpha'}$ . More precisely, for  $\beta \in (1, 2]$  we denote by  $I_{\beta}$  the height between two uniform points in a  $\beta$ -stable tree. We also write  $H_n$  the distance between  $A_0$  and  $A_1$  in  $\mathbf{T}_{\alpha}(n)$ . Then we have by Theorem 5

$$\frac{H_n}{n^{\bar{\alpha}}} \xrightarrow[n \rightarrow \infty]{\text{a.s.}} \alpha I_{\alpha}.$$

The distance  $H_n$  is also the distance between  $A_0$  and  $A_1$  in  $\mathbf{T}_{\alpha,\alpha'}(n)$ . Hence, by Lemmas 8 and 9, combined with Theorem 5,

$$\frac{H_n}{n^{\bar{\alpha}}} = \frac{H_n}{L_{\alpha,\alpha'}(n)^{\bar{\alpha}'}} \times \left( \frac{L_{\alpha,\alpha'}(n)}{n^{\bar{\alpha}/\bar{\alpha}'}} \right)^{\bar{\alpha}'} \xrightarrow[n \rightarrow \infty]{\text{a.s.}} \alpha' I_{\alpha'} (\text{ML}_{\bar{\alpha}/\bar{\alpha}',\bar{\alpha}})^{\bar{\alpha}'},$$

where, in the limit, the Mittag-Leffler random variable and  $I_{\alpha'}$  are independent. Thus if for  $\alpha < \alpha' \in (1, 2]$  we set

$$Q_{\alpha \rightarrow \alpha'} \stackrel{(d)}{=} \frac{\alpha'}{\alpha} (\text{ML}_{\bar{\alpha}/\bar{\alpha}',\bar{\alpha}})^{\bar{\alpha}'} = (M_{\alpha,\alpha'})^{\bar{\alpha}'},$$

we can easily check that for  $a < b < c \in (1, 2]$ , we have  $Q_{a \rightarrow c} = Q_{a \rightarrow b} \cdot Q_{b \rightarrow c}$  in distribution where the two last variables are independent. Hence,  $Q$  can be interpreted as a transition kernel. Furthermore the random height in the stable trees form an invariant family for  $Q$  in the sense that

$$I_{\alpha} \stackrel{(d)}{=} Q_{\alpha \rightarrow \alpha'} \cdot I_{\alpha'} \tag{13}$$

with the two variables on the right-hand side independent. This could have been proved directly using moments and using the fact (see e.g. [22]) that  $\alpha I_\alpha$  is distributed as  $ML_{\bar{\alpha}, \bar{\alpha}}$  for all  $\alpha \in (1, 2]$ . Hence (13) reduces to (11).

We now aim at a backward analog of (13). As in Theorem 1, let  $J_\alpha$  be distributed as the power of a Gamma distribution, namely  $\alpha(\Gamma_{1+\bar{\alpha}})^{\bar{\alpha}}$  for  $\alpha \in (1, 2]$ . In particular the positive moments of  $J_\alpha$  are given by

$$\mathbb{E}[J_\alpha^p] = \alpha^p \frac{\Gamma((p+1)\bar{\alpha} + 1)}{\Gamma(\bar{\alpha} + 1)}, \quad \text{for } p \geq 0.$$

These variables have been designed to satisfy the ‘‘dual’’ relations of (13), namely, for  $1 < \alpha < \alpha' \leq 2$ , one can check using moments that the following identity in distribution holds

$$J_\alpha \cdot Q_{\alpha \rightarrow \alpha'} = J_{\alpha'}, \tag{14}$$

when the two first random variables are independent.

*Proof of Theorem 1* Let  $\mathcal{T}_\alpha$  be a stable tree  $\mathcal{T}_\alpha$  that has been rescaled by an independent copy of  $J_\alpha$ , that is  $\mathcal{T}_\alpha = J_\alpha \cdot \mathcal{T}_\alpha$ . For  $\alpha' \in (\alpha, 2]$ , using Proposition 2, we can find inside  $\mathcal{T}_\alpha$  a subtree  $\mathfrak{T}_{\alpha, \alpha'}$  which is distributed as an  $\alpha'$ -stable tree rescaled by an independent factor distributed as  $Q_{\alpha \rightarrow \alpha'}$ . Hence, after rescaling by  $J_\alpha$  this provides a subtree of  $\mathcal{T}_\alpha$  which is distributed as

$$J_\alpha \cdot Q_{\alpha \rightarrow \alpha'} \mathcal{T}_{\alpha'} \stackrel{(d)}{=} J_{\alpha'} \cdot \mathcal{T}_{\alpha'} \stackrel{(d)}{=} \mathcal{T}_{\alpha'},$$

since  $Q_{\alpha \rightarrow \alpha'}$  and  $J_\alpha$  are independent. Iterating, we obtain for all finite increasing sequences  $1 < \alpha_1 < \dots < \alpha_n \leq 2$  a sequence of nested trees  $\mathcal{T}_{\alpha_n} \subset \dots \subset \mathcal{T}_{\alpha_1}$ . Using Kolmogorov’s extension theorem we can thus construct a process  $(\mathcal{T}_\alpha)_{1 < \alpha \leq 2}$  of nested rescaled stable trees. □

Still using moments, note that

$$J_\alpha I_\alpha \stackrel{(d)}{=} \Gamma_2,$$

when  $I_\alpha$  and  $J_\alpha$  are independent. This implies that the distance between two independent uniform points in  $\mathcal{T}_\alpha$  has a Gamma distribution of parameter 2 (that is with density  $x e^{-x}$  on  $(0, \infty)$ ) for all  $\alpha \in (1, 2]$ . For the nested sequence of trees  $(\mathcal{T}_\alpha)_{1 < \alpha \leq 2}$ , it is then possible to couple the choice of the two uniform independent points so that their distance is the *same* in all trees  $\mathcal{T}_\alpha$ .

### 3.3 Zero measure

**Proposition 10** *For all  $\alpha' \in (\alpha, 2]$ ,*

$$\mu_\alpha(\mathfrak{T}_{\alpha, \alpha'}) = 0 \quad \text{a.s.}$$

*Proof* We will use that  $(\mathcal{T}_\alpha, \mu_\alpha)$  is obtained from a Marchal’s Markov chain  $\mathbf{T}_\alpha$  via (12), jointly with the fact that  $\mathfrak{T}_{\alpha,\alpha'}$  is the scaling limit of the subchain  $\mathbf{T}_{\alpha,\alpha'}$ . Since, by Lemma 9, the number of leaves of  $\mathbf{T}_{\alpha,\alpha'}(n)$  is proportional to  $n^{\bar{\alpha}/\bar{\alpha}'} \ll n$ , it should be intuitively clear that  $\mathfrak{T}_{\alpha,\alpha'}$  has zero mass inside  $\mathcal{T}_\alpha$ . We sketch here a more rigorous argument. Let  $\gamma$  denote the expectation of the  $\mu_\alpha$ -mass of  $\mathfrak{T}_{\alpha,\alpha'}$ :

$$\gamma = \mathbb{E} [\mu_\alpha(\mathfrak{T}_{\alpha,\alpha'})],$$

our goal being to show that  $\gamma = 0$ . For that purpose we use the self-similarity of Marchal’s construction. Indeed, if we stop this construction after one step, we obtain the tree  $\mathbf{T}_\alpha(2)$  made of a “Y” with three leaves  $A_0, A_1$  and  $A_2$  joined by a branch point denoted by  $\mathcal{B}_2$ . In the future evolution of the process, the three edges  $e_0 = \{\mathcal{B}_2, A_0\}, e_1 = \{\mathcal{B}_2, A_1\}$  and  $e_2 = \{\mathcal{B}_2, A_2\}$  will give rise in the scaling limit to three copies of rescaled stable trees denoted by  $\tau_0, \tau_1$  and  $\tau_2$ . Similarly, the edges that will later be grafted on  $\mathcal{B}_2$  will also give rise to a countable collection of continuous trees  $\tau_3, \tau_4, \dots$  such that the tree  $\mathcal{T}_\alpha$  is made of the glueing of all the  $\tau_i$ ’s by one vertex. It is clear from Marchal’s construction, and (12), that the measured trees  $(\tau_i, \mu_{\alpha,i}), i \geq 0$ , where  $\mu_{\alpha,i}$  denotes the restriction of the measure  $\mu_\alpha$  to  $\tau_i$ , satisfy

$$\tau_i = \mu_\alpha(\tau_i) \bar{\alpha} \mathcal{T}_\alpha^{(i)}, \quad \mu_{\alpha,i} = \mu_\alpha(\tau_i) \mu_\alpha^{(i)},$$

where the measured trees  $(\mathcal{T}_\alpha^{(i)}, \mu_\alpha^{(i)}), i \geq 0$  are i.i.d. copies of  $(\mathcal{T}_\alpha, \mu_\alpha)$ , that are moreover independent of the masses  $\mu_\alpha(\tau_i), i \geq 0$ .

