

Inverting the cut-tree transform

Louigi Addario-Berry^a, Daphné Dieuleveut^b and Christina Goldschmidt^c

^a*Department of Mathematics and Statistics, McGill University, 805 Sherbrooke Street West, Montréal, Québec, H3A 2K6, Canada.*

E-mail: louigi.addario@mcgill.ca, *url:* <http://www.problab.ca/louigi/>

^b*Equipe de Probabilités, Statistiques et Modélisation, Université Paris-Sud, Batiment 430, 91405 Orsay Cedex, France.*

E-mail: daphne.dieuleveut@normalesup.org

^c*Department of Statistics and Lady Margaret Hall, University of Oxford, 24-29 St Giles', Oxford OX1 3LB, UK.*

E-mail: goldschm@stats.ox.ac.uk, *url:* <http://www.stats.ox.ac.uk/goldschm>

Abstract. We consider fragmentations of an \mathbb{R} -tree T driven by cuts arriving according to a Poisson process on $T \times [0, \infty)$, where the first co-ordinate specifies the location of the cut and the second the time at which it occurs. The genealogy of such a fragmentation is encoded by the so-called *cut-tree*, which was introduced by Bertoin and Miermont [17] for a fragmentation of the Brownian continuum random tree. The cut-tree was generalised by Dieuleveut [25] to a fragmentation of the α -stable trees, $\alpha \in (1, 2)$, and by Broutin and Wang [19] to the inhomogeneous continuum random trees of Aldous and Pitman [11]. In the first two cases, the projections of the forest-valued fragmentation processes onto the sequence of masses of their constituent subtrees yield an important family of examples of Bertoin's self-similar fragmentations [14]; in the first case the time-reversal of the fragmentation gives an additive coalescent. Remarkably, in all of these cases, the law of the cut-tree is the same as that of the original \mathbb{R} -tree.

In this paper, we develop a clean general framework for the study of cut-trees of \mathbb{R} -trees. We then focus particularly on the problem of *reconstruction*: how to recover the original \mathbb{R} -tree from its cut-tree. This has been studied in the setting of the Brownian CRT by Broutin and Wang [20], who prove that it is possible to reconstruct the original tree in distribution. We describe an enrichment of the cut-tree transformation, which endows the cut-tree with information we call a *consistent collection of routings*. We show this procedure is well-defined under minimal conditions on the \mathbb{R} -trees. We then show that, for the case of the Brownian CRT and the α -stable trees with $\alpha \in (1, 2)$, the original tree and the Poisson process of cuts thereon can both be almost surely reconstructed from the enriched cut-trees. For the latter results, our methods make essential use of the self-similarity and re-rooting invariance of these trees.

Résumé. Nous considérons des fragmentations d'un \mathbb{R} -arbre T dirigées par des coupures qui arrivent selon un processus de Poisson sur $T \times [0, \infty)$, où la première composante désigne le point auquel se produit la coupure et le deuxième, l'instant auquel elle a lieu. La généalogie d'une telle fragmentation est codée par *l'arbre des coupes*, qui a été introduit par Bertoin et Miermont [17] pour une fragmentation de l'arbre brownien. L'arbre des coupes a ensuite été généralisé par Dieuleveut [25] à une fragmentation des arbres α -stables, pour $\alpha \in (1, 2)$, et par Broutin et Wang [19] aux arbres continus inhomogènes d'Aldous et Pitman [11]. Dans les deux premiers cas, l'évolution de la suite des masses des sous-arbres apparaissant dans le processus de fragmentation constitue une famille importante de processus de fragmentation auto-similaires de Bertoin [14]; dans le premier cas, une fois inversée dans le temps, la fragmentation devient un coalescent additif. Remarquablement, dans tous ces cas, la loi de l'arbre des coupes est la même que celle du \mathbb{R} -arbre initial.

Dans cet article, nous développons un cadre général pour l'étude de l'arbre des coupes d'un \mathbb{R} -arbre. Par la suite, nous nous concentrons particulièrement sur le problème de la *reconstruction* : comment retrouver le \mathbb{R} -arbre initial à partir de son arbre des coupes. Ce problème a été étudié pour l'arbre brownien par Broutin et Wang [20], qui démontrent qu'il est possible de reconstruire l'arbre initial en loi. Nous décrivons un enrichissement de la construction

de l'arbre des coupes, qui dote l'arbre des coupes d'une structure supplémentaire que nous appelons une *collection cohérente de routages*. Nous démontrons que ce procédé est bien défini sous des conditions minimales sur le \mathbb{R} -arbre. Ensuite, nous démontrons que pour l'arbre brownien ainsi que pour l'arbre α -stable avec $\alpha \in (1, 2)$, l'arbre initial muni de son processus de Poisson de coupures peut être reconstruit presque sûrement à partir de l'arbre des coupes enrichi. Pour ces derniers résultats, nous utilisons de façon essentielle l'auto-similarité et l'invariance par réenracinement de ces \mathbb{R} -arbres aléatoires.

Keywords: \mathbb{R} -tree, cut-tree, fragmentation.

1. Introduction

1.1. Cutting down trees

Consider a combinatorial tree T_n with vertices labelled by $1, 2, \dots, n$. A natural cutting operation on T_n consists of picking an edge $\{i, j\}$ uniformly at random and removing it, thus splitting the tree into two subtrees. Iterating on each of these subtrees, we obtain a discrete fragmentation process on the tree, which continues until the state has been reduced to a forest of isolated vertices.

A continuum analogue of this process has played an important role in the theory of coalescence and fragmentation. Let T be a Brownian continuum random tree and consider cuts arriving as a Poisson point process \mathcal{P} on $T \times [0, \infty)$ of intensity $\lambda \otimes dt$, where λ is the length measure on the skeleton of the tree. Careful definitions of these objects will be given below; for the moment, we simply note that λ is an infinite, but σ -finite measure, and that there is also a natural *probability* measure μ on T which allows us to assign masses to its subtrees. This Poisson cutting of T was first introduced and studied by Aldous and Pitman [9]. For $s \geq 0$, let $F(s) = (F_1(s), F_2(s), \dots)$ be the sequence of μ -masses of the connected components of $T \setminus \{p : (p, t) \in \mathcal{P}, t \leq s\}$, listed in decreasing order. Then $(F(s), s \geq 0)$ is an example of a self-similar fragmentation process, in the terminology of Bertoin [14]. Moreover, a time-reversal of this fragmentation process gives a construction of the standard additive coalescent (see [9] for more details).

The uniform cutting operation on trees described in the first paragraph was first considered in the mathematical literature in the early 1970's by Meir and Moon [44, 45]. They applied it repeatedly to the component containing a particular vertex (labelled 1, say) and investigated the number of cuts required to isolate that vertex. In [44] and [45], Meir and Moon focussed on cutting down two particular classes of random trees: uniform random trees, and random recursive trees. In both cases, these models possess a useful *self-similarity property*: the tree containing the vertex labelled 1 after the first cut is, conditioned on its size, again a tree chosen uniformly from the class in question.

This work spawned a line of research focussing primarily on the number of cuts, N_n , required to isolate the root in various models of random trees; see, for example, [26, 33, 40–42, 49]. Janson [42] considered the case where T_n is a Galton–Watson tree with critical offspring distribution of finite variance σ^2 , conditioned to have n vertices. (This includes the case where T_n is a uniform random tree, since this is equivalent, up to a random labelling, to taking the offspring distribution to be Poisson(1).) It is well known that the scaling limit of T_n in this case is the Brownian continuum random tree T [5–7]. Janson made the striking observation that $N_n/(\sigma\sqrt{n})$ converges in law to a Rayleigh distribution. The same limit holds for the rescaled distance between two uniformly chosen vertices in T_n , and is the law of the distance between two uniformly chosen points in the Brownian continuum random tree. It was later shown [3, 15, 17] that this common limit can be understood by using the cuts to couple T_n with a new tree T'_n in such a way that the number of cuts needed to isolate the vertex labelled 1 in T_n is the same as the distance between two uniformly-chosen vertices in T'_n , and where T'_n also converges to the Brownian continuum random tree when suitably rescaled. If T_n is a uniform random tree, this can, in fact, be done in such a way that T_n and T'_n have exactly the same distribution for each n , using a construction called the Markov chainsaw [3]. The Markov chainsaw takes the sequence of subtrees which are severed from the subtree containing the root, and glues them along a path; one obtains a tree T'_n which has the same distribution as T_n along with two marked points (the extremities of the new path) which are uniform random vertices of T'_n . An analogous construction can be performed in the continuum.

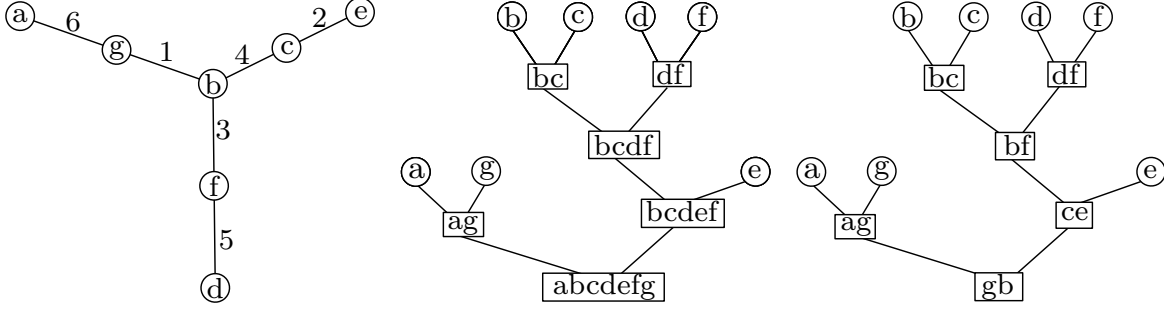


Fig 1: Left: A discrete tree T with vertex labels from $\{a, b, c, d, e, f, g\}$. Edges are marked with the time at which they are cut. Center: The resulting cut tree C . Each internal vertex corresponds to a cut, and is labelled by the block of the fragmentation process which is split by that cut. Right: Cuts correspond to edges in T ; here each internal vertex of C is labelled by the pair of endpoints of the corresponding edge.

1.2. Fragmentation and cut-trees

In this paper, we will focus on a construction which tracks the whole fragmentation, not just the cuts which affect the component of the root. Consider a discrete tree T_n repeatedly cut at uniformly chosen edges. The *cut-tree* C_n of T_n represents the genealogy of this fragmentation process. In this setting, C_n is a binary tree with $n - 1$ internal vertices and n leaves, where the leaves correspond to the vertices of T_n and the internal vertices to the non-singleton *blocks* (that is, the collections of labels of the subtrees) appearing at some stage of the fragmentation. The tree C_n is rooted at a vertex corresponding to $[n] := \{1, 2, \dots, n\}$, and the two children of the root are the two blocks into which the first cut splits $[n]$. More generally, for a non-singleton block $B \subset [n]$, the two children of B are the two blocks into which the next cut to arrive splits B .

The cut-tree was introduced by Bertoin in [15], where he considered the case in which T_n is a uniform random tree with n vertices (although the name “cut-tree” was first coined subsequently in [17]). The idea of using a tree to track the genealogy of a discrete fragmentation process also notably appears earlier in [38]. In [17], Bertoin and Miermont took T_n to be a Galton–Watson tree with critical offspring distribution of finite variance σ^2 , conditioned to have n vertices. Bertoin and Miermont proved the following remarkable result. View T_n (resp. C_n) as a measured metric space by taking $\sigma n^{-1/2}$ (resp. $\sigma^{-1} n^{-1/2}$) times the graph distance as the metric, and in both cases endowing the vertices with the uniform probability measure. Then the pair (T_n, C_n) converges in distribution as $n \rightarrow \infty$ (in the Gromov–Prokhorov sense) to a pair of dependent Brownian continuum random trees (T, C) . The second CRT is obtained from the first by a continuum analogue of the discrete cut-tree construction discussed above; we will describe this in detail (and in greater generality) below, once we have introduced the necessary notation.

Cut-trees have been considered for other models of random discrete trees, notably for random recursive trees in [13, 16]. In that setting, the tree T_n itself (when endowed with the graph distance) does not possess an interesting scaling limit, but the corresponding cut-tree C_n , thought of as a metric space using the graph distance divided by $n/\log n$, and endowed with the uniform probability measure on its leaves, converges in the Gromov–Hausdorff–Prokhorov sense to the unit interval endowed with the Lebesgue measure [16]. (We observe that there are minor variations in the way that discrete cut-trees are defined in the existing literature. We will gloss over these differences since our primary interest is in the continuous case, where there is no ambiguity of definition.)

The fragmentation of the Brownian continuum random tree via Poisson cutting was generalised to a fragmentation process of the α -stable trees T_α , $\alpha \in (1, 2)$, by Miermont in [47]. The stable trees are the scaling limits of Galton–Watson trees T_n with critical offspring distribution in the domain of attraction of an α -stable law, conditioned to have total progeny n . The heavy-tailed nature of the offspring distribution is reflected in the limit by the fact that the branchpoints (or *nodes*) of the tree all have infinite degree almost surely. Despite this, the nodes can be given a notion of size (which is made precise using a local time). In

order to obtain a self-similar fragmentation, it is necessary now for the cuts to occur at the nodes of T_α rather than along the skeleton. This is achieved by using a Poisson process whose intensity at a particular node is proportional to the “size” of that node. The cut-tree C_α corresponding to this fragmentation was introduced and studied in [25]. Again, it is the case that if T_n is a conditioned critical Galton–Watson tree with offspring distribution in the domain of attraction of an α -stable law, and if C_n is the corresponding discrete cut-tree (suitably adapted to take into account of cutting at vertices rather than edges), then (T_n, C_n) suitably rescaled converges (in the Gromov–Prokhorov sense) to the pair (T_α, C_α) , where T_α and C_α are (dependent) α -stable trees. In the sequel, we will often think of and refer to the Brownian CRT as the 2-stable tree.

Broutin and Wang [19] generalised the Brownian cut-tree in a different direction, to the inhomogeneous continuum random trees (ICRT’s) of Aldous and Pitman [11]. (These are the scaling limits of the so-called p -trees [10, 22].) Cutting an ICRT according to the points of a Poisson process, whose intensity measure is now a linear combination of the length measure on the skeleton and a measure on the nodes, again yields a sort of fragmentation process (although no longer, in general, a self-similar or even Markovian one). Broutin and Wang established that (a particular version of) the cut-tree of a p -tree is again a p -tree. They also showed that the cut-tree of a (discrete) p -tree converges to the continuum cut-tree of the scaling limit ICRT. Moreover, the cut-tree of an ICRT again has the same law as the original tree.

Abraham and Delmas [1], working in the continuum, proved an analogue of the Markov chainsaw result for the Lévy trees of Duquesne and Le Gall [27], which form the general family of scaling limits of conditioned Galton–Watson trees. In particular, they showed that there is a measure λ such that if one cuts the tree in a Poisson manner with intensity $\lambda \otimes dt$ and glues the trees which get separated from the component containing the root along a line-segment then, working under the excursion measure, one again obtains a Lévy tree with the same “law” as the original tree. We understand that the cut-tree of a Lévy tree is the subject of work in progress by Broutin and Wang.

1.3. A general framework, and reconstruction

In this paper, we work directly in the continuum, and establish a general framework for the study of the cut-tree of an \mathbb{R} -tree T , where the cutting occurs according to a Poisson random measure on $T \times [0, \infty)$ of intensity $\lambda \otimes dt$, and where λ is any measure on T satisfying certain natural conditions. This encompasses all of the examples which have previously been studied. In this general setting, we establish conditions under which it is possible to make sense of a unique cut-tree C . Under a compactness assumption, we are also able to define the push-forwards of probability measures on T to the cut-tree C , by studying the push-forwards of empirical measures in T .

Our main result concerns the problem of *reconstruction*, that is, recovering the original tree T from its cut-tree C . In the Brownian CRT setting, the paper [3] describes a partial cut-tree, in which only cuts of the component containing a root vertex are considered, and describes how to reconstruct T in distribution from this partial cut-tree. This result is generalised by Broutin and Wang in [19] to a partial reconstruction result for cut-trees of ICRTs, and in [20] to complete reconstruction in the case of the Brownian CRT. More precisely, in [20] they describe what they call a “shuffling” operation on trees. Writing $s(T)$ for the shuffling of the tree T , they show that the pair $(s(T), T)$ has the same law as the pair (T, C) , where C is the cut-tree of T . Thus, the shuffling operation is a *distributional* inverse of the cut-tree operation, for Brownian CRTs. However, the question of whether the original tree can be recovered almost surely was left open. One of the contributions of this paper is to establish that, indeed, α -stable trees can almost surely be reconstructed from their cut-trees, for all $\alpha \in (1, 2]$.

To state our theorems formally requires some technical set-up, which we defer to the next section. It is, however, instructive to consider the discrete reconstruction problem, since what we do in the continuum will be analogous and the discrete version is rather easier to visualise. We will focus on the situation where we cut at edges, so that the cut-tree is defined as at the start of Section 1.2. Then the extra information which is required in order to reconstruct T_n from C_n is precisely the set of labels of the edges in the original tree.

Earlier we thought of an internal vertex of C_n as representing a non-singleton block B of the fragmentation, where B contains all of the vertices labelling the leaves in the subtree above that internal vertex. We may equally think of such an internal vertex as corresponding to the edge $\{i, j\}$ which is cut and causes B to

fragment, and from this perspective it is natural to mark this vertex with the pair $\{i, j\}$. In the discrete setting, such markings provide enough information to recover T_n ; indeed, they fully specify the edge set of T_n , so reconstruction is trivial! However, this does not generalise to the \mathbb{R} -tree setting. A natural question, and one which we partly answer in this paper, is whether an analogue of “labelling by cut-edges” can be defined for the cut-trees of \mathbb{R} -trees; and, if so, whether such a labelling contains enough information to allow reconstruction, as in the discrete setting. In the next paragraph, we sketch the reconstruction procedure whose continuum analogue we develop in the sequel.

Suppose that we wish to recover the path between two vertices i and j . Then we may do so as follows. The subtrees containing i and j were separated by a cut which is represented in C_n by the most recent common ancestor (MRCA) of i and j ; call this node $i \wedge j$. The internal node $i \wedge j$ is marked with two labels, k and ℓ , where the vertex k lies in the subtree above $i \wedge j$ containing i and the vertex ℓ lies in the subtree above $i \wedge j$ containing j . We call the pair (k, ℓ) a *signpost* for i and j . So we now know that $\{k, \ell\}$ is an edge of T_n lying on the path between i and j . In order to recover the rest of the path, we need to determine the path between i and k and the path between ℓ and j . We do this by repeating the same procedure in the subtree above $i \wedge j$ containing i and k and in the subtree above $i \wedge j$ containing j and ℓ . So, for example, we find the MRCA of i and k , $i \wedge k$, and consider its marks, which tell us about the edge which was cut resulting in the separation of i and k into different subtrees. We continue this recursively in each subtree, stopping in a particular subtree only when both of the marks on the MRCA are those of vertices already observed on the path. We will call a *routing* the collection of signposts used in this process. Although somewhat cumbersome in the discrete setting, this procedure turns out to generalise nicely to the continuum, whereas the notion of edges does not.

1.4. Stable trees as fixed points

This work may be viewed in part as a contribution to the literature on transformations with stable trees as a fixed point. The articles [31, 32] were perhaps the first to explicitly take this perspective; motivated by problems from phylogenetics and algorithmic computational biology, they introduce *cutting* and *regrafting* operations on CRTs, and show that these operations have the Brownian CRT as fixed points. The main results of [3, 19, 25] state that the Brownian, inhomogeneous, and α -stable trees, respectively, are all fixed points of suitable cut-tree operations. We also mention the quite recent work of Albenque and the third author [4], which describes a CRT transformation for which the Brownian CRT is the *unique* fixed point, and furthermore shows that the fixed point is attractive. It would be interesting to establish analogous results for cut-tree operations.

1.5. Outline

We conclude this rather lengthy introduction with an outline of the remainder of the paper. Section 2 formally introduces some of the basic objects and random variables of study, including \mathbb{R} -trees and their marked and measured versions, and the α -stable trees. Section 3 presents our general construction of cut-trees of measured \mathbb{R} -trees, and of the routing information which we use for reconstruction.

In Section 4 we specialise our attention to stable trees, and show that almost sure reconstruction is possible in this case. In Section 4.2 we establish a fixed-point identity for size-biased Mittag-Leffler random variables. We use this identity in Section 4.3, together with an endogeneity result and a somewhat subtle martingale argument, to show that it is possible to almost surely reconstruct distances between two random points. We extend the reconstruction from two points to all points in Section 4.4.

Finally, Section 5 contains some more speculative remarks, and presents several questions and avenues for future research.

2. Trees and metric spaces

Fix a measurable space (S, \mathcal{S}) and a finite measure μ on (S, \mathcal{S}) . We write $X \sim \mu$ if X is an S -valued random variable with law $\mu/\mu(S)$. For $R \in \mathcal{S}$, we write $\mu_R = \mu(R)^{-1}\mu|_R$ for the restriction of μ to R , rescaled to form a probability measure.

2.1. \mathbb{R} -trees

We begin by recalling some standard definitions. For a metric space (M, d) and $S \subset M$, we write (S, d) as shorthand for the metric space $(S, d|_{S \times S})$.

Definition 1. A metric space (T, d) is an \mathbb{R} -tree if, for every $u, v \in T$:

- there exists a unique isometry $f_{u,v}$ from $[0, d(u, v)]$ into T such that $f_{u,v}(0) = u$ and $f_{u,v}(d(u, v)) = v$;
- for any continuous injective map $f : [0, 1] \rightarrow T$, such that $f(0) = u$ and $f(1) = v$, we have

$$f([0, 1]) = f_{u,v}([0, d(u, v)]) := \llbracket u, v \rrbracket.$$

A rooted \mathbb{R} -tree is an \mathbb{R} -tree (T, d, ρ) with a distinguished point ρ called the *root*.

Note that we do not require \mathbb{R} -trees to be compact.

Let $T = (T, d)$ be an \mathbb{R} -tree. The *degree* $\deg(x)$ of $x \in T$ is the number of connected components of $T \setminus \{x\}$. An element $x \in T$ is a *leaf* if it has degree 1; we write $\text{leaf}(T)$ for the set of leaves of T . An element $x \in T$ is a *branchpoint* if it has degree at least 3; we write $\text{br}(T)$ for the set of branchpoints of T . For $x, y \in T$, write $\llbracket x, y \rrbracket$ for $\llbracket x, y \rrbracket \setminus \{x, y\}$. The *skeleton* of T , denoted $\text{skel}(T)$, is the set $\bigcup_{x, y \in T} \llbracket x, y \rrbracket$ of vertices with degree at least two. We observe that the metric d gives rise to a *length measure* λ_d , supported by $\text{skel}(T)$, which is the unique σ -finite measure such that $\lambda_d(\llbracket u, v \rrbracket) = d(u, v)$.

A common way to encode a compact rooted \mathbb{R} -tree is via an excursion, that is, a continuous function $h : [0, 1] \rightarrow \mathbb{R}_+$ such that $h(0) = h(1) = 0$ and $h(x) > 0$ for $x \in (0, 1)$. For $x, y \in [0, 1]$, let

$$d(x, y) = h(x) + h(y) - 2 \min_{x \wedge y \leq z \leq x \vee y} h(z),$$

and define an equivalence relation \sim by declaring $x \sim y$ if $d(x, y) = 0$. Let $T = [0, 1] / \sim$. Then it can be checked that (T, d) is an \mathbb{R} -tree which may be naturally rooted at the equivalence class ρ of 0, and endowed with the measure μ which is the push-forward of the Lebesgue measure on $[0, 1]$ onto T .

A rooted \mathbb{R} -tree $T = (T, d, \rho)$ comes with a genealogical order \preceq such that $x \preceq y$ if and only if $x \in \llbracket \rho, y \rrbracket$; we say that x is an *ancestor* of y . (The symbol \prec then has the obvious meaning.) The *most recent common ancestor* $x \wedge y$ of $x, y \in T$ is the point of $\{z : z \preceq x, z \preceq y\}$ which maximises $d(\rho, z)$. For $z \in T$, write $T_z = \{x \in T : x \wedge z = z\}$, and $T_z = (T_z, d, z)$ for the subtree above z . Next, for $y \in T_z$ write $T_z^y = \{x \in T : x \notin \llbracket x, y \rrbracket\} \cup \{z\}$ and $T_z^y = (T_z^y, d, z)$ for the subtree rooted at z containing y .

A *measured* \mathbb{R} -tree is a triple (T, d, μ) , where (T, d) is an \mathbb{R} -tree and μ is a Borel probability measure on T . A *pointed* \mathbb{R} -tree is a triple (T, d, s) , where s is a finite or infinite sequence of elements of T . We combine adjectives in the natural way; thus, for example, a rooted measured pointed \mathbb{R} -tree is a quintuple (T, d, ρ, μ, s) .