Let us now examine the evolution of the “blue” subprocess  $\mathbf{T}_{\alpha,\alpha'}$  inside each of these trees. First, it is plain that the construction of  $\mathbf{T}_{\alpha,\alpha'}$  restricted to each offspring subtree of the three initial edges  $e_0, e_1$  and  $e_2$  follows exactly the rules of a subchain of parameters  $\alpha, \alpha'$  inside this subtree. This gives in the limit a subtree  $t_i \subset \tau_i$ , for  $i \in \{0, 1, 2\}$  that satisfies

$$\gamma = \frac{\mathbb{E}[\mu_\alpha(t_i)]}{\mathbb{E}[\mu_\alpha(\tau_i)]}.$$

For  $i \geq 3$ , the situation is almost the same, provided that the first edge grafted on  $\mathcal{B}_2$  that will give rise to  $\tau_i$  is blue. If this is not the case, that is the ancestral edge creating  $\tau_i$  is red, then the blue subprocess does not “enter” this part of the tree and we set  $t_i = \emptyset$ . If the ancestral edge creating  $\tau_i$  is blue, then the blue subprocess enters this part of the tree and provides in the limit a subtree  $t_i \subset \tau_i$  such that

$$\gamma = \frac{\mathbb{E}[\mu_\alpha(t_i) \mid t_i \neq \emptyset]}{\mathbb{E}[\mu_\alpha(\tau_i)]}.$$

The subtree  $\mathfrak{T}_{\alpha,\alpha'}$  is thus made of the union of the  $t_i, i \geq 0$  and we have

$$\gamma = \mathbb{E} \left[ \sum_{i \geq 0} \mu_\alpha(t_i) \mathbf{1}_{\{t_i \neq \emptyset\}} \right] = \gamma \sum_{i \geq 0} \mathbb{P}(t_i \neq \emptyset) \mathbb{E} [\mu_\alpha(\tau_i)]. \tag{15}$$

To conclude, note that  $\mathbb{P}(\exists i : t_i = \emptyset) = 1$ . Indeed, the probability that for all  $\tau_i, i \geq 3$  the ancestral edges are blue is

$$\prod_{d \geq 3} \frac{(d - 1 - \alpha')(\alpha - 1)}{(d - 1 - \alpha)(\alpha' - 1)} = 0,$$

since  $\alpha' > \alpha$ . Together with the fact that  $\sum_{i \geq 0} \mu_\alpha(\tau_i) = 1$  a.s., this implies that the sum in (15) is strictly less than 1. Hence  $\gamma = 0$ .

### 4 Pruning of discrete and continuous trees

The goal of this section is to present a “geometric” way of retrieving  $\mathbf{T}_{\alpha,\alpha'}(n)$  from the labeled tree  $\mathbf{T}_\alpha(n)$ . Although equivalent to the iterative construction of  $\mathbf{T}_{\alpha,\alpha'}(n)$  this new procedure, called “pruning operation” is well-suited to pass to the continuous limit and yields Theorem 3.

#### 4.1 Pruning

Let  $1 < \alpha < \alpha' \leq 2$ . We recall and extend the pruning procedure of stable trees described in the Introduction to the context of continuous or discrete trees. This procedure depends on the probabilities  $p_{\alpha,\alpha',d,d'}$  and uses the sequence of uniform random variables  $(U_i)_{i \geq 2}$  introduced in Sect. 2.3. Let  $(t; x_0, x_1, x_2, \dots, x_n)$  be a discrete or continuous tree given with an ordered subset of distinct leaves. We remove randomly some edges in this labeled tree so as to obtain a subtree of  $t$  (given with an ordered subset of its leaves) that we denote by  $\text{Prun}_{\alpha,\alpha'}(t; x_0, \dots, x_n)$  and which is constructed recursively as follows:

- the tree  $\text{Prun}_{\alpha,\alpha'}(t; x_0, x_1)$  is  $[[x_0, x_1]]$ , the tree spanned by  $x_0$  and  $x_1$  in  $t$ ;
- for  $i \geq 2$ , consider two trees:  $t_{i-1}$ , the tree spanned by  $x_0, \dots, x_{i-1}$  in  $t$ , and  $\tau_{i-1} = \text{Prun}_{\alpha,\alpha'}(t; x_0, \dots, x_{i-1})$ . In  $t$ , the vertex  $x_i$  is attached to  $t_{i-1}$  via the point  $\Delta_i$ , that is

$$t_{i-1} \cap [[x_i, \Delta_i]] = \{\Delta_i\}.$$

Let then  $d_i$  denote the degree of  $\Delta_i$  in  $t_{i-1}$  and  $d'_i$  its degree in  $\tau_{i-1}$  with the usual convention that  $d'_i = 0$  if  $\Delta_i \notin \tau_{i-1}$ . See Fig. 2. We then set:

$$\begin{aligned} \text{Prun}_{\alpha,\alpha'}(t; x_0, \dots, x_i) &= \tau_{i-1} \cup [[x_i, \Delta_i]] && \text{if } U_i \leq p_{\alpha,\alpha',d_i,d'_i}, \\ \text{Prun}_{\alpha,\alpha'}(t; x_0, \dots, x_i) &= \tau_{i-1} && \text{otherwise.} \end{aligned}$$

Note that  $\text{Prun}_{\alpha,\alpha'}(t; x_0, \dots, x_n)$  is a function of  $(t; x_0, \dots, x_n)$  and of the random variables  $(U_i)_{i \geq 2}$ . This tree can be given with an ordered sequence of leaves, which is the subsequence of elements of  $\{x_0, x_1, \dots, x_n\}$  that end-up in  $\text{Prun}_{\alpha,\alpha'}(t; x_0, \dots, x_n)$  and following our principles we denote this labeled tree by  $\mathbf{Prun}_{\alpha,\alpha'}(t; x_0, \dots, x_n)$ . To stick with the color coding that we defined in Sect. 3.1 we think of the pruned

subtree as a blue subtree of  $t$  whose complement in  $t_n$  is red. In particular a leaf  $x_i$  is blue if it belongs to the pruned subtree. All these constructions are coupled *via* the  $U'_i$ 's in such a way that

$$\text{Prun}_{\alpha,\alpha'}(t; x_0, \dots, x_{n-1}) \subset \text{Prun}_{\alpha,\alpha'}(t; x_0, \dots, x_n), \quad \forall n \geq 1,$$

and for every  $\alpha$  the map  $\alpha' \in (\alpha, 2] \mapsto \text{Prun}_{\alpha,\alpha'}(t; x_0, \dots, x_n)$  is decreasing for the inclusion in  $t$ . Thus, when  $t$  is a continuous tree and  $x_0, \dots, x_n, \dots$  is an infinite sequence of (distinct) leaves in  $t$ , we can define

$$\text{Prun}_{\alpha,\alpha'}(t; (x_i)_{i \geq 0}) = \overline{\bigcup_{n \geq 1} \text{Prun}_{\alpha,\alpha'}(t; x_0, \dots, x_n)}.$$

Note also that when  $\alpha' = 2$ , since  $p_{\alpha,2,d,3} = 0$ , the tree  $\text{Prun}_{\alpha,\alpha'}(t; x_0, \dots, x_n)$  is binary for all  $n \geq 1$ . In this case, the pruning procedure is deterministic (it does not depend on the  $U_i$ 's): the leaves  $x_i$  are just added one by one provided that the tree remains binary. Notice also that the pruning constructions are coupled with the chain  $\mathbf{T}_{\alpha,\alpha'}$  via the  $U_i$ 's.

Now, recall that  $\mathbf{T}_\alpha(n)$  is a discrete tree with  $n + 1$  ordered leaves  $A_0, \dots, A_n$ . The pruning has been defined such that the blue subtree  $\mathbf{T}_{\alpha,\alpha'}(n)$  can also be seen as  $\mathbf{Prun}_{\alpha,\alpha'}(\mathbf{T}_\alpha(n))$ :

**Proposition 11** *For all  $1 < \alpha < \alpha' \leq 2$  and for every  $n \geq 1$ , we have*

$$\mathbf{Prun}_{\alpha,\alpha'}(\mathbf{T}_\alpha(n)) = \mathbf{T}_{\alpha,\alpha'}(n).$$

*Proof* This is easy to prove by induction on  $n$  and is safely left to the reader.

### 4.2 Proof of Theorem 3

We start with a continuity property of the application  $\text{Prun}(\cdot)$  with respect to the leaves. It applies both in the discrete and continuous settings and will be useful to prove Theorem 3.

**Proposition 12** *Let  $(t; x_0, \dots, x_n)$  be a discrete or continuous tree given with a subset of  $n + 1$  leaves. Then for any  $k \in \{0, 1, \dots, n\}$  we have*

$$\begin{aligned} & \mathcal{D}(\text{Prun}_{\alpha,\alpha'}(t; x_0, \dots, x_n), \text{Prun}_{\alpha,\alpha'}(t; x_0, \dots, x_k)) \\ & \leq \inf \{ \varepsilon > 0 : x_0, \dots, x_k \text{ is an } \varepsilon\text{-net in } (t, \mathcal{D}) \}, \end{aligned}$$

where  $\mathcal{D}$  denotes either the graph distance if  $t$  is discrete or its metric if  $t$  is a continuous tree.

*Proof* Let  $0 \leq k \leq n$ . By the monotonicity of the pruning operation, the pruned tree  $\text{Prun}_{\alpha,\alpha'}(t; x_0, \dots, x_n)$  can be obtained from  $\text{Prun}_{\alpha,\alpha'}(t; x_0, \dots, x_k)$  by grafting on each of its vertices and edges some trees. The maximal height of these grafted

trees is thus equal to the (Hausdorff) distance between  $\text{Prun}_{\alpha,\alpha'}(t; x_0, \dots, x_k)$  and  $\text{Prun}_{\alpha,\alpha'}(t; x_0, \dots, x_n)$ . We just have to remark that none of the points  $x_0, \dots, x_k$  belong to the grafted trees that is

$$\{x_0, \dots, x_k\} \cap (\text{Prun}_{\alpha,\alpha'}(t; x_0, \dots, x_n) \setminus \text{Prun}_{\alpha,\alpha'}(t; x_0, \dots, x_k)) = \emptyset.$$

The statement of the proposition follows.