2.2. Random \mathbb{R} -trees

The topological prerequisites for the study of random \mathbb{R} -trees have been addressed by several authors [2, 21, 31, 32, 34, 36, 37, 48]. Many of the aforementioned papers study random metric spaces more generally, but the theory specialises nicely to the setting of \mathbb{R} -trees, which is all we require in the current work. In this section we summarise the definitions and results which we require.

For the rest of the paper we restrict our attention to complete, locally compact \mathbb{R} -trees. Fix rooted measured pointed \mathbb{R} -trees $T = (T, d, \rho, \mu, s)$ and $T' = (T', d', \rho', \mu', s')$. We say T and T' are *isometric* if there exists a metric space isometry $\phi : T \rightarrow T'$ which sends ρ to ρ' , μ to μ' , and s to s' . More precisely, s and s' must have the same cardinality, and ϕ must satisfy the following:

- $\phi(\rho) = \rho'$;
- μ' is the pushforward of μ under ϕ ;
- writing s_k and s'_k for the k 'th elements of s and s' , respectively, then $\phi(s_k) = s'_k$ for all k .

We define isometry for less adjective-heavy \mathbb{R} -trees by relaxing the constraints on ϕ correspondingly. For example, measured \mathbb{R} -trees (T, d, μ) and (T', d', μ') are isometric if there exists a metric space isometry $\phi : T \rightarrow T'$ which sends μ to μ' .

Let \mathbb{T} be the set of isometry equivalence classes of (complete, locally compact) rooted measured \mathbb{R} -trees. Endowing \mathbb{T} with the Gromov–Hausdorff–Prokhorov (GHP) distance turns \mathbb{T} into a Polish space [2], which allows us to consider \mathbb{T} -valued random variables.

The GHP convergence theory for rooted measured pointed \mathbb{R} -trees with a finite number of marked points is described in Section 6 of [48]. In order to apply this theory in the current setting, two comments are in order. First, the theory is described for compact, rather than locally compact spaces. However, the development proceeds identically for locally compact spaces, so we omit the details. Second, in the present paper, we will in fact have a countably infinite number of marks. We briefly comment on the additional, rather standard, topological considerations. For each $1 \leq n \leq \infty$, fix $T_n = (T_n, d_n, \mu_n) \in \mathbb{T}$ and let $s_n = (s_{n,i}, i \geq 1)$ be a sequence of elements of T_n . We say the sequence of marked spaces (T_n, d_n, μ_n, s_n) converges to $(T_\infty, d_\infty, \mu_\infty, s_\infty)$ if, for each $m \in \mathbb{N}$, $(T_n, d_n, \mu_n, (s_{n,i}, i \leq m))$ converges to $(T_\infty, d_\infty, \mu_\infty, (s_{\infty,i}, i \leq m))$ in the sense described in [48].

We conclude by noting a sufficient condition for two \mathbb{R} -trees to have the same law. Let $T = (T, d, \rho, \mu)$ and $T' = (T', d', \rho', \mu')$ be \mathbb{T} -valued random variables such that μ and μ' both almost surely have full support. Conditional on T , let $(U_i)_{i \in \mathbb{N}}$ be a sequence of i.i.d. points of T with common law μ and, conditional on T' , let $(U'_i)_{i \in \mathbb{N}}$ be a sequence of i.i.d. points of T' with common law μ' . (See Section 6.5 of [48] for a treatment of the measurability issues involved in randomly sampling from a random metric space.) Letting $U_0 = \rho$ and $U'_0 = \rho'$, if

$$(d(U_i, U_j))_{i,j \geq 0} \stackrel{(d)}{=} (d'(U'_i, U'_j))_{i,j \geq 0},$$

then T and T' are identically distributed (see [37], Theorem 3 $\frac{1}{2}$.5).

2.3. The stable trees

Let $(X_t, t \geq 0)$ be a spectrally positive α -stable Lévy process with Laplace exponent $\mathbf{E}[\exp(-\lambda X_t)] = \exp(t\lambda^\alpha)$. A *stable tree* of index $\alpha \in (1, 2]$ is a random measured \mathbb{R} -tree $T = (T, d_T, \mu)$ encoded by an excursion of X of length 1. This normalisation of the stable tree agrees with that used by Duquesne and Le Gall [28]; see Section 3.3.

Equivalently, T is the scaling limit of large conditioned Galton–Watson trees whose offspring distribution is critical and lies in the domain of attraction of an α -stable law. As the scaling limit of a sequence of rooted random trees, T is naturally equipped with a root ρ . (The root ρ also arises as the equivalence class of 0 in the coding of T by the continuous height function associated with a length-1 excursion of $(X_t, t \geq 0)$ as discussed in Section 2.1.) The following theorem says that the root acts as a sample from the mass measure μ .

Theorem 2. *Let (T, d_T, ρ, μ) be a rooted stable tree of index $\alpha \in (1, 2]$. Let r be sampled from T according to μ . Then (T, d_T, r, μ) has the same (unconditional) distribution as (T, d_T, ρ, μ) .*

This follows from Aldous [5] for $\alpha = 2$ and Proposition 4.8 of Duquesne and Le Gall [28] for $\alpha \in (1, 2)$.

The stable tree of index $\alpha = 2$ corresponds to the Brownian continuum random tree encoded by $(\sqrt{2}\mathfrak{e}(t))_{0 \leq t \leq 1}$, where \mathfrak{e} denotes a normalised Brownian excursion. A significant difference between the Brownian CRT and the stable trees of index $\alpha \in (1, 2)$ is the fact that the Brownian CRT is almost surely binary (i.e. $\deg(b) = 3$ almost surely for all branchpoints b), whereas a stable tree of index $\alpha \in (1, 2)$ almost surely has only branchpoints of infinite degree. In the latter case, the “size” of a branchpoint $b \in \text{br}(T)$ can be described by the almost sure limit

$$A(b) = \lim_{\epsilon \rightarrow 0^+} \frac{1}{\epsilon} \mu\{v \in T : b \in [\rho, v], d_T(b, v) < \epsilon\},$$

whose existence was proved in [47] (see also [28]). It is useful to define a measure Λ on T , as follows. If $\alpha = 2$ then Λ is twice the length measure on $\text{skel}(T)$, and if $\alpha \in (1, 2)$ then $\Lambda = \sum_{b \in \text{br}(T)} A(b) \cdot \delta_b$.

The next theorem concerns the self-similarity of the stable trees, and will play an important role in Section 4. Let x, y, z be independent points of T with common law μ , and let b be the common branchpoint of x, y, z (i.e. the unique element of $\llbracket x, y \rrbracket \cap \llbracket x, z \rrbracket \cap \llbracket y, z \rrbracket$). Recall that T_b^x is the subtree rooted at b containing x , and write T^x and T^y for the measured \mathbb{R} -trees induced by $(T_b^x, \mu(T_b^x)^{-1+1/\alpha} \cdot d, b, \mu_{T_b^x})$ and $(T_b^y, \mu(T_b^y)^{-1+1/\alpha} \cdot d, b, \mu_{T_b^y})$ respectively.

Theorem 3. 1. *The trees T^x and T^y are independent α -stable trees, independent of the vector $(\mu(T_b^x), \mu(T_b^y))$.*
 2. *Conditionally on T^x (resp. T^y), the point x (resp. y) is sampled from T^x (resp. T^y) according to its rescaled mass measure.*

Proof. See Theorem 2 of Aldous [8] for the case $\alpha = 2$ and Corollary 10 of Haas, Pitman and Winkel [39] for $\alpha \in (1, 2)$. \square

We refer the reader to [27, 28, 46, 47] for more on the theory of stable trees.

3. The cut-tree of an \mathbb{R} -tree: general theory

3.1. Defining branch lengths for the cut-tree

Throughout Section 3 we let $T = (T, d_T, \mu)$ be an \mathbb{R} -tree with $\mu(T) = \mu(\text{leaf}(T)) = 1$. We further fix a σ -finite Borel measure λ on T with $\lambda(\text{leaf}(T)) = 0$ and such that $\lambda(\llbracket x, y \rrbracket) < \infty$ for all $x, y \in T$; this implies that λ is σ -finite. Let $\mathcal{P} = ((p_i, t_i), i \in I)$ be a Poisson point process on $T \times [0, \infty)$ with intensity measure $\lambda \otimes dt$.

We view each point p_i as a *cut*, which arrives at time t_i . For all $t \geq 0$ and $x \in T \setminus \{p_i : t_i \leq t\}$, let $T(x, t)$ be the connected component of $T \setminus \{p_i : t_i \leq t\}$ containing x , and let $T(x, t)$ be the corresponding \mathbb{R} -tree. For $x \in \{p_i : t_i \leq t\}$, let $T(x, t)$ be the subtree of T containing only the point x .

For points $x, y \in T$, let

$$t(x, y) = \inf\{t : \exists (p, t) \in \mathcal{P} \text{ with } p \in \llbracket x, y \rrbracket\};$$

if $t(x, y)$ is finite then let $(p(x, y), t(x, y))$ be the point of \mathcal{P} which first separates x and y . Next, for $S \subset T$, let $t(S) = \inf\{t(x, y) : x, y \in S\}$ be the first time a cut hits S , and write $t(x) = t(\{x\}) = t(x, x)$ for the first time x is hit.

Next, for $t \geq 0$, we define

$$\ell(x, t) = \int_0^{t \wedge t(x)} \mu(T(x, s)) ds,$$

and $\ell(x) = \ell(x, t(x))$. Then, for $x, y \in T$, let

$$D(x, y) = \int_{t(x, y)}^{t(x)} \mu(T(x, s)) ds + \int_{t(x, y)}^{t(y)} \mu(T(y, s)) ds.$$

If $t(x, y) = \infty$ this means that $D(x, y) = 0$. Note that if $\ell(x) < \infty$ and $\ell(y) < \infty$ then

$$D(x, y) = \ell(x) + \ell(y) - 2\ell(x, t(x, y)) = \ell(x) + \ell(y) - 2\ell(y, t(x, y)).$$

Proposition 4. *D is a pseudo-metric (possibly taking infinite values).*

Proof. Clearly $D(x, y) = D(y, x)$, so it suffices to verify that D satisfies the triangle inequality. Fix $x, y, z \in T$; we aim to show that $D(x, z) \leq D(x, y) + D(y, z)$. If $D(x, y) = \infty$ or $D(y, z) = \infty$ then there is nothing to prove, so we assume these quantities are finite. We have $t(x, z) \geq \min(t(x, y), t(y, z))$, so we henceforth assume without loss of generality that $t(x, z) \geq t(x, y)$ and $t(y, z) \geq t(x, y)$.

If $t(x, y) = \infty$ then $t(y, z) = t(x, z) = \infty$ and so $D(x, y) = D(y, z) = D(x, z) = 0$. If $t(x, y) < \infty$, using the fact that $T(y, s) = T(z, s)$ for $s \in [t(x, y), t(y, z))$, we get

$$\begin{aligned} D(x, y) + D(y, z) &\geq \int_{t(x, y)}^{t(x)} \mu(T(x, s)) ds + \int_{t(x, y)}^{t(y)} \mu(T(y, s)) ds + \int_{t(y, z)}^{t(z)} \mu(T(z, s)) ds \\ &\geq \int_{t(x, z)}^{t(x)} \mu(T(x, s)) ds + \int_{t(x, y)}^{t(y, z)} \mu(T(z, s)) ds + \int_{t(y, z)}^{t(z)} \mu(T(z, s)) ds \\ &\geq \int_{t(x, z)}^{t(x)} \mu(T(x, s)) ds + \int_{t(x, z)}^{t(z)} \mu(T(z, s)) ds = D(x, z), \end{aligned}$$

as required. \square

Note that $t(x, y)$ is exponentially distributed with parameter $\lambda(\llbracket x, y \rrbracket) < \infty$. Thus $t(x, y)$ is a.s. finite and so $D(x, y)$ is a.s. positive. However, it is possible that $D(x, y) = 0$ for some pairs x, y with $x \neq y$.

Now fix a sequence $u = (u_i, i \geq 1)$ of distinct points of T . The next proposition describes a tree encoding the genealogical structure that \mathcal{P} , viewed as a cutting (or fragmentation) process, induces on the elements of u .

Proposition 5. *Suppose that almost surely $\ell(u_i) < \infty$ for all $i \geq 1$. Then almost surely, up to isometry-equivalence there is a unique \mathbb{R} -tree $C^\circ = C^\circ(T, \mathcal{P}, u) := (C^\circ, d^\circ, \rho)$ containing points $\{\rho\} \cup \mathbb{N}$ and satisfying the following properties.*

1. *The points C° of C° satisfy $C^\circ = \bigcup_{i \in \mathbb{N}} \llbracket \rho, i \rrbracket$.*
2. *For all $i, j \in \mathbb{N}$, $d^\circ(\rho, i) = \ell(u_i)$ and $d^\circ(i, j) = D(u_i, u_j)$.*

Proof. First, the integrability condition $\ell(u_i) < \infty$ for all $i \geq 1$ implies that, almost surely, $d^\circ(i, j) < \infty$ for all $i, j \in \mathbb{N} \cup \{\rho\}$. For $n \in \mathbb{N}$ we write $[n] = \{1, \dots, n\}$.

We next show that for any metric Δ on $\mathbb{N} \cup \{\rho\}$, up to isometry there is at *most* one \mathbb{R} -tree containing \mathbb{N} whose restriction to \mathbb{N} is isometric to (\mathbb{N}, Δ) . Indeed, suppose $R = (R, d_R)$ and $R' = (R', d_{R'})$ are two such trees. Then for all $n \in \mathbb{N}$, the subtrees of R and R' induced by $\bigcup_{i, j \in [n] \cup \{\rho\}} \llbracket i, j \rrbracket$ are easily seen to be isometric. Further, any isometry between them induces an isometry of the subtrees spanned by $\bigcup_{i, j \in [k] \cup \{\rho\}} \llbracket i, j \rrbracket$, for any $k < n$. We may thus take a projective limit to obtain an isometry between R and R' .

It remains to prove existence, which in fact follows in much the same way once we verify that the distances specified by D are “tree-like”. More precisely, suppose that for each $n \in \mathbb{N}$, there exists an \mathbb{R} -tree $R_n = (R_n, d_n)$ containing $[n] \cup \{\rho\}$ such that $D(i, j) = d_n(i, j)$ for all $i, j \in [n] \cup \{\rho\}$ and such that $R_n = \bigcup_{i \leq n} \llbracket \rho, i \rrbracket$. Then a projective limit of the sequence R_n has the required properties.

Finally, for any $t \geq 0$, the collection of cuts $\{p_i : t_i \leq t\}$ induces a partition of $[n]$: for $j, k \in [n]$, j and k lie in the same part at time t if $t(u_j, u_k) > t$. This partition-valued process has an evident genealogical structure, and so describes a rooted discrete tree F with leaves $[n]$. In this picture, ρ is simply the root of F . Note that ρ has degree one; see Figure 2.

To each internal node v of F , let $L(v)$ be the set of leaves which are descendants of v . For each edge vw of F , with v an internal node and w a child of v , give vw length

$$\int_{t(\{u_j, j \in L(v)\})}^{t(\{u_j, j \in L(w)\})} \mu(T(u_i, s)) ds,$$

where i is any element of $L(v)$. Finally, the child of ρ is the unique internal node v with $L(v) = [n]$; let the

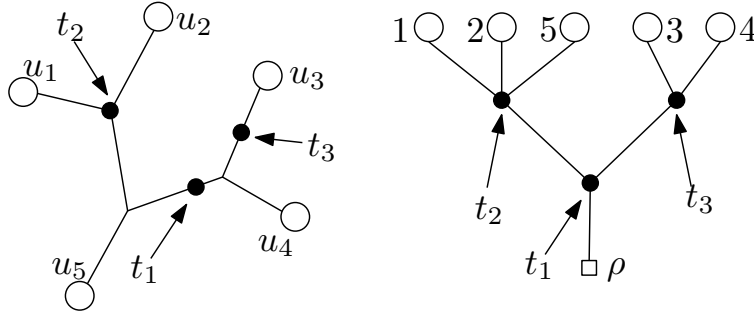


Fig 2: Left: the subtree of T spanned by vertices u_1, \dots, u_5 . Cuts on this subtree arrive at times $t_1 < t_2 < t_3$. Right: the resulting genealogical tree F with leaves $1, \dots, 5$. The figure indicates the correspondence between cuts and branchpoints of F .

edge ρv have length

$$\int_0^{t(\{u_j, j \in [n]\})} \mu(T(u_i, s)) ds,$$

where $i \in [n]$ is arbitrary. The resulting \mathbb{R} -tree has the correct distance between any pair $a, b \in [n] \cup \{\rho\}$, and is spanned by the paths between such pairs. (The only subtlety to note is that some branch lengths in this tree may be zero, if there exist points among $\{u_i, i \leq [n]\}$ which are cut.) This completes the proof of existence. \square

It deserves emphasis that the elements of \mathbb{N} are *random* points of the random tree C° : it is not possible to recover their locations from C° alone. In fact an analogue of the preceding proposition also holds in the setting where $\ell(u_i)$ may be infinite for some i . Though we do not require this case in the current work, its elaboration introduces several ideas we do use, so we now describe it. For $t \geq 0$, let

$$D_t(x, y) = \ell(x, t) + \ell(y, t) - 2\ell(x, t \wedge t(x, y)).$$

This is almost surely finite for any fixed $x, y \in T$, and $D_t(x, y) \uparrow D(x, y)$ as $t \rightarrow \infty$. Following the proof of Proposition 5 shows that there is a unique (up to isometry-equivalence) \mathbb{R} -tree $C_t = (C_t, d_t, \rho)$ satisfying the obvious modifications of conditions (1) and (2).

The trees $C_t = (C_t, d_t, \rho)$ are increasing in the sense that for $t < t'$, C_t may be realised as a subtree of $C_{t'}$. We may therefore define C° as the increasing limit of the process $(C_t, t \geq 0)$. This definition agrees with that of Proposition 5 when $\ell(u_i)$ is almost surely finite for all $i \in \mathbb{N}$. It additionally endows $\text{skel}(C^\circ)$ with a labelling by “arrival time”: for $x \in \text{skel}(C^\circ)$, let $\alpha(x) = \inf\{t : x \in C_t\}$.

To see that this is a measurable quantity, first note that for fixed $j \in \mathbb{N}$, the geodesic $[\rho, j]$ is isometric to the line segment $[0, \ell(u_j)]$. The function $m : [0, \infty) \rightarrow [0, 1)$ given by $m(s) = \mu(T(u_j, s))$ is clearly measurable. Thus, for $x \in [\rho, j]$, let

$$\alpha(x) = \inf \left\{ t : \int_0^t m(s) ds \geq d^\circ(\rho, x) \right\} = \inf \{ t : \ell(u_j, t) \geq d^\circ(\rho, x) \}. \quad (1)$$

It is easily seen that these labellings are consistent in that the label $\alpha(x)$ does not depend on the choice of j with $x \in [\rho, j]$.

Write (C, d, ρ) for the completion of $C^\circ = (C^\circ, d^\circ, \rho)$. Note that the elements of \mathbb{N} are points of C° and thus of (C, d, ρ) . For $n \in \mathbb{N}$, let $C(n) = (C(n), d, \rho)$ be the subtree of C° spanned by $\{\rho\} \cup [n]$, so having points $\bigcup_{j \leq n} [\rho, j]$. This is essentially the tree R_n from within the proof of Proposition 5.

3.2. Measures on the cut-tree

We next define, for each $n \in \mathbb{N}$, a measure ν_n on (C, d, ρ) whose support is $C(n)$. Let

$$\mathcal{P}_n = \{(p, t) \in \mathcal{P} : \exists j \leq n \text{ such that } T(u_j, t) \neq T(u_j, t-)\}.$$

This is the set of points whose cuts fall within a subtree containing some point $\{u_j, j \leq n\}$. For $s \geq 0$, let $\mathcal{P}_n(s) = \{(p, t) \in \mathcal{P}_n : t \leq s\}$. We likewise define $\mathcal{P}(s) = \{(p, t) \in \mathcal{P} : t \leq s\}$.

For a set $S \subset T$, we say a connected component T' of $T \setminus S$ is *trivial* if $|T'| = 1$; otherwise it is *nontrivial*. For the remainder of Section 3 we assume that the fragmentation induced by \mathcal{P} conserves mass in that almost surely, for all $s > 0$

$$\sum_{\substack{T' \text{ a nontrivial component} \\ \text{of } T \setminus \{p : (p, t) \in \mathcal{P}(s)\}}} \mu(T') = 1. \quad (2)$$

Note that the preceding summation is over a countable collection of components, since any nontrivial component contains an interval of positive length. We also assume for the remainder of the section that $\lambda(\llbracket x, y \rrbracket) > 0$ for all distinct $x, y \in T$.

Denote the set of nontrivial connected components of $T \setminus \{p : (p, t) \in \mathcal{P}_n\}$ by $\{T_i : i \in I_n\}$.

Proposition 6. *Almost surely no component of $\{T_i, i \in I_n\}$ contains an element of $\{u_j, j \leq n\}$.*

Proof. Fix a countable set $S \subset T \setminus \{u_j, j \leq n\}$ with the property that for all distinct $x, y \in T$, $|S \cap \llbracket x, y \rrbracket| > 0$. Then for all $i \in I_n$ and $j \leq n$,

$$\begin{aligned} \mathbf{P}\{u_j \in T_i\} &\leq \sum_{s \in S} \mathbf{P}\{\llbracket u_j, s \rrbracket \subset T_i\} \\ &\leq \sum_{s \in S} \mathbf{P}\{t(\llbracket u_j, s \rrbracket) = \infty\} = 0, \end{aligned}$$

since $\lambda(\llbracket u_j, s \rrbracket) > 0$ for all $s \in S$. □

For each $i \in I_n$, let

$$\sigma_i = \inf\{s \geq 0 : T_i \text{ is a nontrivial connected component of } T \setminus \{p : (p, t) \in \mathcal{P}_n(s)\}\}$$

be the creation time of T_i . Let

$$m_i = \min\{j \leq n : u_j \text{ and } T_i \text{ lie in the same component of } T \setminus \{p : (p, t) \in \mathcal{P}_n(\sigma_i-)\}\}$$

be the index of the last point from $\{u_j, j \leq n\}$ to separate from T_i , breaking ties by taking the smallest such. Then let x_i be the unique point of $C(n)$ on $\llbracket \rho, m_i \rrbracket$ satisfying $\alpha(x_i) = \sigma_i$.

Proposition 7. *Suppose that $\mu(T(u_j, s)) \rightarrow 0$ as $s \rightarrow \infty$ for all $j \geq 1$, almost surely. Then for all $n \geq 1$, almost surely*

$$\sum_{i \in I_n} \mu(T_i) = 1.$$

Proof. For $s > 0$, write $\{T_{i,s}, i \in I_{n,s}\}$ for the set of nontrivial components of $T \setminus \{p : (p, t) \in \mathcal{P}_n(s)\}$. Each nontrivial component of $T \setminus \{p : (p, t) \in \mathcal{P}(s)\}$ is contained within an element of $\{T_{i,s}, i \in I_{n,s}\}$; together with (2), this implies $\sum_{i \in I_{n,s}} \mu(T_{i,s}) = 1$.

Now note that the only components of $\{T_{i,s}, i \in I_{n,s}\}$ which are not components of $\{T_i, i \in I_n\}$ are those containing points from $\{u_j, j \leq n\}$. Therefore,

$$\begin{aligned} \sum_{i \in I_n} \mu(T_i) &\geq \sum_{i \in I_{n,s}} \mu(T_{i,s}) - \sum_{j \leq n} \mu(T(u_j, s)) \\ &= 1 - \sum_{j \leq n} \mu(T(u_j, s)). \end{aligned}$$

The final sum tends to zero as $s \rightarrow \infty$ by assumption. \square

Now define a measure ν_n on (C, d, ρ) with support $C(n)$ by

$$\nu_n = \sum_{i \in I_n} \mu(T_i) \delta_{x_i}.$$

We may view the tree T_i as “frozen” at time σ_i and attached to $C(n)$ at point x_i . With this perspective, ν_n is obtained by projecting the masses of the frozen subtrees onto their attachment points in $C(n)$. We do not explicitly need this construction, however, so do not formalise it. The preceding proposition gives a sufficient condition for ν_n to be a probability measure almost surely.

Proposition 8. *If (C, d, ρ) is compact then ν_n is almost surely a Cauchy sequence in the space of Borel measures on C , so has a weak limit ν .*

Proof. Fix $m > n$ and write $\nu_n = \sum_{i \in I_n} \mu(T_i) \delta_{x_i}$ as above. Note that \mathcal{P}_n is increasing in n , so we may view $\{T_j, j \in I_m\}$ as a “refinement” of $\{T_i, i \in I_n\}$ in the sense that each tree T_i in the latter set is split by the cuts associated with points of $\mathcal{P}_m \setminus \mathcal{P}_n$ into a collection of trees $\{T_{i,j}, j \in J_i\}$ which all lie in the former. Furthermore, we exhaust $\{T_j, j \in I_m\}$ in this manner, in that $I_m = \bigcup_{i \in I_n} J_i$. Finally, the compactness of C implies that $\ell(u_j) < \infty$ for all $j \geq 1$, so $\mu(T(u_j, s)) \rightarrow 0$ as $s \rightarrow \infty$. By Proposition 7, it follows that $\sum_{i \in I_n} \mu(T_i) = 1$ for all n , from which it is clear that we also have $\mu(T_i) = \sum_{j \in J_i} \mu(T_{i,j})$.