*Proof of Theorem 3* Starting with an  $\alpha$ -stable tree  $\mathcal{T}_\alpha$  and a sample  $(X_i, i \geq 0)$  of i.i.d. random leaves with distribution  $\mu_\alpha$  (given  $\mathcal{T}_\alpha$ ), we consider the chain  $\mathbf{T}_\alpha$  such that (12) holds, as well as its blue subchain  $\mathbf{T}_{\alpha,\alpha'}$ . We have already established (see the proof of Proposition 2) that once re-normalized by  $\alpha n^{\bar{\alpha}}$  the blue tree  $T_{\alpha,\alpha'}(n)$  converges almost surely towards  $\mathfrak{T}_{\alpha,\alpha'} \subset \mathcal{T}_\alpha$ . By Proposition 11 we just have to show that  $\text{Prun}_{\alpha,\alpha'}(\mathbf{T}_\alpha(n))$  converges after renormalization towards  $\text{Prun}_{\alpha,\alpha'}(\mathcal{T}_\alpha; (X_i)_{i \geq 0})$ . For this, we will use that for every  $k \geq 0$  (see Theorem 5(ii))

$$\left( \frac{T_\alpha(n)}{\alpha n^{\bar{\alpha}}}; A_0, A_1, \dots, A_k \right) \xrightarrow[n \rightarrow \infty]{\text{a.s.}} (\mathcal{T}_\alpha; X_0, \dots, X_k), \tag{16}$$

in the  $k + 1$ -pointed Gromov–Hausdorff topology. Let  $\varepsilon > 0$ . Almost surely, since the  $X_i$ 's form a dense subset of  $\mathcal{T}_\alpha$ , which is compact, we can find a (random)  $0 \leq k < \infty$  such that  $\{X_0, \dots, X_k\}$  is an  $\varepsilon$ -net in  $\mathcal{T}_\alpha$ . In particular, we eventually have that  $\{A_0, \dots, A_k\}$  is a  $\lfloor 2\alpha\varepsilon n^{\bar{\alpha}} \rfloor$ -net in  $T_\alpha(n)$ . Applying Proposition 12 twice we get that for  $n$  large enough,

$$\begin{aligned} \delta(\text{Prun}_{\alpha,\alpha'}(\mathcal{T}_\alpha; (X_i)_{i \geq 0}), \text{Prun}_{\alpha,\alpha'}(\mathcal{T}_\alpha; X_0, \dots, X_k)) &\leq \varepsilon, \\ \frac{1}{\alpha n^{\bar{\alpha}}} d_{\text{gr}}(\text{Prun}_{\alpha,\alpha'}(\mathbf{T}_\alpha(n)), \text{Prun}_{\alpha,\alpha'}(T_\alpha(n); A_0, \dots, A_k)) &\leq 2\varepsilon, \end{aligned}$$

where  $\delta$  is the metric in  $\mathcal{T}_\alpha$ . Hence the proof will be completed when we will have shown that for every fixed  $k_0 \geq 0$ ,

$$\frac{1}{\alpha n^{\bar{\alpha}}} \text{Prun}_{\alpha,\alpha'}(T_\alpha(n); A_0, \dots, A_{k_0}) \xrightarrow[n \rightarrow \infty]{\text{a.s.}} \text{Prun}_{\alpha,\alpha'}(\mathcal{T}_\alpha; X_0, \dots, X_{k_0}), \tag{17}$$

in the Gromov–Hausdorff sense as  $n \rightarrow \infty$ . To see this, we will obviously use (16). Let us however emphasize here that the convergence (16) in the  $k + 1$ -pointed Gromov–Hausdorff sense *does not imply* the convergence of the pruned subtrees in general. Counter-examples can easily be set up when some of the marked points in the limiting tree are not leaves. However, here, the  $X_i$ 's are distinguished leaves of  $\mathcal{T}_\alpha$ , therefore the discrete labeled trees obtained by removing vertices of degree 2 (and glueing the adjacent edges) in  $\text{Span}(T_\alpha(n); A_0, \dots, A_{k_0})$ ,  $n \geq k_0$  and forgetting lengths on the edges of  $\text{Span}(\mathcal{T}_\alpha; X_0, \dots, X_{k_0})$  are all *identical*. From this and the definition of the pruning procedure (with, for reminder, the same sequence of uniform random variables  $U_i, i \geq 2$  in all cases), we get that the discrete labeled trees obtained by removing vertices of degree 2 in  $\text{Prun}_{\alpha,\alpha'}(T_\alpha(n); A_0, \dots, A_{k_0})$  for  $n$  large enough and forgetting lengths on the edges of  $\text{Prun}_{\alpha,\alpha'}(\mathcal{T}_\alpha; X_0, \dots, X_{k_0})$  are all identical. Re-incorporating



distances, the conclusion follows from the convergence of  $\text{Span}(T_\alpha(n); A_0, \dots, A_{k_0})$ , after normalization, towards  $\text{Span}(\mathcal{T}_\alpha; X_0, \dots, X_{k_0})$ , which is a consequence of (16), see e.g. Lemma 14 in [15].

With the notation of the proof of Proposition 2 we have thus proved that

$$\text{Prun}_{\alpha,\alpha'}(\mathcal{T}_\alpha; (X_i)_{i \geq 0}) = \mathfrak{T}_{\alpha,\alpha'} = M_{\alpha,\alpha'}^{\bar{\alpha}'} \cdot \mathcal{T}_{\alpha'}.$$

With a slight abuse of notation, we also denote by  $\mu_{\alpha'}$  the uniform mass measure on the tree  $\text{Prun}_{\alpha,\alpha'}(\mathcal{T}_\alpha; (X_i)_{i \geq 0})$ . Let then  $B_i, i \geq 0$  be the indices of the leaves that are kept in the pruning procedure applied to  $(\mathcal{T}_\alpha; (X_i)_{i \geq 0})$ . By the above proof, these indices also correspond to the indices of the leaves  $A_i, i \geq 0$  that end-up being blue in the construction of  $\mathbf{T}_{\alpha,\alpha'}$  from  $\mathbf{T}_\alpha$ . We then claim that conditionally on  $\text{Prun}_{\alpha,\alpha'}(\mathcal{T}_\alpha; (X_i)_{i \geq 0})$ ,

the leaves  $X_{B_i}, i \geq 0$  are i.i.d. according to  $\mu_{\alpha'}$ .

In words, the blue leaves of the pruning procedure form an i.i.d. sequence of leaves distributed according to the uniform mass measure on the remaining pruned subtree. Although this is a purely continuous setup statement we give a proof relying on a discrete argument. Indeed, it follows from the preceding proof that for every  $k \geq 0$

$$\left( \frac{T_{\alpha,\alpha'}(n)}{\alpha n^{\bar{\alpha}}}; A_{B_0}, \dots, A_{B_k} \right) \xrightarrow[n \rightarrow \infty]{\text{a.s.}} (\text{Prun}_{\alpha,\alpha'}(\mathcal{T}_\alpha; (X_i)_{i \geq 0}); X_{B_0}, \dots, X_{B_k}),$$

in the  $k + 1$ -pointed Gromov–Hausdorff sense. However, since  $\mathbf{T}_{\alpha,\alpha'}$  has the same distribution has a Marchal chain of parameter  $\alpha'$  time-changed by an independent increasing process, we deduce from Theorem 5(ii) that  $X_{B_0}, \dots, X_{B_k}$  are i.i.d. according to  $\mu_{\alpha'}$  conditionally on  $M_{\alpha,\alpha'}^{\bar{\alpha}'} \cdot \mathcal{T}_{\alpha'}$  as demanded.

### 5 A fragmentation point of view

In this section, we exploit the pruning operation of the last section in order to give another point of view on the construction of  $\mathfrak{T}_{\alpha,\alpha'}$  which is based on the fragmentation properties of the stable tree. Indeed, it is by now standard that the stable tree can be constructed from a certain self-similar fragmentation process whose conservative dislocation measure is explicitly known. We will show that we can trim this fragmentation process by throwing certain fragments away and such that the genealogy of the resulting *dissipative* fragmentation process is coded by the tree  $\mathfrak{T}_{\alpha,\alpha'}$ . This, finally, gives another interpretation of the random scaling appearing in Proposition 2 as the limit of the Malthusian martingale associated to the dissipative fragmentation.

#### 5.1 Fragmentation of the stable tree

**General self-similar fragmentation processes.** This paragraph is deliberately brief and we refer to Bertoin [5–7] for a detailed and rigorous introduction to the theory of

self-similar fragmentations. Let  $a \in \mathbb{R}$  and  $\nu$  be a sigma-finite measure on the set of decreasing sequences

$$\mathcal{S} = \left\{ \mathbf{s} = (s_1, s_2, \dots) : s_1 \geq s_2 \geq \dots \geq 0 \text{ and } \sum_{i \geq 1} s_i \leq 1 \right\}$$

(endowed with the topology of pointwise convergence), such that

$$\nu(1, 0, \dots) = 0 \quad \text{and} \quad \int_{\mathcal{S}} (1 - s_1) \nu(d\mathbf{s}) < \infty.$$

Such measure is called a *conservative dislocation measure* when  $\nu(\sum_i s_i < 1) = 0$  and a *dissipative dislocation measure* otherwise.

A pure-jump self-similar fragmentation process  $F$  with index of self-similarity  $a$  and dislocation measure  $\nu$  is a  $\mathcal{S}$ -valued càdlàg Markov process modelling the evolution of a system of splitting masses. The sequence  $F(t)$  represents the masses present at time  $t$ , ranked in the decreasing order. When  $\nu$  is finite the dynamic of the process is easy to describe:

- (i) different masses present at a given time evolve independently
- (ii) each mass  $m$  splits after a random time with an exponential distribution of parameter  $m^a \nu(\mathcal{S})$  to give sub-masses  $mS_1, mS_2, \dots$  where  $(S_1, S_2, \dots) \in \mathcal{S}$  is distributed according to  $\nu/\nu(\mathcal{S})$ , independently of the splitting time.