We write

$$\nu_m = \sum_{i \in I_n} \sum_{j \in J_i} \mu(T_{i,j}) \delta_{x_{i,j}}$$

where, by analogy with the above, $x_{i,j}$ is the point of attachment of $T_{i,j}$ to $C(m)$. Now note that $x_{i,j} \in C_{x_i}$ by construction. (Recall that C_{x_i} is the subtree of C rooted at x_i .) Since the fragmentation conserves mass, it follows that for all $i \in I_n$,

$$\mu(T_i) = \sum_{x_{i,j} \in C_{x_i}} \mu(T_{i,j}).$$

We may thus obtain ν_n from ν_m by projecting all mass of ν_m onto the closest point of $C(n)$. Writing d_P and d_H for the Prokhorov and Hausdorff distances respectively, it follows that $d_P(\nu_n, \nu_m) \leq d_H(C(n), C(m))$. Since $C(m)$ contains $C(n)$ and is increasing in m , this implies that $d_P(\nu_n, \nu_m) \leq d_H(C(n), C)$, and the final quantity tends to zero as $n \rightarrow \infty$ by compactness. \square

In the case when (C, d, ρ) is compact, we write $C = (C, d, \rho, \nu)$, and call $C = C(T, \mathcal{P}, u)$ the cut-tree of (T, \mathcal{P}, u) , or sometimes simply the cut-tree of T .

The tree T comes endowed with measure μ . If it happens that

$$\mu = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j \leq n} \delta_{u_j},$$

then we say that μ is the empirical measure of the sequence u . In particular, if the elements of u are i.i.d. with law μ then this holds by the Glivenko-Cantelli theorem.

Proposition 9. *If μ is the empirical measure of u and (C, d, ρ) is compact then almost surely ν is the empirical measure of $\mathbb{N} \subset C$, i.e.*

$$\nu = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j \leq n} \delta_j.$$

Proof. For $n < m$ and for $i \in I_n$ let

$$\hat{\mu}_{n,m}(T_i) = \frac{1}{m-n} \# \{n < j \leq m : u_j \in T_i\}.$$

Since μ is the empirical measure of u , for fixed n we have

$$\lim_{m \rightarrow \infty} \hat{\mu}_{n,m}(T_i) = \lim_{m \rightarrow \infty} \frac{1}{m} \# \{1 \leq j \leq m : u_j \in T_i\} = \mu(T_i).$$

Now let $\nu_{n,m} = \sum_{i \in I_n} \hat{\mu}_{n,m}(T_i) \delta_{x_i}$. Since $\sum_{i \in I_n} \hat{\mu}_{n,m}(T_i) = \sum_{i \in I_n} \mu(T_i) = 1$, it follows that

$$\nu_n = \lim_{m \rightarrow \infty} \nu_{n,m}.$$

Finally, writing

$$\hat{\nu}_{n,m} = \frac{1}{m-n} \sum_{n < j \leq m} \delta_j,$$

we have

$$\lim_{n \rightarrow \infty} \sup_{m > n} d_P(\hat{\nu}_{n,m}, \nu_{n,m}) \leq \lim_{n \rightarrow \infty} d_H(C(n), C) = 0,$$

the last equality holding by compactness. This yields

$$\lim_{n \rightarrow \infty} \nu_n = \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \nu_{n,m} = \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \hat{\nu}_{n,m} = \lim_{m \rightarrow \infty} \frac{1}{m} \sum_{1 \leq j \leq m} \delta_j,$$

as required. □

3.3. Images in C of points of T .

In this subsection we make the following assumptions on T , the intensity measure $\lambda \otimes dt$ and the sequence u of points of T .

1. Almost surely, $\ell(u_i) < \infty$ for all $i \geq 1$. (By Proposition 5 we may then define $C^\circ(T, \mathcal{P}, u)$ and its completion (C, d, ρ) .)
2. The cut-tree (C, d, ρ) is compact.
3. The set $\{u_i : i \geq 1\}$ is dense in T .

For $k \in \mathbb{N}$ it is natural to associate the point $u_k \in T$ with the point $k \in C$; we call k an *image* of u_k . We now define the image (or *images*) in C of a fixed point $x \in T$.

First suppose that x is not one of the cut points p_i ; this holds a.s. precisely if $\lambda(\{x\}) = 0$. Fix a sequence $(i_k, k \geq 1)$ such that $u_{i_k} \rightarrow x$ as $k \rightarrow \infty$. We aim to prove that $(i_k, k \geq 1)$ is convergent when viewed as a sequence of points of C .

We first prove that a particular subsequence $(j_k, k \geq 1)$ is convergent; the subsequence is defined as follows. Let $j_1 = i_1$. Recalling that $t(x, y)$ is the first separation time of x and y , for $k \geq 1$ let

$$j_{k+1} = \min \{i_\ell : \ell \geq 1, u_{i_\ell} \in T(x, t(x, u_{j_k}))\}.$$

In words, $u_{j_{k+1}}$ is the first point in the sequence which is remains in the same component as x just after u_{j_k} separates from x . The subtree of C spanned by $\{\rho\} \cup \{j_k, k \geq 1\}$ consists of a single path P , hanging from which there is a distinct branch leading to each of the j_k . The branches are attached in the order given by the sequence. Note that the full subtree of C hanging from the path and containing j_k is simply $C_{j_k \wedge j_{k+1}}^{j_k}$.

One endpoint of the path P is ρ ; call the other x' . The point x' will be the image of x in C . By the definition of P , the sequence of branchpoints $(j_k \wedge j_{k+1}, k \geq 1)$ must converge to x' . Since C is compact, it must also hold that $(j_k, k \geq 1)$ converges to x' .

Turning now to the full sequence, observe that since $u_{i_\ell} \rightarrow x$ as $\ell \rightarrow \infty$, for any $k \geq 1$, only finitely many points of the sequence $(u_{i_\ell}, \ell \geq 1)$ lie within the subtree $C_{j_k \wedge j_{k+1}}^{j_k}$. Finally, compactness implies that the diameters of these subtrees tend to zero with k ; it follows that $i_\ell \rightarrow x'$ as $\ell \rightarrow \infty$.

It is immediate from the construction that

$$d(\rho, x') = \ell(x) = \int_0^\infty \mu(T(x, s)) ds < \infty.$$

Likewise, for all $k \geq 1$, $d(x', k) = \ell(x) + \ell(u_k) - 2\ell(x, t(x, u_k))$. This is important as, letting $u' = (x, u_1, u_2, \dots)$, it implies that we may view $C^\circ(T, \mathcal{P}, u')$ as a subtree of (C, d, ρ) . With this perspective, (C, d, ρ) is also the completion of $C^\circ(T, \mathcal{P}, u')$.

Next suppose x is one of the cut points, so $(x, t(x)) \in \mathcal{P}$. In this case, x has multiple images in C . There is one image which is a branchpoint, which is easily described as follows. Let u_i and u_j be distinct points such that $t(u_i, u_j) = t(x)$. Then the required image in C is $i \wedge j$.

Additionally, there is an image of x corresponding to each connected component of $T \setminus \{x\}$. We next describe how to find the image of x corresponding to a fixed such component \hat{T} . The idea is to view the cut-tree process as acting on \hat{T} , starting at time $t(x)$, and thereby identify the image of x . The key point is that the cut-tree behaves nicely under restriction, in a sense we now explain. Let

$$\hat{\mathcal{P}} = \{(p, t) \in \mathcal{P} : p \in T(x, t(x)-) \cap \hat{T}, t > t(x)\}.$$

The set $T(x, t(x)-) \cap \hat{T}$ is the subtree separated at time $t(x)$ which is contained in \hat{T} , and $\hat{\mathcal{P}}$ is the set of Poisson points falling in this subtree after time $t(x)$.

Next, let v be the set of points $u_i \in T(x, t(x)-) \cap \hat{T}$; for concreteness we list these in increasing order of index as $v = (v_i, i \geq 1)$. Let $T' = \{x\} \cup (T(x, t(x)-) \cap \hat{T})$, and let $T' = (T', d_{T'})$ be the subtree of T induced by T' .

Let $\mathcal{P}' = \{(p, t - t(x)) : (p, t) \in \hat{\mathcal{P}}\}$, and observe that \mathcal{P}' is a Poisson process on $T' \times [0, \infty)$ with intensity measure $\lambda|_{T' \setminus \{x\}} \otimes dt$. The completion of the tree $C^\circ(T', \mathcal{P}', v)$ is now isometric to a subtree C' of (C, d, ρ) . Furthermore, this isometry is uniquely specified by requiring that the images of the points of v agree with the images of the corresponding elements of the sequence u . Since x is not hit by \mathcal{P}' , it also has an image x' in $C^\circ(T', \mathcal{P}', v)$, which we view as contained in $C^\circ(T, \mathcal{P}, u)$ using the isometry just described.

It is important that x' depends on the choice of a component \hat{T} of $T \setminus \{x\}$. When we need to make this dependence explicit we will write $x'(\hat{T})$.

3.4. Routing

In this subsection we let (C, d, ρ) be a rooted \mathbb{R} -tree. As this notation may suggest, we will apply the following constructions to a cut-tree; however, the quantities make sense more generally. While reading the following definitions, the reader may find it helpful to refer to the description of reconstruction in the discrete setting given in Section 1.3 for intuition.

Given $v, w \in C$, a *signpost* for v and w is a pair (v_1, w_1) with $v_1 \in C_{v \wedge w}^v$ and $w_1 \in C_{v \wedge w}^w$. Let $B = \{\emptyset\} \cup \bigcup_{n \in \mathbb{N}} \{0, 1\}^n$, and view B as indexing the vertices of a complete infinite binary tree with root \emptyset . If $b \in \{0, 1\}^n$ write $|b| = n$. For $b, b' \in B$, if b is a prefix of b' we write $b \leq b'$. This agrees with the genealogical order when B is viewed as a tree.

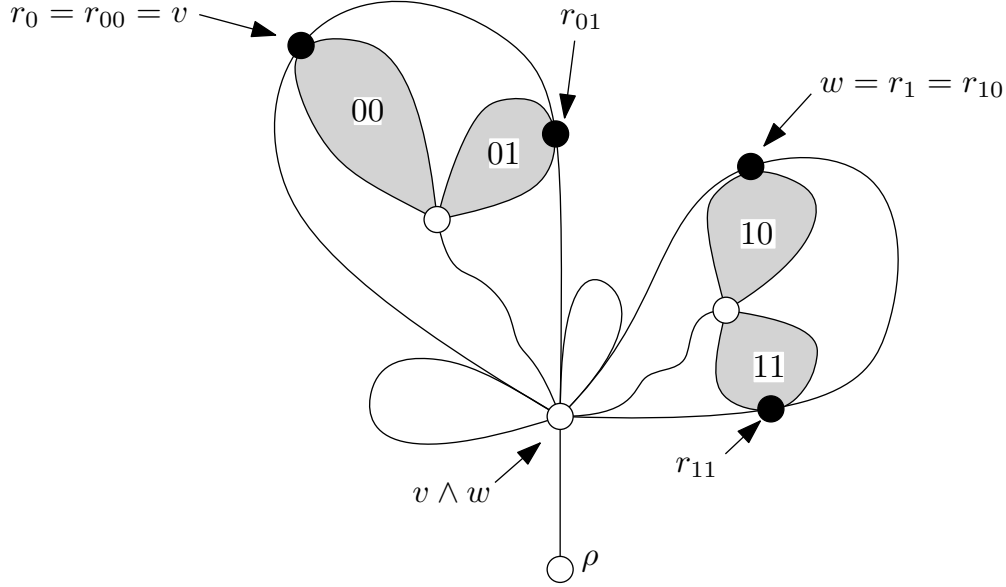


Fig 3: Part of a routing for a pair v, w of points of C . The pair (r_{01}, r_{11}) is a signpost for (v, w) . For each $b \in \{0, 1\}^2$, we must have $r_{b0} = r_b$, and r_{b1} must lie within the grey subtree labelled with b .

A *routing* for v and w is a collection $(r_b, b \in B \setminus \{\emptyset\})$ with the following properties. First, $r_0 = v$ and $r_1 = w$. Next, for all $n \geq 1$, for all $b \in \{0, 1\}^n$ we have $r_{b0} = r_b$; and for all $b \in \{0, 1\}^{n-1}$, (r_{b01}, r_{b11}) is a signpost for (r_{b0}, r_{b1}) .

Now fix a set $N \subset C$ such that $\overline{\bigcup_{i \in N} [\rho, i]} = C$. For each pair (i, j) of distinct elements of N let $r^{ij} = (r_b^{ij}, b \in B \setminus \{\emptyset\})$ be a routing for i and j . We say the collection $(r^{ij}, i, j \in N, i \neq j)$ of routings is *consistent* if the following three properties hold. In words, the first says that r^{ji} is always obtained from r^{ij} by swapping the subtrees at the root of B . The second says that sub-routings are themselves routings for the appropriate pairs. The third says that for any branchpoint x of C and any $y, z \in C_x$, if the subtrees C_x^y and C_x^z are the same then all signposts at x in directions y and z are also the same.

1. For any distinct $i, j \in N$, for all $b \in B$, $r_{0b}^{ij} = r_{1b}^{ji}$.
2. For any distinct $i, j \in N$, distinct $k, l \in N$ and $b \in B$, if $(k, l) = (r_{b0}^{ij}, r_{b1}^{ij})$ then $r_a^{kl} = r_{ba}^{ij}$ for all $a \in B \setminus \{\emptyset\}$.
3. For any distinct $i, j \in N$ and distinct $k, l \in N$, if $i \wedge j = k \wedge l$ and $C_{i \wedge j}^i = C_{k \wedge l}^k$ then $r_{01}^{kl} = r_{01}^{ij}$.

Suppose $(r^{ij}, i, j \in N, i \neq j)$ is consistent. Then, for each $x \in \text{br}(C)$ and each subtree $C' = C_x^y$ above x , fix $i, j \in N$ with $i \wedge j = x$ and $C_{i \wedge j}^i = C'$, and let $f(x, C') = r_{01}^{ij}$. The consistency conditions guarantee that the value of $f(x, C')$ does not depend on i and j satisfying these properties.

Conversely, suppose that we are given the data $f(x, C')$ for all such pairs (x, C') . Then we may reconstruct the collection of routings by setting $r_{01}^{ij} = f(x, C')$ for all pairs $i, j \in N$ with $i \wedge j = x$ and $C_{i \wedge j}^i = C'$; the consistency conditions then uniquely determine all other routing data.

3.5. Routing in cut-trees

Now suppose that (C, d, ρ) is constructed as in Section 3.3, so (C, d, ρ) is the completion of $C^\circ = C^\circ(T, \mathcal{P}, u)$ for a suitable triple (T, \mathcal{P}, u) with the property that $\ell(u_i) < \infty$ a.s. for all $i \in \mathbb{N}$. Recall that we are treating \mathbb{N} as a collection of random points of C . We assume that $\lambda(\llbracket x, y \rrbracket) > 0$ for all distinct $x, y \in T$.

For any distinct $i, j \in \mathbb{N} \subset C$, the cut-tree construction described above then yields a routing $R^{ij} = (R_b^{ij}, b \in B \setminus \{\emptyset\})$, built as follows. In reading the description, Figures 4 and 5 should be useful. In both figures, the superscripts ij are omitted for readability.

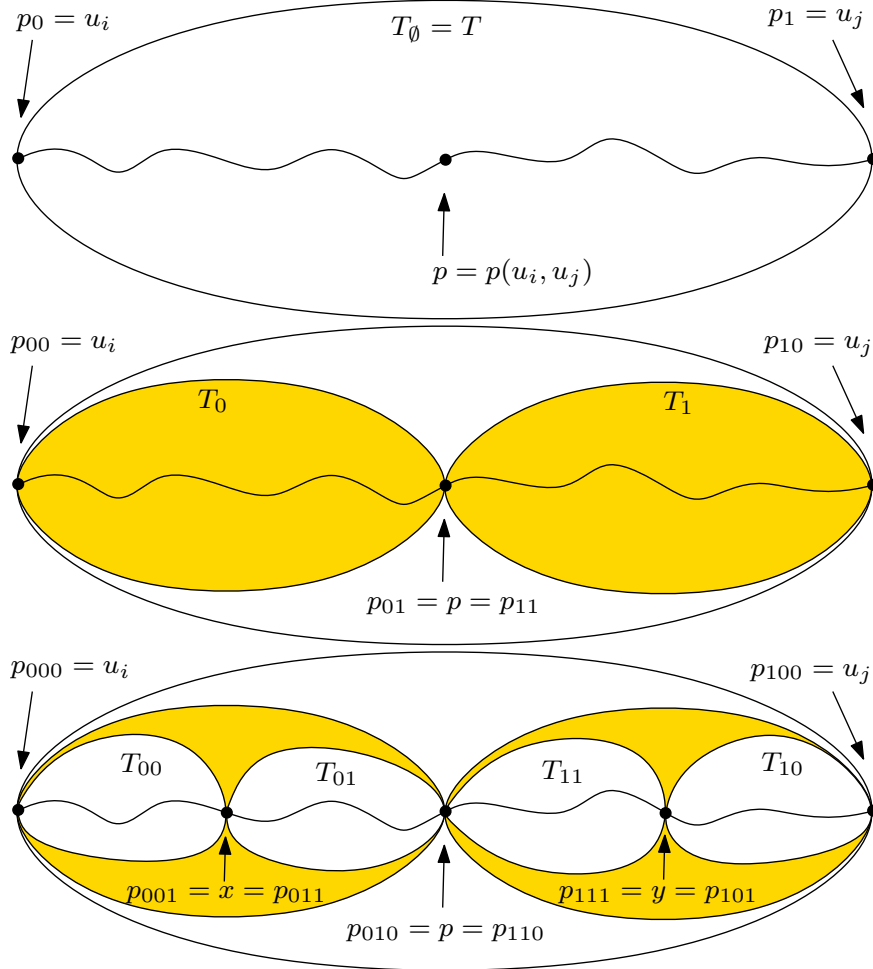


Fig 4: Top: The tree T with points u_i and u_j marked. Middle: the first cut point p and the trees T_0 and T_1 . Bottom: x and y are the first cuts to separate u_i and u_j from p , respectively.

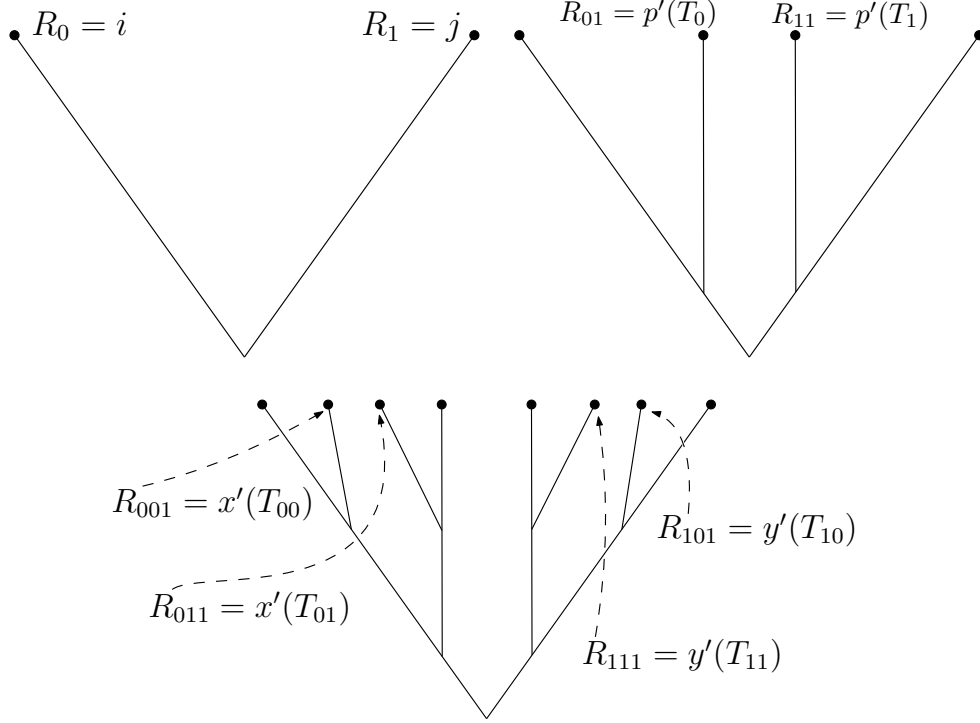


Fig 5: Top left: The images in C of u_i and u_j , and the path between them. Top right: The images R_{01} and R_{11} of $p = p_{01} = p_{11}$. Bottom: The images of $x = p_{001} = p_{011}$ and of $y = p_{111} = p_{101}$.

First, let $R_0^{ij} = i$ and $R_1^{ij} = j$. Also let $p_0^{ij} = u_i$ and $p_1^{ij} = u_j$. It will later be convenient to set $T_0^{ij} = T$.

Suppose inductively that $(R_b^{ij}, 0 < |b| \leq n)$ and $(p_b^{ij}, 0 < |b| \leq n)$ are already defined. Fix $b \in \{0, 1\}^{n-1}$ and let (p, t) be the first Poisson point separating p_{b0}^{ij} and p_{b1}^{ij} . In earlier notation, we have $(p, t) = (p(p_{b0}^{ij}, p_{b1}^{ij}), t(p_{b0}^{ij}, p_{b1}^{ij}))$. Such a point almost surely exists since $\lambda(\llbracket p_{b0}^{ij}, p_{b1}^{ij} \rrbracket) > 0$ by assumption.

Let T_{b0}^{ij} and T_{b1}^{ij} be the components of $T \setminus \{p_k : (p_k, t_k) \in \mathcal{P}, t_k \leq t\}$ containing p_{b0}^{ij} and p_{b1}^{ij} , respectively. In earlier notation, $T_{b0}^{ij} = T(p_{b0}^{ij}, t)$ and $T_{b1}^{ij} = T(p_{b1}^{ij}, t)$.

Recall that $p'(T_{b0}^{ij})$ and $p'(T_{b1}^{ij})$ are the images of p in C corresponding to components T_{b0}^{ij} and T_{b1}^{ij} . Then set $R_{b00}^{ij} = R_{b0}^{ij}$ and $R_{b01}^{ij} = p'(T_{b0}^{ij})$, and set $R_{b10}^{ij} = R_{b1}^{ij}$ and $R_{b11}^{ij} = p'(T_{b1}^{ij})$. Finally, set $p_{b00}^{ij} = p_{b0}^{ij}$, $p_{b10}^{ij} = p_{b1}^{ij}$ and $p_{b01}^{ij} = p = p_{b11}^{ij}$.

Observe that with the above definitions, for all $b \in B \setminus \{\emptyset\}$, R_b^{ij} is an image of p_b^{ij} in C . However, it need not be the unique such image, and indeed it is typically not. It is worth recording that the points $(p_b^{ij}, |b| = n)$ are all elements of $\llbracket u_i, u_j \rrbracket$ (with repetition). It follows from our recursive labelling convention that

$$d_T(u_i, u_j) = \sum_{|b|=n} d_T(p_{b0}^{ij}, p_{b1}^{ij}), \quad (3)$$

which will be useful in the next section.

We write $R = (R^{ij}, i, j \in \mathbb{N}, i \neq j)$ for the collection of such routings, or $R(T, \mathcal{P}, u)$ when we need such dependence to be explicit.

Proposition 10. *For all distinct $i, j \in \mathbb{N}$, R^{ij} is a routing for i and j . Furthermore, R is a consistent collection of routings.*

Proof. The first statement is by construction. Consistency is immediate from the fact that, in the notation

just preceding the proposition, for any i', j', b' with $R_{b'0}^{i'j'} = R_{b0}^{ij}$ and $R_{b'1}^{i'j'} = R_{b1}^{ij}$ we will have $R_{b'00}^{i'j'} = R_{b00}^{ij}$, $R_{b'10}^{i'j'} = R_{b10}^{ij}$, $R_{b'01}^{i'j'} = p'(T_{b0}^{ij})$ and $R_{b'11}^{i'j'} = p'(T_{b1}^{ij})$. \square

We view the triple (C, \mathbb{N}, R) as the image of (T, \mathcal{P}, u) under the cut-tree transformation. It should be understood as a random metric measure space with a countable infinity of marked points, in the sense discussed at the end of Section 2.1. The marks are the points \mathbb{N} together with the points $(R_b^{ij}, b \in B \setminus \{\emptyset\}, i, j \in \mathbb{N}, i \neq j)$. Since this is a countable collection, we may re-index it by the natural numbers according to some arbitrary but fixed rule. We will describe a convenient such rule in the course of proving Proposition 12, below.