When  $\nu$  is infinite, the splitting times are dense in  $\mathbb{R}_+$  and the description of the process is more tricky. Informally, the item (i) is still valid, whereas the second item is replaced by the more general fact that a mass  $m$  splits in masses  $m\mathbf{s}$ ,  $\mathbf{s} \in \mathcal{S}$  at rate  $m^a \nu(d\mathbf{s})$ .

In all cases, this can be made rigorous by using Poisson point processes (see [4,5] for details). We start with the description of *homogeneous* fragmentations, the fragmentations with an index of self-similarity equal to 0. Let  $(S(t), k(t))_{t \geq 0}$  be a Poisson point process with intensity measure  $\nu \otimes \#$ , where  $\#$  denotes the counting measure on  $\mathbb{N}$ . Then it is possible to build a pure-jump càdlàg  $\mathcal{S}$ -valued process  $H$  such that  $H(0) = (1, 0, \dots)$  and  $H$  jumps only when an atom  $(S(t), k(t))$  of the Poisson point process occurs:  $H(t)$  is then obtained from  $H(t-)$  by keeping all elements of the sequence  $H(t-)$  but  $H_{k(t)}(t-)$  which is replaced by the sequence of masses  $H_{k(t)}(t-)S_i(t)$ ,  $i \geq 1$ , that is

$$H(t) = \downarrow \{ H_j(t-), H_{k(t)}(t-)S_i(t), i, j \geq 1, j \neq k(t) \},$$

where the symbol  $\downarrow$  means that we have ranked the terms in the decreasing order. This defines a fragmentation process  $H$  with parameters  $(0, \nu)$ , and all  $(0, \nu)$ -fragmentations can be constructed similarly from a Poisson point process.

Next, from a homogeneous fragmentation  $H$  with dislocation measure  $\nu$  and any  $a \in \mathbb{R}$ , one can construct a fragmentation  $F$  with parameters  $(a, \nu)$  by using a family of suitable time-changes. And vice-versa, any  $(a, \nu)$ -fragmentation can be turned into a homogeneous one with same dislocation measure by suitable time-changes. We do

not specify these time-changes here, as it requires some additional technical material, but refer to [6] for details. Let us just mention that the time-changes only depend on the past evolution of the masses. They modify the splitting times of the masses but neither change the masses themselves, nor the genealogical relations between them. It turns out that for all  $t \geq 0$  and all  $i \in \mathbb{N}$ , there exists a pair  $(s, j) \in [0, \infty) \times \mathbb{N}$  such that  $F_i(t) = H_j(s)$ , and vice-versa.

**Fragmentation trees.** A *rooted* continuous tree is a pair consisting of a continuous tree and one of its points, this distinguished point being called the *root*. To simplify notation we write  $\text{ht}(\cdot)$  for the distance to the root in a rooted tree.

In [16] it is proved that when  $a < 0$  and  $\nu$  is conservative, if  $F$  is an  $(a, \nu)$  fragmentation process then one can construct a random compact rooted continuous tree  $\mathcal{T}^\bullet$  endowed with a (random) probability measure  $\mu$  which is supported on its leaves such that  $\mathcal{T}^\bullet$  codes “the genealogy” of the process  $F$  in the sense that almost surely for all  $t \geq 0$  we have

$$F(t) = \downarrow \mu\text{-masses of the connected components of } \{v \in \mathcal{T}^\bullet : \text{ht}(v) > t\}. \tag{18}$$

The boundedness of this genealogical tree is due to a loss of mass by formation of dust in the fragmentation, which is itself due to the strictly negative index of self-similarity. The tree  $\mathcal{T}^\bullet$  is called the fragmentation tree associated with  $F$ . Recently, this result has been extended to the dissipative dislocation measures [26], the main difference being that the probability measure  $\mu$  is then no more concentrated on the set of leaves of the genealogical tree and also charges its skeleton. When  $\nu$  is dissipative and  $\nu(0, 0, \dots) = 0$  it is even fully supported by the skeleton.

A particular instance of fragmentation trees is given by the stable trees. Specifically, let  $\mathcal{T}_\alpha^\bullet$  denote an  $\alpha$ -stable tree endowed with its uniform mass measure  $\mu_\alpha$  and rooted at  $X_0$  (a uniform random leaf). If  $F_\alpha$  denotes the  $\mathcal{S}$ -valued process associated with  $\mathcal{T}_\alpha^\bullet$  by (18) then Bertoin [6] for the Brownian case and Miermont [22] for the general setting showed that  $F_\alpha$  is a pure-jump self-similar fragmentation process with self-similarity index  $-\bar{\alpha}$  and computed its dislocation measure. For the fragmentation  $F_2$  this measure is given for all test functions  $f : \mathcal{S} \rightarrow \mathbb{R}$  by

$$\int_{\mathcal{S}} f(\mathbf{s}) \nu_{\text{Br}}(d\mathbf{s}) = \int_{1/2}^1 f(x, 1-x, 0, \dots) \left(\pi x^3(1-x)^3\right)^{-1/2} dx,$$

and for the stable fragmentation  $F_\alpha$ ,  $1 < \alpha < 2$ , it is given by

$$\int_{\mathcal{S}} f(\mathbf{s}) \nu_\alpha(d\mathbf{s}) = C_\alpha \mathbb{E} \left[ \sigma_1 f \left( \frac{\Xi_i}{\sigma_1}, i \geq 1 \right) \right],$$

where

$$C_\alpha = \frac{\alpha(\alpha - 1)\Gamma(\bar{\alpha})}{\Gamma(2 - \alpha)} \tag{19}$$

and  $(\sigma_t, t \geq 0)$  is a stable subordinator of Laplace exponent  $\lambda^{1/\alpha}$  and  $(\Xi_i, i \geq 1)$  the sequence of its jumps before time 1, ranked in the decreasing order. Note that  $\nu_\alpha(\mathcal{S}) = \infty$  for all  $\alpha \in (1, 2]$ , which is related to the fact that the set of leaves of  $\mathcal{T}_\alpha$  is dense in  $\mathcal{T}_\alpha$ . Note also that  $\nu_\alpha$  is conservative.

We need some more notation. Note that for all  $t \geq 0$ , the connected components of  $\{v \in \mathcal{T}_\alpha^\bullet : \text{ht}(v) > t\}$  are open subtrees of  $\mathcal{T}_\alpha$ . We denote these subtrees by  $\mathcal{T}_\alpha^i(t), i \geq 1$  so that

$$F_\alpha^i(t) = \mu_\alpha(\mathcal{T}_\alpha^i(t)),$$

with  $F_\alpha(t) = (F_\alpha^i(t), i \geq 1)$ . These subtrees are all distributed as  $\mathcal{T}_\alpha \setminus \{X_0\}$ , up to random scalings. Specifically, conditionally on  $F_\alpha(t)$ , the trees  $\mathcal{T}_\alpha^i(t), i \geq 1$  are independent, such that for all  $i, \mathcal{T}_\alpha^i(t)$  endowed with the probability measure  $\mu_\alpha(\cdot | \mathcal{T}_\alpha^i(t))$  is distributed as  $F_\alpha^i(t)^{\bar{\alpha}} \cdot (\mathcal{T}_\alpha^i \setminus \{X_0^i\})$  endowed with  $\mu_\alpha^i$ , where  $(\mathcal{T}_\alpha^i, \mu_\alpha^i)$  is an independent copy of the  $\alpha$ -stable tree and  $X_0^i$  is a random point of  $\mathcal{T}_\alpha^i$  with distribution  $\mu_\alpha^i$ . This is called the self-similar property of the stable tree  $\mathcal{T}_\alpha$ . A similar property holds for all fragmentation trees.

### 5.2 The tree $\mathfrak{T}_{\alpha,\alpha'}$ is a dissipative fragmentation tree

From now on and in the remainder of this paper, we let  $1 < \alpha < \alpha' \leq 2$  and consider a stable tree  $\mathcal{T}_\alpha$  together with a sample of uniform leaves  $X_i, i \geq 0$ . We let

$$\mathfrak{T}_{\alpha,\alpha'} = \text{Prun}_{\alpha,\alpha'}(\mathcal{T}_\alpha, (X_i)_{i \geq 0})$$

and denote by  $\mathcal{T}_\alpha^\bullet$  and  $\mathfrak{T}_{\alpha,\alpha'}^\bullet$  the trees  $\mathcal{T}_\alpha$  and  $\mathfrak{T}_{\alpha,\alpha'}$  rooted at the point  $X_0$ . Since  $\mathfrak{T}_{\alpha,\alpha'}$  is a closed subtree of  $\mathcal{T}_\alpha$  for every  $x \in \mathcal{T}_\alpha$  there exists a unique point  $\text{Proj}(x) \in \mathfrak{T}_{\alpha,\alpha'}$  minimizing the distance to  $x$  in  $\mathfrak{T}_{\alpha,\alpha'}$ . We then consider the measure  $\mu_{\alpha,\alpha'}$  on  $\mathfrak{T}_{\alpha,\alpha'}$  obtained as the push-forward of  $\mu_\alpha$  by the application  $\text{Proj}(\cdot)$ .

**Proposition 13** *The probability  $\mu_{\alpha,\alpha'}$  is purely atomic and supported by the skeleton of  $\mathfrak{T}_{\alpha,\alpha'}$ .*

*Proof* If  $x \notin \mathfrak{T}_{\alpha,\alpha'}$  then a moment of thought using the definition of the pruning procedure shows that  $\text{Proj}(x)$  is a branch point of  $\mathfrak{T}_{\alpha,\alpha'}$ . Thus the image of the measure  $\mu_\alpha$  restricted to  $\mathcal{T}_\alpha \setminus \mathfrak{T}_{\alpha,\alpha'}$  is supported by the countable branch points of  $\mathcal{T}_\alpha$  that belong to  $\mathfrak{T}_{\alpha,\alpha'}$ . The statement of the proposition follows since  $\mu_\alpha(\mathcal{T}_\alpha \setminus \mathfrak{T}_{\alpha,\alpha'}) = 1$  by Proposition 10.

The goal of this section is to show that  $(\mathfrak{T}_{\alpha,\alpha'}^\bullet, \mu_{\alpha,\alpha'})$  is the genealogical tree of a dissipative self-similar fragmentation process. To make this precise, we start by constructing its dislocation measure  $\nu_{\alpha,\alpha'}$  from  $\nu_\alpha$ . For this, we design a procedure that extract randomly certain fragments of a dislocation of the unit mass.

Let  $\mathbf{s} = (s_1, s_2, \dots) \in \mathcal{S}$ . We first consider  $\mathbf{s}^* = (s_1^*, s_2^*, \dots)$  a size-biased random reordering of this sequence from which we will extract a subsequence  $(\widehat{s}_i^*)_{i \geq 1}$  as follows. When  $\alpha' = 2$  the extraction is deterministic: set  $s_1^* = s_1^*$  and  $s_2^* = s_2^*$  and

$\widehat{s}_n^* = 0$  for  $n \geq 3$ . For  $\alpha' < 2$ , we use an extra randomness and the definition of the probabilities  $p_{\alpha,\alpha',d,d'}$  defined by (1) in the Introduction. We proceed recursively on  $n$  as follows:

- (i)  $\widehat{s}_1^* = s_1^*$  and  $\widehat{s}_2^* = s_2^*$ ,
- (ii) for  $n \geq 3$ , let  $d = n$  and  $d' = \#\{1 \leq i \leq n - 1 : \widehat{s}_i^* = s_i^*\} + 1$  and put

$$\begin{aligned} \widehat{s}_n^* &= s_n^* && \text{with probability } p_{\alpha,\alpha',d,d'}, \\ \widehat{s}_n^* &= 0 && \text{otherwise.} \end{aligned}$$

The dislocation measure  $\nu_{\alpha,\alpha'}$  is then defined for all test functions  $f : \mathcal{S} \rightarrow \mathbb{R}$  by

$$\int_{\mathcal{S}} f(\mathbf{s}) \nu_{\alpha,\alpha'}(d\mathbf{s}) := \int_{\mathcal{S}} \nu_{\alpha}(d\mathbf{s}) \mathbb{E} \left[ f(\downarrow (s_i^*)) \right], \tag{20}$$

where we recall that the symbol  $\downarrow$  means reordering in decreasing order. Note that  $\nu_{\alpha,\alpha'}$  is dissipative since  $\alpha < \alpha'$ .

Our goal, now, is to prove the following result which clearly leads to Proposition 4. The proof does not rely on the particular expression of the dislocation measure of the stable tree and the result could be generalized to any fragmentation tree pruned by a similar procedure.

**Proposition 14** *The process  $F_{\alpha,\alpha'}$  obtained by considering for each  $t \geq 0$  the decreasing reordering of the*

$$\mu_{\alpha,\alpha'}\text{-masses of the connected components of } \{v \in \mathfrak{T}_{\alpha,\alpha'}^\bullet : ht(v) > t\}$$

*is a self-similar fragmentation, with index  $-\bar{\alpha}$  and dislocation measure  $\nu_{\alpha,\alpha'}$ .*

To be coherent with the notation introduced for the fragmentation  $F_{\alpha}$  of the stable tree  $\mathcal{T}_{\alpha}^\bullet$ , we denote by  $\mathcal{T}_{\alpha,\alpha'}^i(t)$ ,  $i \geq 1$  the open subtrees forming the connected components of  $\{v \in \mathfrak{T}_{\alpha,\alpha'}^\bullet : ht(v) > t\}$ , so that the elements of  $F_{\alpha,\alpha'}(t)$  satisfy for  $i \geq 1$ ,

$$F_{\alpha,\alpha'}^i(t) = \mu_{\alpha,\alpha'}(\mathcal{T}_{\alpha,\alpha'}^i(t)).$$

Note that for each  $i \geq 1$ , there exists a unique integer  $j (\geq i)$  such that  $\mathcal{T}_{\alpha,\alpha'}^i(t) \subset \mathcal{T}_{\alpha,\alpha'}^j(t)$ , and then that

$$F_{\alpha,\alpha'}^i(t) = \mu_{\alpha}(\mathcal{T}_{\alpha,\alpha'}^j(t)) = F_{\alpha,\alpha'}^j(t).$$

We then say that the mass  $F_{\alpha,\alpha'}^j(t)$  contributes to  $F_{\alpha,\alpha'}(t)$ . Note that  $F_{\alpha,\alpha'}^j(t)$  contributes to  $F_{\alpha,\alpha'}(t)$  if and only if  $\mathcal{T}_{\alpha,\alpha'}^j(t) \cap \mathfrak{T}_{\alpha,\alpha'} \neq \emptyset$ .

*Proof (Sketch)* A complete proof of the last proposition would be technical and we deliberately chose to stay rather informal. We first use the pruning construction of  $\mathfrak{T}_{\alpha,\alpha'}$  to give a practical criterion to decide which of the elements of  $F_\alpha(t)$  contribute to  $F_{\alpha,\alpha'}(t)$ . Let  $F_\alpha^i(t)$  be an element of  $F_\alpha(t)$  and consider the smallest  $k \geq 1$  such that  $X_k \in \mathcal{T}_\alpha^i(t)$ . Then  $F_\alpha^i(t)$  contributes to  $F_{\alpha,\alpha'}(t)$  if and only if  $X_k \in \text{Prun}_{\alpha,\alpha'}(\mathcal{T}_\alpha; X_0, \dots, X_k)$ , that is if  $X_k$  is a blue leaf, with the color coding of Sect. 4.1. Consider then a jump of the process  $F_\alpha$ : Pick a time  $t$  such that a fragment  $F_\alpha^i(t-)$  splits into a countable collection of masses

$$F_\alpha^j(t) \text{ for } j \in J = \{\text{indices offspring of } i \text{ at time } t\}.$$

In geometric terms, there exists a branch point at height  $t$  in  $\mathcal{T}_\alpha^\bullet$  that branches into the subtrees  $\mathcal{T}_\alpha^j(t)$ ,  $j \in J$ . Let us then assume that  $F_\alpha^i(t-)$  contributes to  $F_{\alpha,\alpha'}(t-)$ . This means that during the pruning operation the first leaf  $X_k$  falling into  $\mathcal{T}_\alpha^i(t-)$  is blue. Now, let us condition on that fact and on  $(F_\alpha(s), s \leq t)$  and set

$$s_j = \frac{\mu_\alpha(\mathcal{T}_\alpha^j(t))}{\mu_\alpha(\mathcal{T}_\alpha^i(t-))} \text{ for } j \in J.$$

Our goal is to show that the relative masses  $s_k$  of the fragments  $F_\alpha^k(t)$ ,  $k \in J$  that will contribute to  $F_{\alpha,\alpha'}(t)$  are distributed as  $(\widehat{s}_j^*)_{j \in J}$ . We denote by  $Y_k$ ,  $k \geq 1$  the leaves among  $X_k$ ,  $k \geq 1$  that belong to  $\mathcal{T}_\alpha^i(t-)$ . Clearly, on our conditioning, the  $(Y_k)_{k \geq 1}$  are i.i.d. sampled according to  $\mu_\alpha(\cdot \mid \mathcal{T}_\alpha^i(t-))$ . Thus, if  $\mathcal{T}_\alpha^{a_1}(t)$ ,  $\mathcal{T}_\alpha^{a_2}(t)$ ,  $\dots$  are the first, second,  $\dots$  subtrees into which one of the points  $Y_i$  falls then the sequence  $s_{a_1}, s_{a_2}, \dots$  is just a sized-biased ordering of the sequence  $(s_j)_{j \in J}$ ,

$$(s_{a_n})_{n \geq 1} = (s_j^*)_{j \in J}.$$

Moreover, for every  $n \geq 1$ , the first leaf that falls into  $\mathcal{T}_\alpha^{a_n}(t)$  is kept in the pruning construction (otherwise said is blue) with probability  $p_{\alpha,\alpha',d,d'}$ , where  $d = a_n$  and  $d' = \{k \leq n - 1 : \mathcal{T}_\alpha^{a_k}(t) \cap \mathfrak{T}_{\alpha,\alpha'} \neq \emptyset\} + 1$ . From this observation and the definition of  $(\widehat{s}_j^*)$  it should be clear that the  $s_k$ 's,  $k \in J$ , corresponding to offspring subtrees of  $\mathcal{T}_\alpha^i(t-)$  intersecting  $\mathfrak{T}_{\alpha,\alpha'}$  are thus distributed as  $\widehat{s}_j^*$ ,  $j \in J$  as desired. Obviously this procedure has to be done for *each* jump of the fragmentation process. It is then possible to conclude by using the Poisson point process construction of the fragmentation  $F_\alpha$ . One important point is to notice that the time-changes used to pass to the homogenous counterparts of  $F_\alpha$  and  $F_{\alpha,\alpha'}$  are identical in both fragmentations for each element of  $F_\alpha$  that contributes to  $F_{\alpha,\alpha'}$ . We leave subtleties to the careful reader.  $\square$

### 5.3 From the dissipative fragmentation to a randomized conservative fragmentation

Since the random tree  $\mathfrak{T}_{\alpha,\alpha'}$  is distributed as the stable tree  $\mathcal{T}_{\alpha'}$  (up to a random scaling) it supports a uniform mass measure distributed as  $\mu_{\alpha'}$ . However, the measure  $\mu_{\alpha,\alpha'}$  (see last section) does not charge the set of leaves of the tree  $\mathfrak{T}_{\alpha,\alpha'}$  but only its skeleton

and so is not connected, a priori, to  $\mu_{\alpha'}$ . We will see in this section how to construct from  $\mu_{\alpha, \alpha'}$  a measure on the leaves of  $\mathfrak{T}_{\alpha, \alpha'}$  that will be related to  $\mu_{\alpha'}$ .