4. The case of stable trees

Throughout Section 4, we fix $\alpha \in (1, 2]$ and let $T = (T, d_T, \mu)$ be an α -stable tree. We further let $U = (U_i, i \in \mathbb{N})$ be an i.i.d. sequence of samples from μ , and let \mathcal{P} be a Poisson process on $T \times [0, \infty)$ with intensity measure $\Lambda \otimes dt$, where Λ is as defined in Section 2.3. For $s \geq 0$, let $F(s) = (F_1(s), F_2(s), \dots)$ be the sequence of μ -masses of the nontrivial connected components of $T \setminus \{p : (p, t) \in \mathcal{P}, t \leq s\}$, listed in decreasing order. Then, as discussed in the introduction, $(F(s), s \geq 0)$ is a self-similar fragmentation process [9, 47]. We observe, in particular, that defining $(F(s), s \geq 0)$ as above for any $\alpha \in (1, 2]$, by Lemma 10 of [9] and Lemma 8(iii) of [47], for each $s \geq 0$, we have $\sum_{i \geq 1} F_i(s) = 1$ almost surely, so that the fragmentation process conserves mass.

We let $C = C(T, \mathcal{P}, U)$ be the cut-tree, and as usual write $C = (C, d, \rho, \nu)$. The following result, which is due to Bertoin and Miermont [17] in the case $\alpha = 2$ and to Dieuleveut [25] in the case $\alpha \in (1, 2)$, states that the cut-tree of a stable tree is again a stable tree.

Theorem 11. *Let $U_0 \sim \mu$ be a random point of T independent of the points in U . Then we have*

$$(d_T(U_i, U_j), i, j \in \{0\} \cup \mathbb{N}) \stackrel{d}{=} (d(i, j), i, j \in \{\rho\} \cup \mathbb{N}).$$

In particular, Theorem 11 implies that the assumption of Proposition 5 holds, so C is well-defined. The theorem then implies that C has the same distribution as (T, d_T, U_0, μ) . Since C is an α -stable tree, it is compact.

In T , the points $(U_i, i \in \{0\} \cup \mathbb{N})$ are i.i.d samples from the mass measure. Also, almost surely $\Lambda(\llbracket x, y \rrbracket) > 0$ for all distinct $x, y \in T$, which places us in the setting of Section 3.2. Thus, by Proposition 9, ν is the empirical measure of \mathbb{N} in C , so $(C, \{\rho\} \cup \mathbb{N})$ is distributed as an α -stable tree together with a sequence of i.i.d. samples from ν .

As just noted, Λ a.s. charges every interval of positive length, which allows us the collection $R = (R^{ij}, i, j \in \mathbb{N}, i \neq j)$ of routings described in Section 3.5 to be defined. By Proposition 10, R is consistent.

4.1. Routings in stable trees

For each $x \in \text{br}(C)$ and $y \in C_x$, let Z_x^y be the common value of all the random variables R_{01}^{ij} for which $C_{i \wedge j}^i = C_x^y$, where as before C_x^y is the subtree above x containing y . We emphasise that there is redundancy in our notation for the set of routing variables $Z = \{Z_x^y : x \in \text{br}(C), y \in C_x\}$ since there are multiple ways of specifying the same tree C_x^y . However, each random variable appears only once; see Figure 6.

Proposition 12. *The law of (C, \mathbb{N}, Z) is as follows.*

1. C is a stable tree endowed with its mass measure.
2. The elements of $\{\rho\} \cup \mathbb{N}$ are i.i.d. with law ν .
3. For each $x \in \text{br}(C)$ and $y \in C_x$, $Z_x^y \sim \nu_{C_x^y}$. Moreover, these random variables are conditionally independent given $(C, \{\rho\} \cup \mathbb{N})$.

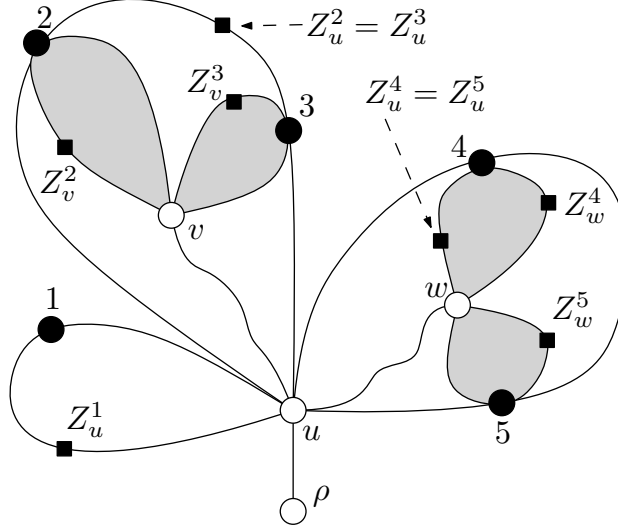


Fig 6: The routing variables Z_x^y for $x \in \{u, v, w\}$ and $y \in \{1, \dots, 5\}$. The redundancy in the notation is witnessed by the fact that $Z_u^2 = Z_u^3$ and $Z_u^4 = Z_u^5$.

Note that since R is consistent, it is completely determined by Z . The preceding proposition therefore fully specifies the joint law of the triple (C, \mathbb{N}, R) . Observe that (3) implies that the only dependence between the Z_x^y and $\{\rho\} \cup \mathbb{N}$ is via the labelling of subtrees.

Proof of Proposition 12. The first two statements are contained in Theorem 11. Next, for $S \subset \mathbb{N}$ write $Z(S) = \{Z_{i \wedge j}^i : i, j \in S, i \neq j\}$. To prove the third statement, it suffices to verify that the triple $(C, \mathbb{N}, Z(S))$ has the appropriate law for all finite subsets S of \mathbb{N} . In doing so we may assume without loss of generality that $S = [k]$ for some $k \geq 1$.

We argue by induction on k , but first introduce a small amount of notation. For $k \geq 2$, list the elements of $Z([k])$ without repetition as $Z_{x(1)}^{y(1)}, \dots, Z_{x(m)}^{y(m)}$. (So, for example, in Figure 6, with $k = 5$ we have $m = 7$ and we may take $(x(1), y(1)) = (u, 1)$, $(x(2), y(2)) = (u, 2)$, $(x(3), y(3)) = (u, 4)$, $(x(4), y(4)) = (v, 2)$, $(x(5), y(5)) = (v, 3)$, $(x(6), y(6)) = (w, 4)$ and $(x(7), y(7)) = (w, 5)$.) We omit the dependence on k from our notation.

We will show by induction that given $(C, \{\rho\} \cup \mathbb{N})$, for all $1 \leq i \leq m$, $Z_{x(i)}^{y(i)}$ has law $\nu_{C_{x(i)}^{y(i)}}$ and, moreover, the random variables $Z_{x(i)}^{y(i)}$ are conditionally independent. This identifies the law of the Z_x^y .

The case $k = 1$ is immediate as $Z([1])$ is empty, and the joint law of the points in $\{\rho\} \cup \mathbb{N}$ is given by Theorem 11. We hereafter assume $k \geq 2$.

Let b be the nearest branchpoint to ρ in $\bigcup_{i \in [k]} \llbracket \rho, i \rrbracket$. List the elements of $\{C_b^i, i \in [k]\}$ without repetition as $C_b(1), \dots, C_b(\ell)$, with $C_b(j) = (C_b(j), d, b)$. Note that $2 \leq \ell \leq k$, so that for each $1 \leq j \leq \ell$ the subtree $C_b(j)$ contains $k_j \leq k - 1$ elements of $[k]$. We list the points of \mathbb{N} lying in $C_b(j)$ in increasing order as $N_j = (n(j, m), m \geq 1)$. Observe that the sets $N_j \cap [k]$, with $1 \leq j \leq \ell$, partition $[k]$.

There is a unique point (p, t) of \mathcal{P} which first separates $U_{n(1,1)}$ from $U_{n(2,1)}$ in T . In the notation of Section 3.1, this is the point $(p, t) = (p(U_{n(1,1)}, U_{n(2,1)}), t(U_{n(1,1)}, U_{n(2,1)}))$. Note that this is also the first separator of any $U_{n(i,1)}$ and $U_{n(j,1)}$ for distinct i and j .

Since $U_{n(i,1)}$ is a.s. not a cut point, $T(U_{n(i,1)}, t)$ is a.s. the connected component of $T \setminus \{p_j : t_j \leq t\}$ containing $U_{n(i,1)}$. Since this does not contain the point p we let $T^*(U_{n(i,1)}) = T(U_{n(i,1)}, t) \cup \{p\}$ and $T^*(U_{n(i,1)})$ be the subtree of T induced by $T^*(U_{n(i,1)})$.

In [47], Miermont showed that $T^*(U_{n(i,1)})$ is a stable tree of mass $\mu(T^*(U_{n(i,1)}))$; see, in particular, the proof of his Lemma 9. Furthermore, writing $\mathcal{P}^*(U_{n(i,1)}) = \{(p_j, t_j - t) : p_j \in T(U_{n(i,1)}, t)\}$, then $\mathcal{P}^*(U_{n(i,1)})$

is a Poisson process with intensity measure $\lambda|_{T^*(U_{n(i,1)})} \otimes dt$. The pairs $((T^*(U_{n(i,1)}), \mathcal{P}^*(U_{n(i,1)})), i \in [\ell])$ only depend on each other via the vector of masses $(\mu(T^*(U_{n(i,1)})), i \in [\ell])$. Finally, for each $i \in [\ell]$, the points $U_i = (U_j : j \in N_i)$ are precisely those U_j lying in $T^*(U_{n(i,1)})$, and are i.i.d. with law $\mu_{T^*(U_{n(i,1)})}$. By Theorem 2, p also has law $\mu_{T^*(U_{n(i,1)})}$.

For each $1 \leq i \leq \ell$, the tree $C_b(i)$, which is rooted at b , is obtained as the cut-tree of $(T^*(U_{n(i,1)}), \mathcal{P}^*(U_{n(i,1)}), U_i)$. It follows that $C_b(i)$ is a stable tree of mass $\nu(C_b(i)) = \mu(T^*(U_{n(i,1)}))$. The facts from the preceding paragraph also imply that the trees $(C_b(i), i \in [\ell])$ are conditionally independent given their masses. Furthermore, the elements of N_i are precisely the images in $C_b(i)$ of the points in U_i , and $Z_b^{n(i,1)}$ is the image of p . These facts are special cases of the observation about cut-trees of subtrees described at the end of Section 3.3. Finally, note that we may also view $C_b(i)$ as the cut-tree of $(T^*(U_{n(i,1)}), \mathcal{P}^*(U_{n(i,1)}), \{p\} \cup U_i)$, as described near the start of Section 3.3.

Now apply the inductive hypothesis to

$$\left(C_b(i), \left(b, Z_b^{n(i,1)}, n(i,1), n(i,2), \dots \right), Z(N_i \cap [k]) \setminus \{Z_b^{n(i,1)}\} \right),$$

for each $i \in [\ell]$. This is permitted since b , $Z_b^{n(i,1)}$, and N_i have the correct conditional joint law, and since $|N_i \cap [k]| = k_i < k$.

We obtain by induction that $Z_b^{n(1,1)}, \dots, Z_b^{n(\ell,1)}$ are conditionally independent, with $Z_b^{n(i,1)} \sim \nu_{C_b(i)}$. We emphasise that this is because $Z_b^{n(i,1)}$ is the image of a uniform point in $T^*(U_{n(i,1)})$; in the induction $Z_b^{n(i,1)}$ is no longer playing the role of a routing variable.