**Keeping the mass.** Recall that the measure  $\nu_\alpha$  is conservative, that is the mass is conserved at each splitting. However there is a loss of mass in  $F_\alpha$  by formation of dust (due to the negative auto-similarity index). That is why, in this section, we rather focus on the homogeneous counterpart  $H_\alpha$  of the fragmentation  $F_\alpha$ , which is obtained by applying suitable time-changes in  $F_\alpha$ , see Sect. 5.1. In this case we have for all  $t \geq 0$

$$\sum_{i \geq 1} H_\alpha^i(t) = 1.$$

Obviously, this conservation of mass does not hold for homogeneous fragmentations with a dissipative dislocation measure. However there is then a natural way to define a random measure called the Malthusian measure which is stochastically conserved [7, 8, 26]. Although this new random measure is not a probability measure, it has a finite total mass with mean 1.

Let us now introduce rigorously this object. For this we use results established by Stephenson [26] for general dissipative fragmentation trees, following the ideas of Bertoin [7] for dissipative fragmentations with finite dislocation measures. From now on, we focus on the case of  $\nu_{\alpha, \alpha'}$  (see (20)), that is we consider  $H_{\alpha, \alpha'}$  the homogeneous counterpart of the process  $F_{\alpha, \alpha'}$ . The starting point is the existence of a Malthusian index  $p_{\alpha, \alpha'}^*$  which is characterized by

$$\int_S \left( 1 - \sum_{i \geq 1} s_i^{p_{\alpha, \alpha'}^*} \right) \nu_{\alpha, \alpha'}(ds) = 0.$$

See below for the discussion about existence and computation of this exponent. Once the existence of  $p_{\alpha, \alpha'}^*$  is granted, the process

$$\text{Mart}_{1,0}(s) = \sum_{i \geq 1} (H_{\alpha, \alpha'}^i(s))^{p_{\alpha, \alpha'}^*}, \quad s \geq 0$$

is a positive martingale, see e.g. [7, 26]. More generally, for each  $t \geq 0$  and  $i \in \mathbb{N}$ , the process

$$\text{Mart}_{i,t}(s) = \sum_{i \rightarrow j} (H_{\alpha, \alpha'}^j(s))^{p_{\alpha, \alpha'}^*}, \quad s \geq t, \tag{21}$$

where the sum is over all fragments whose ancestor at time  $t$  is  $H_{\alpha, \alpha'}^i(t)$ , is a positive càdlàg martingale which therefore converges almost surely towards a limiting value denoted by  $\text{Mart}_{i,t}$ ,  $i \in \mathbb{N}$ ,  $t \geq 0$ . Actually, one can check that

$$\mathbb{E} [\text{Mart}_{1,0}] = 1 \tag{22}$$

(this will be done during the proof of Theorem 15), which then ensures that the random variables  $\text{Mart}_{i,t}, i \in \mathbb{N}, t \geq 0$  can be chosen so that the convergences hold for *all*  $i, t$ , almost surely (see [26]). Besides, by the homogeneity property of the fragmentation process  $H_{\alpha,\alpha'}$ , we clearly have for all  $t \geq 0$ ,

$$(\text{Mart}_{i,t}, i \geq 1) = \left( (H_{\alpha,\alpha'}^i(t))^{p_{\alpha,\alpha'}^*} \text{Mart}_{1,0}^{(i,t)}, i \geq 0 \right),$$

where the r.v.  $\text{Mart}_{1,0}^{(i,t)}, i \geq 0$  are independent, all distributed as  $\text{Mart}_{1,0}$  and also independent of  $(H_{\alpha,\alpha'}^i(t), i \geq 1)$ . Finally, recall the notation  $\mathcal{T}_\alpha^i(t)$  and  $\mathcal{T}_{\alpha,\alpha'}^i(t), i \geq 1$  used in the last section for the open subtrees above level  $t$  in  $\mathcal{T}_\alpha^\bullet$  and  $\mathcal{T}_{\alpha,\alpha'}^\bullet$  respectively. We similarly denote by  $\mathcal{H}_\alpha^i(t)$  and  $\mathcal{H}_{\alpha,\alpha'}^i(t)$  the open subtrees of  $\mathcal{T}_\alpha^\bullet$  and  $\mathcal{T}_{\alpha,\alpha'}^\bullet$  so that

$$H_\alpha^i(t) = \mu_\alpha(\mathcal{H}_\alpha^i(t)), \quad H_{\alpha,\alpha'}^i(t) = \mu_{\alpha,\alpha'}(\mathcal{H}_{\alpha,\alpha'}^i(t)).$$

The set of subtrees  $\{\mathcal{H}_\alpha^i(t), i \in \mathbb{N}, t \geq 0\}$  is in bijection with the set of subtrees  $\{\mathcal{T}_\alpha^i(t), i \in \mathbb{N}, t \geq 0\}$ , and it is the same for the subtrees indexed by  $\alpha, \alpha'$ . We now have the material to build the Malthusian measure  $\mu_{\alpha,\alpha'}^*$ . On the event with probability 1 where all the martingales (21) converge, we set

$$\mu_{\alpha,\alpha'}^* \left( \mathcal{H}_{\alpha,\alpha'}^i(t) \right) = \text{Mart}_{i,t} \tag{23}$$

for all  $i \in \mathbb{N}$  and  $t \geq 0$ . By [26], this indeed defines uniquely a  $\sigma$ -finite measure  $\mu_{\alpha,\alpha'}^*$  on the tree  $\mathcal{T}_{\alpha,\alpha'}^\bullet$ , which is fully supported by the set of leaves. Note that the set  $\{\mathcal{H}_{\alpha,\alpha'}^i(t), i \in \mathbb{N}, t \geq 0\}$  generates the Borel  $\sigma$ -field on  $\mathcal{T}_{\alpha,\alpha'}^\bullet$ .

**Computing  $p_{\alpha,\alpha'}^*$ .** Apart from the Brownian case when  $\alpha' = 2$ , it is not easy to calculate directly from the expression of  $\nu_{\alpha,\alpha'}$  given in (20) its Malthusian index  $p_{\alpha,\alpha'}^*$ . However we can use information from the proof of Proposition 2, based on the coupling in Marchal’s construction, to get that

$$p_{\alpha,\alpha'}^* = \frac{\bar{\alpha}}{\bar{\alpha}'}$$

This will be done during the proof of Theorem 15. There is another indirect way to get the value of  $p_{\alpha,\alpha'}^*$ , provided we know that this index exists and that the measure  $\nu_{\alpha,\alpha'}$  satisfies suitable (weak) integrability properties. Indeed, on the one hand, the Hausdorff dimension of the  $\alpha'$ -stable tree, hence also that of  $\mathcal{T}_{\alpha,\alpha'}^\bullet$ , is  $1/\bar{\alpha}'$  a.s. [13, 16]. On the other hand, according to results of [26] on general dissipative fragmentation trees, we know that the Hausdorff dimension of the  $(-\bar{\alpha})$ -self-similar fragmentation tree  $\mathcal{T}_{\alpha,\alpha'}^\bullet$  is almost surely  $p_{\alpha,\alpha'}^*/\bar{\alpha}$ , provided  $\nu_{\alpha,\alpha'}$  satisfies the above mentioned integrability properties. After identification this gives back  $p_{\alpha,\alpha'}^* = \bar{\alpha}/\bar{\alpha}'$ .

The main result of this section is the fact that  $\mathcal{T}_{\alpha,\alpha'}^\bullet$  endowed with the measure  $\mu_{\alpha,\alpha'}^*$  is distributed as an  $\alpha'$ -stable tree endowed with its uniform mass measure, up to a random scaling. Recall from (10) the definition of a generalized Mittag-Leffler random variable  $\text{ML}_{\bar{\alpha}/\bar{\alpha}', \bar{\alpha}}$  with parameters  $(\bar{\alpha}/\bar{\alpha}', \bar{\alpha})$  and set  $M_{\alpha,\alpha'} = (\alpha'/\alpha)^{1/\bar{\alpha}'} \text{ML}_{\bar{\alpha}/\bar{\alpha}', \bar{\alpha}}$ .



**Theorem 15** *There is the following identity in distribution, with  $M_{\alpha,\alpha'}$  independent of  $(\mathcal{T}_{\alpha'}^\bullet, \mu_{\alpha'})$ ,*

$$\left(\mathfrak{T}_{\alpha,\alpha'}^\bullet, \mathbb{E}[M_{\alpha,\alpha'}] \mu_{\alpha,\alpha'}^*\right) \stackrel{(d)}{=} \left(M_{\alpha,\alpha'}^{\bar{\alpha}'} \mathcal{T}_{\alpha'}^\bullet, M_{\alpha,\alpha'} \mu_{\alpha'}\right),$$

where  $\mathbb{E}[M_{\alpha,\alpha'}] = (\alpha/\alpha')^{1/\bar{\alpha}'} \bar{\alpha}' \Gamma(\bar{\alpha}) / \Gamma(\bar{\alpha}(1 + 1/\bar{\alpha}'))$ , and where the measure  $\mu_{\alpha'}$  in the right-hand side of this identity actually represents the push-forward of  $\mu_{\alpha'}$  on  $\mathcal{T}_{\alpha'}$  by the homothety that sends  $\mathcal{T}_{\alpha'}$  on  $M_{\alpha,\alpha'}^{\bar{\alpha}'} \mathcal{T}_{\alpha'}$ .

In other words, the dissipative fragmentation tree  $\mathfrak{T}_{\alpha,\alpha'}^\bullet$ , endowed with a rescaled version of its Malthusian measure, is distributed as the genealogical tree of an  $\alpha'$ -stable fragmentation of the random mass  $M_{\alpha,\alpha'}$ . Note that the change of measures on the tree, from  $\mu_{\alpha,\alpha'}$  to  $\mu_{\alpha,\alpha'}^*$ , changes the index of self-similarity of the underlying fragmentation, which passes from  $-\bar{\alpha}$  to  $-\bar{\alpha}'$ .

*Proof of Theorem 15* This proof is divided into three steps. We recall that in addition to proving the identity in distribution of the statement, we also have to prove the existence of  $p_{\alpha,\alpha'}^*$ , as well as (22), so that the measure  $\mu_{\alpha,\alpha'}^*$  can indeed be defined via (23) on an event with probability one.

STEP 1. Recall that  $\mathfrak{T}_{\alpha,\alpha'} = \text{Prun}_{\alpha,\alpha'}(\mathcal{T}_\alpha, (X_i)_{i \geq 0})$  and that a leaf  $X_i$  is said to be blue if it belongs to  $\mathfrak{T}_{\alpha,\alpha'}$  and red otherwise. Write  $\mathbb{B}$  for the indices of the blues leaves and  $\mathbb{B}(n) = \mathbb{B} \cap \{0, 1, 2, \dots, n\}$ . Recalling from the proof of Theorem 3 that  $\mathbb{B}(n)$  is the number of leaves of a chain  $\mathbf{T}_{\alpha,\alpha'}$ , we apply Lemma 9 to get

$$\frac{\#\mathbb{B}(n)}{n^{\bar{\alpha}/\bar{\alpha}'}} \xrightarrow[n \rightarrow \infty]{\text{a.s.}} \text{ML}_{\bar{\alpha}/\bar{\alpha}', \bar{\alpha}}, \tag{24}$$

where  $\text{ML}_{\bar{\alpha}/\bar{\alpha}', \bar{\alpha}}$  has a generalized Mittag-Leffler distribution with parameters  $(\bar{\alpha}/\bar{\alpha}', \bar{\alpha})$ . In particular, its mean is

$$\mathbb{E}[\text{ML}_{\bar{\alpha}/\bar{\alpha}', \bar{\alpha}}] = \frac{\bar{\alpha}' \Gamma(\bar{\alpha})}{\Gamma(\bar{\alpha}(1 + 1/\bar{\alpha}'))},$$

using (10) and the relation  $\Gamma(a + 1) = a\Gamma(a)$  to get simplifications. Moreover, we know from (the proof of) Proposition 2 and the discussion at the end of Sect. 4.2 that there exists an  $\alpha'$ -stable tree  $\mathcal{T}_{\alpha'}$  such that

$$\mathfrak{T}_{\alpha,\alpha'} = \frac{\alpha'}{\alpha} \text{ML}_{\bar{\alpha}/\bar{\alpha}', \bar{\alpha}} \mathcal{T}_{\alpha'}$$

and that  $(X_i)_{i \in \mathbb{B}}$  is an i.i.d. sample of the uniform mass measure on (the leaves of)  $\mathfrak{T}_{\alpha,\alpha'}$ , which, with a slight abuse of notation, is also denoted by  $\mu_{\alpha'}$ . Consequently,

$$\mu_{\alpha'} \text{ is the almost sure empirical limit of } \frac{\sum_{i \in \mathbb{B}(n)} \delta_{X_i}}{\#\mathbb{B}(n)} \text{ as } n \rightarrow \infty, \tag{25}$$

which will be useful in the following. Theorem 15 will therefore be proved if we show the existence of  $p_{\alpha, \alpha'}^*$  and (22) and that almost surely for all  $i \in \mathbb{N}$ ,  $t \geq 0$ ,

$$ML_{\bar{\alpha}/\bar{\alpha}', \bar{\alpha}} \mu_{\alpha'} \left( \mathcal{H}_{\alpha, \alpha'}^i(t) \right) = \mathbb{E}[ML_{\bar{\alpha}/\bar{\alpha}', \bar{\alpha}}] \mu_{\alpha, \alpha'}^* \left( \mathcal{H}_{\alpha, \alpha'}^i(t) \right). \tag{26}$$

For this, we will use (24) and (25).

STEP 2. As for their self-similar counterparts,  $H_{\alpha, \alpha'}$  is a sub-fragmentation of  $H_\alpha$ . More precisely, for  $i \in \mathbb{N}$  and  $t \geq 0$ , there exists a unique integer  $j$  such that  $\mathcal{H}_{\alpha, \alpha'}^i(t) \subset \mathcal{H}_\alpha^j(t)$  and moreover,

$$H_{\alpha, \alpha'}^i(t) = \mu_{\alpha, \alpha'}(\mathcal{H}_{\alpha, \alpha'}^i(t)) = \mu_\alpha(\mathcal{H}_\alpha^j(t)) = H_\alpha^j(t).$$

Then, similarly to (24), introduce the random variable

$$ML_i(t) = \lim_{n \rightarrow \infty} \frac{\#\{k \in \mathbb{B}(n) : X_k \in \mathcal{H}_\alpha^j(t)\}}{\#\{k \leq n : X_k \in \mathcal{H}_\alpha^j(t)\}^{\bar{\alpha}/\bar{\alpha}'}}$$

which, by the self-similarity property of  $\mathcal{T}_\alpha$  and of the blue-coloring operation exists a.s., is distributed as  $ML_{\bar{\alpha}/\bar{\alpha}', \bar{\alpha}}$  and is independent of  $H_\alpha(t)$ . Moreover, the random variables  $ML_i(t)$ ,  $i \geq 1$  are independent. Now note the following easy but crucial fact: almost surely,

$$\begin{aligned} ML_{\bar{\alpha}/\bar{\alpha}', \bar{\alpha}} \mu_{\alpha'} \left( \mathcal{H}_{\alpha, \alpha'}^i(t) \right) &= \lim_{n \rightarrow \infty} \frac{\#\mathbb{B}(n)}{n^{\bar{\alpha}/\bar{\alpha}'}} \times \frac{\#\{k \in \mathbb{B}(n) : X_k \in \mathcal{H}_\alpha^j(t)\}}{\#\mathbb{B}(n)} \\ &= \lim_{n \rightarrow \infty} \frac{\#\{k \in \mathbb{B}(n) : X_k \in \mathcal{H}_\alpha^j(t)\}}{\#\{k \leq n : X_k \in \mathcal{H}_\alpha^j(t)\}^{\bar{\alpha}/\bar{\alpha}'}} \\ &\quad \times \left( \frac{\#\{k \leq n : X_k \in \mathcal{H}_\alpha^j(t)\}}{n} \right)^{\bar{\alpha}/\bar{\alpha}'} \\ &= ML_i(t) \left( H_\alpha^j(t) \right)^{\bar{\alpha}/\bar{\alpha}'} = ML_i(t) \left( H_{\alpha, \alpha'}^i(t) \right)^{\bar{\alpha}/\bar{\alpha}'}. \end{aligned} \tag{27}$$

Note also that a.s.

$$\sum_{i \geq 1} \mu_{\alpha'} \left( \mathcal{H}_{\alpha, \alpha'}^i(t) \right) = 1, \quad \forall t \geq 0.$$

Indeed, since  $H_\alpha$  is the homogeneous version of the fragmentation  $F_\alpha$ , the union  $\cup_i \mathcal{H}_\alpha^i(t)$  contains all the leaves of  $\mathcal{T}_\alpha$  but the root,  $\forall t \geq 0$  a.s. Consequently,  $\cup_i \mathcal{H}_{\alpha, \alpha'}^i(t)$  also contains all the leaves (except the root) of  $\mathfrak{T}_{\alpha, \alpha'}^\bullet$ , hence the last display. Together

with (27), this leads to

$$ML_{\bar{\alpha}/\bar{\alpha}', \bar{\alpha}} = \sum_{i \geq 1} ML_i(t) \left( H_{\alpha, \alpha'}^i(t) \right)^{\bar{\alpha}/\bar{\alpha}'}, \quad \forall t \geq 0 \text{ a.s.} \tag{28}$$

Taking expectations and recalling that  $0 < \mathbb{E}[ML_{\bar{\alpha}/\bar{\alpha}', \bar{\alpha}}] < \infty$ , this implies that

$$1 = \mathbb{E} \left[ \sum_{i \geq 1} \left( H_{\alpha, \alpha'}^i(t) \right)^{\bar{\alpha}/\bar{\alpha}'} \right] = \exp \left( -t \int_S \left( 1 - \sum_{i \geq 1} s_i^{\bar{\alpha}/\bar{\alpha}'} \right) \nu_{\alpha, \alpha'}(ds) \right),$$

using for the second equality that  $\nu_{\alpha, \alpha'}$  is the dislocation measure of  $H_{\alpha, \alpha'}$  and then the well-known fact in the theory of homogeneous fragmentations

$$\mathbb{E} \left[ \sum_{i \geq 1} \left( H_{\alpha, \alpha'}^i(t) \right)^q \right] = \exp \left( -t \int_S \left( 1 - \sum_{i \geq 1} s_i^q \right) \nu_{\alpha, \alpha'}(ds) \right), \quad \forall q \in \mathbb{R},$$

see e.g. Bertoin [5]. This shows that  $p_{\alpha, \alpha'}^*$  indeed exists and is equal to  $\bar{\alpha}/\bar{\alpha}'$ . In particular, we can now consider the martingales (21).

STEP 3. It remains to show (22) and then use (27) to get (26). Following an idea of [8] we get from (28) and the definition of the martingale  $(Mart_{1,0}(t), t \geq 0)$  that

$$\begin{aligned} & \mathbb{E} \left[ \left( ML_{\bar{\alpha}/\bar{\alpha}', \bar{\alpha}} - \mathbb{E}[ML_{\bar{\alpha}/\bar{\alpha}', \bar{\alpha}}] Mart_{1,0}(t) \right)^2 \right] \\ &= \mathbb{E} \left[ \left( \sum_{i \geq 1} \left( H_{\alpha, \alpha'}^i(t) \right)^{\bar{\alpha}/\bar{\alpha}'} (ML_i(t) - \mathbb{E}[ML_{\bar{\alpha}/\bar{\alpha}', \bar{\alpha}}]) \right)^2 \right] \\ &= \text{Var}(ML_{\bar{\alpha}/\bar{\alpha}', \bar{\alpha}}) \mathbb{E} \left[ \sum_{i \geq 1} \left( H_{\alpha, \alpha'}^i(t) \right)^{2\bar{\alpha}/\bar{\alpha}'} \right] \\ &= \text{Var}(ML_{\bar{\alpha}/\bar{\alpha}', \bar{\alpha}}) \exp \left( -t \int_S \left( 1 - \sum_{i \geq 1} s_i^{2\bar{\alpha}/\bar{\alpha}'} \right) \nu_{\alpha, \alpha'}(ds) \right), \end{aligned}$$

where the second equality comes from the fact that the random variables  $ML_i(t), i \geq 1$  are i.i.d distributed as  $ML_{\bar{\alpha}/\bar{\alpha}', \bar{\alpha}}$ , and independent of  $H_{\alpha}(t)$ . The function

$$q \mapsto \int_S \left( 1 - \sum_{i \geq 1} s_i^q \right) \nu_{\alpha, \alpha'}(ds)$$

is strictly increasing on its domain of definition and is equal to 0 for  $q = \bar{\alpha}/\bar{\alpha}'$ , as seen above. Hence, noticing that  $\text{Var}(\text{ML}_{\bar{\alpha}/\bar{\alpha}',\bar{\alpha}}) < \infty$ ,

$$\mathbb{E} \left[ (\text{ML}_{\bar{\alpha}/\bar{\alpha}',\bar{\alpha}} - \mathbb{E}[\text{ML}_{\bar{\alpha}/\bar{\alpha}',\bar{\alpha}}] \text{Mart}_{1,0}(t))^2 \right] \rightarrow 0 \text{ as } t \rightarrow \infty,$$

which implies the almost sure equality

$$\text{ML}_{\bar{\alpha}/\bar{\alpha}',\bar{\alpha}} = \mathbb{E}[\text{ML}_{\bar{\alpha}/\bar{\alpha}',\bar{\alpha}}] \text{Mart}_{1,0},$$

since moreover  $\text{Mart}_{1,0}(t) \rightarrow \text{Mart}_{1,0}$  a.s. Note that this implies the expected equality (22), hence the existence of  $\mu_{\alpha,\alpha'}^*$ . Similarly, by the homogeneity property of the fragmentation  $H_\alpha$ , for each  $(i, t)$ ,

$$\text{ML}_i(t) \left( H_{\alpha,\alpha'}^i(t) \right)^{\bar{\alpha}/\bar{\alpha}'} = \mathbb{E}[\text{ML}_{\bar{\alpha}/\bar{\alpha}',\bar{\alpha}}] \text{Mart}_{i,t}, \text{ a.s.},$$

which, together with (27), finally gives the expected

$$\text{ML}_{\bar{\alpha}/\bar{\alpha}',\bar{\alpha}} \mu_{\alpha,\alpha'} \left( \mathcal{H}_{\alpha,\alpha'}^i(t) \right) = \mathbb{E}[\text{ML}_{\bar{\alpha}/\bar{\alpha}',\bar{\alpha}}] \text{Mart}_{i,t} = \mathbb{E}[\text{ML}_{\bar{\alpha}/\bar{\alpha}',\bar{\alpha}}] \mu_{\alpha,\alpha'}^* \left( \mathcal{H}_{\alpha,\alpha'}^i(t) \right) \text{ a.s.}$$

Note that this identity is proved for each fixed  $(i, t)$  almost surely, and we actually want it almost surely for all  $(i, t) \in \mathbb{N} \times [0, \infty)$ . This is indeed true: for each  $(i, t)$ , consider a sequence  $(t_n)$  with  $t_n \in \mathbb{Q} \cap [0, \infty)$  decreasing towards  $t$ ; then one has

$$\mathcal{H}_{\alpha,\alpha'}^i(t) = \bigcup_{n \geq 1} \bigcup_{j \in J_n(i,t)} \mathcal{H}_{\alpha,\alpha'}^j(t_n)$$

where  $J_n(i, t)$  is the set of all indices  $j$  such that the mass  $H_{\alpha,\alpha'}^j(t_n)$  is a descendant of  $H_{\alpha,\alpha'}^i(t)$ . The above union on  $j \in J_n(i, t)$  is disjoint, whereas that on  $n$  is an increasing union. Consequently, a measure on  $\mathfrak{T}_{\alpha,\alpha'}$  is entirely characterized by the weights it assigns to the subtrees  $\mathcal{H}_{\alpha,\alpha'}^j(q)$ ,  $j \in \mathbb{N}$ ,  $q \in \mathbb{Q} \cap [0, \infty)$ . This ends the proof (and the paper).

**Acknowledgments** This work is partially supported by ANR-08-BLAN-0190 and ANR-08-BLAN-0220-01.

**References**

1. Abraham, R., Delmas, J.-F., Voisin, G.: Pruning a Lévy continuum random tree. *Electron. J. Probab* **15**(2010), 1429–1473 (2010)
2. Aldous, D.: The continuum random tree I. *Ann. Probab.* **19**, 1–28 (1991)
3. Aldous, D.: The continuum random tree III. *Ann. Probab.* **21**, 248–289 (1993)
4. Berestycki, J.: Ranked fragmentations, *ESAIM Probab. Statist.*, 6, pp. 157–175 (2002) (electronic, <http://www.emath.fr/ps/>)
5. Bertoin, J.: Homogeneous fragmentation processes. *Probab. Theory Relat. Fields* **121**, 301–318 (2001)
6. Bertoin, J.: Self-similar fragmentations. *Ann. Inst. H. Poincaré Probab. Statist.*, 38 (2002)

7. Bertoin, J.: Random fragmentations and coagulation processes. No. 102 in Cambridge studies in advanced mathematics, Cambridge University Press, Cambridge (2006)
8. Bertoin, J., Gnedin, A.: Asymptotic laws for nonconservative self-similar fragmentations. *Electron. J. Probab.* **9**, 575–593 (2004)
9. Bertoin, J., Le Gall, J.-F., Le Jan, Y.: Spatial branching processes and subordination. *Canad. J. Math.* **49**, 24–54 (1997)
10. Burago, D., Burago, Y., Ivanov, S.: A course in metric geometry. vol. 33 of graduate studies in Mathematics, American Mathematical Society, Providence, RI (2001)
11. Duquesne, T.: A limit theorem for the contour process of conditioned Galton-Watson trees. *Ann. Probab.* **31**, 996–1027 (2003)
12. Duquesne, T., Le Gall, J.-F.: Random trees, Lévy processes and spatial branching processes. *Astérisque*, pp. vi+147 (2002)
13. Duquesne, T., Le Gall, J.-F.: Probabilistic and fractal aspects of Lévy trees. *Probab. Theory Related Fields* **131**, 553–603 (2005)
14. Evans, S.N.: Probability and real trees, vol. 1920 of lecture notes in mathematics, Springer, Berlin, 2008. In: Proceedings of the lectures from the 35th summer school on probability theory held in Saint-Flour, 6–23 July 2005
15. Haas, B., Miermont, G.: Scaling limits of Markov branching trees, with applications to Galton-Watson and random unordered trees. *Ann. Probab.* **40**(6), 2589–2666 (2012)
16. Haas, B., Miermont, G.: The genealogy of self-similar fragmentations with negative index as a continuum random tree. *Electron. J. Probab.* **9**(4), 57–97 (2004) (electronic)
17. Haas, B., Miermont, G., Pitman, J., Winkel, M.: Continuum tree asymptotics of discrete fragmentations and applications to phylogenetic models. *Ann. Probab.* **36**, 1790–1837 (2008)
18. Haas, B., Pitman, J., Winkel, M.: Spinal partitions and invariance under re-rooting of continuum random trees. *Ann. Probab.* **37**, 1381–1411 (2009)
19. Le Gall, J.-F.: Random real trees. *Ann. Fac. Sci. Toulouse Math.* (6) **15**, 35–62 (2006)
20. Le Gall, J.-F., Le Jan, Y.: Branching processes in Lévy processes: the exploration process. *Ann. Probab.* **26**, 213–252 (1998)
21. Marchal, P.: A note on the fragmentation of a stable tree. In: Proceedings of the 15th colloquium on mathematics and computer science, discrete Math. Theor., AI, Assoc. Discrete Math. Theor. Comput. Sci., Nancy, pp. 489–499 (2008)
22. Miermont, G.: Self-similar fragmentations derived from the stable tree. I. Splitting at heights. *Probab. Theory Related Fields* **127**, 423–454 (2003)
23. Miermont, G.: Tessellations of random maps of arbitrary genus. *Ann. Sci. Éc. Norm. Supér.* (4) **42**, 725–781 (2009)
24. Pitman, J.: Combinatorial stochastic processes, vol. 1875 of lecture notes in mathematics, Springer, Berlin, 2006. Lectures from the 32nd summer school on probability Theory held in Saint-Flour, 7–24 July 2002
25. Rémy, J.-L.: Un procédé itératif de dénombrement d’arbres binaires et son application à leur génération aléatoire. *RAIRO Inform. Théor.* **19**, 179–195 (1985)
26. Stephenson, R.: General fragmentation trees (in preparation)