We further obtain that for each $1 \leq i \leq \ell$, the elements of $Z(N_i \cap [k]) \setminus \{Z_b^{n(i,1)}\}$ have the correct conditional joint law and are independent of $Z_b^{n(i,1)}$. Here we are again using the difference in roles between $Z_b^{n(i,1)}$ and the other elements of $Z(N_i \cap [k])$. Finally, the subtrees $C_b(i)$ are conditionally independent given their masses, which yields the requisite conditional independence of the collections $Z(N_i \cap [k])$ for $1 \leq i \leq \ell$. Since

$$Z([k]) = \bigcup_{i=1}^{\ell} Z(N_i \cap [k]),$$

this fully identifies the conditional joint law of the random variables $Z_{x(1)}^{y(1)}, \dots, Z_{x(m)}^{y(m)}$ which comprise $Z([k])$, and so completes the proof. \square

4.2. Distributional identities for stable trees

We recall the definitions of some distributions that play a role in the sequel. Write $\Delta_n = \{(x_1, \dots, x_n) \in \mathbb{R}_+^n : \sum_{i=1}^n x_i = 1\}$. A Δ_n -valued random vector $\mathbf{X} = (X_1, \dots, X_n)$ has the *Dirichlet distribution* $\text{Dir}(\theta_1, \dots, \theta_n)$ if its density with respect to Lebesgue measure on Δ_n is

$$\frac{\Gamma(\sum_{i=1}^n \theta_i)}{\prod_{i=1}^n \Gamma(\theta_i)} \prod_{i=1}^n x_i^{\theta_i - 1}.$$

A non-negative random variable Y has the *Mittag-Leffler distribution with parameter* $\beta \in (0, 1)$, denoted $\text{ML}(\beta)$, if it satisfies

$$\mathbf{E}[Y^p] = \frac{\Gamma(p+1)}{\Gamma(p\beta+1)}, \quad (4)$$

for $p \geq -1$. This equation determines the law of Y (see [50, p.10] and [18, p.391]). Write g_β for the density of $\text{ML}(\beta)$ with respect to Lebesgue measure. Write $\widehat{\text{ML}}(\beta)$ for the size-biased distribution, which has density

$$\hat{g}_\beta(r) := \Gamma(\beta+1) r g_\beta(r), \quad r \geq 0.$$

Our proofs exploit the following characterization of the size-biased Mittag-Leffler distribution,

Lemma 13. *Fix $\beta \in [1/2, 1)$, and let M , M_1 and M_2 be independent and identically distributed non-negative random variables with*

$$\mathbf{E}[M] = \frac{2\Gamma(\beta + 1)}{\Gamma(2\beta + 1)}.$$

Let $(X_1, X_2, X_3) \sim \text{Dir}(\beta, \beta, 1 - \beta)$ be independent of M_1 and M_2 . Then M solves the recursive distributional equation

$$X_1^\beta M_1 + X_2^\beta M_2 \stackrel{(d)}{=} M, \quad (5)$$

if and only if $M \sim \widehat{\text{ML}}(\beta)$.

Let x, y, z be independent points of T with common law μ , and let b be the common branchpoint of x, y, z (i.e. the unique element of $\llbracket x, y \rrbracket \cap \llbracket x, z \rrbracket \cap \llbracket y, z \rrbracket$). Recall from Section 2.1 that T_b^x is the subtree of T rooted at b and containing x .

Theorem 14. 1. *The random variable $\alpha \cdot d(x, y)$ is $\widehat{\text{ML}}(1 - \frac{1}{\alpha})$ -distributed.*
 2. *The vector $(\mu(T_b^x), \mu(T_b^y), \mu(T \setminus (T_b^x \cup T_b^y)))$ is $\text{Dir}(1 - \frac{1}{\alpha}, 1 - \frac{1}{\alpha}, \frac{1}{\alpha})$ -distributed.*

Proof. For the first assertion, see Theorem 3.3.3 of [27]. The second follows from Corollary 10 of [39] after using distributional identities established in [35], and is explicitly noted as [35, (4.1)]. When $\alpha = 2$, the law of the distance between two uniform points was earlier proved to follow the Rayleigh distribution, by Aldous [7]. This is in agreement with the current result, up to a choice of normalization, since if Z is standard Rayleigh then the law of $\sqrt{2}Z$ is $\widehat{\text{ML}}(1/2)$; see Section 1.1 of [35]. \square

We can now proceed to the proof of Lemma 13.

Proof of Lemma 13. Write $\beta = 1 - 1/\alpha$; then $\alpha \in (1, 2]$. Consider again the α -stable tree T with x, y, z independent points sampled according to the law μ . Let $M = \alpha \cdot d(x, y)$, $M_1 = \alpha \mu(T_b^x)^{-\beta} \cdot d(x, b)$, $M_2 = \alpha \mu(T_b^y)^{-\beta} \cdot d(y, b)$, $X_1 = \mu(T_b^x)$, $X_2 = \mu(T_b^y)$ and $X_3 = \mu(T_b^z)$. Then $M = X_1^\beta M_1 + X_2^\beta M_2$. Moreover, by Theorems 3 and 14, M is $\widehat{\text{ML}}(\beta)$ -distributed, and M_1, M_2 are independent $\widehat{\text{ML}}(\beta)$ random variables, independent of $(X_1, X_2, X_3) \sim \text{Dir}(\beta, \beta, 1 - \beta)$. It follows that $\widehat{\text{ML}}(\beta)$ satisfies the RDE (5).

For the converse, we use that the left-hand side of (5) is an instance of the *smoothing transform* applied to the law of M . The fixed points of the smoothing transform have been completely characterised by Durrett and Liggett [30]. Indeed, the space of fixed points is determined by the analytical properties of the function $\nu : \mathbb{R}_+ \rightarrow \mathbb{R}$ defined (in our setting) by

$$\nu(s) = \log \left(\mathbf{E} \left[X_1^{\beta s} \mathbb{1}_{X_1 > 0} + X_2^{\beta s} \mathbb{1}_{X_2 > 0} \right] \right),$$

for $s \geq 0$. Since X_1 and X_2 are marginally both distributed as $\text{Beta}(\beta, 1)$, it is easily checked that $X_1^\beta \stackrel{(d)}{=} X_2^\beta \stackrel{(d)}{=} U$, where U is uniform on $[0, 1]$. Hence,

$$\nu(s) = \log(2\mathbf{E}[U^s]) = \log(2) - \log(s + 1), \quad s \geq 0,$$

for any $\alpha \in (1, 2]$. Observe that ν has its unique zero in $(0, 1]$ at $s = 1$, and that $\nu'(1) = -1/2 < 0$. In this case, [30, Theorem 2(a)] entails that (5) has a unique distributional solution, up to multiplication by a non-negative constant; we have already identified this solution as the $\widehat{\text{ML}}(\beta)$ distribution. \square

4.3. Reconstructing the distance between a pair of points in T

Recall that $C = (C, d, \rho, \nu)$ is the cut-tree of T , that the points of \mathbb{N} are i.i.d. with law ν and that R is the collection of routings for the elements of \mathbb{N} . For the remainder of Section 4.3 we fix distinct $i, j \in \mathbb{N}$ and recall that the routing for i and j is denoted R^{ij} .

Recall that for distinct nodes y and z of C , the subtree of C above z containing y is denoted by C_z^y . Now fix $b \in B \setminus \{\emptyset\}$, and let b' be the sibling of b in B ; so if $b = \hat{b}0$ then $b' = \hat{b}1$ and vice versa. Then let $M_b = \nu(C_{R_b \wedge R_{b'}}^{R_b})$. In Figure 3, for example, M_{00} is the mass of the shaded subtree labelled 00, and likewise for M_{01} , M_{10} and M_{11} . It is crucial in what follows that, in the notation of Section 3.5, we also have $M_b = \mu(T_b^{ij})$ since $C_{R_b \wedge R_{b'}}^{R_b}$ is the cut-tree of T_b^{ij} . We also set $M_\emptyset = \nu(C) = 1 = \mu(T)$.

It is convenient to write

$$(\Delta_{b0}, \Delta_{b1}, 1 - \Delta_{b0} - \Delta_{b1}) = \frac{1}{M_b} (M_{b0}, M_{b1}, M_b - M_{b0} - M_{b1}).$$

We will repeatedly use that for all $b \in B$,

$$(\Delta_{b0}, \Delta_{b1}, 1 - \Delta_{b0} - \Delta_{b1}) \sim \text{Dir}\left(1 - \frac{1}{\alpha}, 1 - \frac{1}{\alpha}, \frac{1}{\alpha}\right),$$

and that the vectors $\{(\Delta_{b0}, \Delta_{b1}, 1 - \Delta_{b0} - \Delta_{b1}), b \in B\}$ are mutually independent; these properties follow from Theorems 3 and 14, and Proposition 12.

Next, for $n \geq 0$, let

$$Y_n = Y_n(i, j) = \sum_{|b|=n} M_b^{1-1/\alpha}. \quad (6)$$

Proposition 15. *As $n \rightarrow \infty$, $Y_n \xrightarrow{\text{a.s.}} Y$ where the limit $Y = Y(i, j)$ satisfies*

$$\frac{2\Gamma(2 - \frac{1}{\alpha})}{\Gamma(3 - \frac{2}{\alpha})} Y \sim \widehat{\text{ML}}(1 - 1/\alpha).$$

Proof. Let $\mathcal{G}_n = \sigma((M_b, |b| \leq n))$, and let $\mathcal{G}_\infty = \sigma((M_b, b \in B)) = \sigma(\bigcup_n \mathcal{G}_n)$.

Explicit calculation (as in the proof of Lemma 13) shows that $\Delta_{b0}^{1-1/\alpha}$ and $\Delta_{b1}^{1-1/\alpha}$ are both uniformly distributed on $[0, 1]$, so

$$\mathbf{E}[Y_{n+1} \mid \mathcal{G}_n] = \sum_{|b|=n} M_b^{1-1/\alpha} \cdot \mathbf{E}\left[\Delta_{b0}^{1-1/\alpha} + \Delta_{b1}^{1-1/\alpha} \mid \mathcal{G}_n\right] = Y_n$$

and so (Y_n) is a (\mathcal{G}_n) -martingale. Since $Y_n \geq 0$ for all n , it follows that $Y_n \xrightarrow{\text{a.s.}} Y$ for some random variable Y by the martingale convergence theorem; it remains to show that Y has the correct distribution. For this we use a second martingale argument together with a result of [12].

Fix $n \in \mathbb{N}$ and let $(Z_b^{(n)} : |b| = n)$ be i.i.d. $\widehat{\text{ML}}(1 - \frac{1}{\alpha})$. Then for $b \in B$ with $|b| = m < n$, define $Z_b^{(n)}$ inductively by

$$Z_b^{(n)} = \Delta_{b0}^{1-1/\alpha} Z_{b0}^{(n)} + \Delta_{b1}^{1-1/\alpha} Z_{b1}^{(n)}.$$

By Lemma 13, $Z_b^{(n)} \sim \widehat{\text{ML}}(1 - \frac{1}{\alpha})$ for all b with $|b| \leq n$. Furthermore, the families $(Z_b^{(n)}, |b| \leq n)$ are consistent in n , in that

$$(Z_b^{(n)}, |b| \leq n-1) \sim (Z_b^{(n-1)}, |b| \leq n-1),$$

and so have a projective limit by Kolmogorov's extension theorem. Let $(Z_b, b \in B)$ be such that $(Z_b, |b| \leq n) \sim (Z_b^{(n)}, |b| \leq n)$ for all n ; in particular, for all b we have $Z_b \sim \widehat{\text{ML}}(1 - \frac{1}{\alpha})$, and

$$Z_b = \Delta_{b0}^{1-1/\alpha} Z_{b0} + \Delta_{b1}^{1-1/\alpha} Z_{b1}. \quad (7)$$

The families $(Z_b, b \in B)$ and $(\Delta_b, b \in B)$ together define a *recursive tree process* in the sense of [12]. This process is easily seen to verify the conditions of Corollary 17 of [12] (briefly: $\mathbf{E}[\Delta_b^x]$ is decreasing in x and $\mathbf{P}\{Z_\emptyset = 0\} = 0$). We conclude that the recursive tree process is *endogenous*, which means that for all $b \in B$, the random variable Z_b is measurable with respect to $\sigma(\Delta_{b'}, b < b')$. In particular, Z_\emptyset is integrable and \mathcal{G}_∞ -measurable, and so the martingale convergence theorem gives that

$$\mathbf{E}[Z_\emptyset \mid \mathcal{G}_n] \xrightarrow{\text{a.s.}} Z_\emptyset$$

as $n \rightarrow \infty$.

Finally, by (7) and induction we have $Z_\emptyset = \sum_{|b|=n} M_b^{1-1/\alpha} Z_b$ for all n . Also, Z_b is $\sigma(\Delta_{b'}, b < b')$ -measurable, so if $|b| = n$ then Z_b is independent of \mathcal{G}_n . On the other hand, M_b is \mathcal{G}_n -measurable and so

$$\mathbf{E}[Z_\emptyset \mid \mathcal{G}_n] = \sum_{|b|=n} M_b^{1-1/\alpha} \mathbf{E}[Z_b] = \frac{2\Gamma(2 - \frac{1}{\alpha})}{\Gamma(3 - \frac{2}{\alpha})} \cdot Y_n,$$

where the last equality holds as $Z_b \sim \widehat{\text{ML}}(1 - \frac{1}{\alpha})$. It follows that $\frac{2\Gamma(2 - \frac{1}{\alpha})}{\Gamma(3 - \frac{2}{\alpha})} \cdot Y = Z_\emptyset$ almost surely. \square

Let

$$\delta_C(i, j) = \frac{2\alpha\Gamma(2 - \frac{1}{\alpha})}{\Gamma(3 - \frac{2}{\alpha})} \cdot Y(i, j). \quad (8)$$

Observe that, by Lemma 13 and Theorem 14 (1), this has the same law as $d_T(U_i, U_j)$. In Theorem 16 below, we will show that the two quantities are, in fact, almost surely equal.

Next, for $b \in B$ write

$$Y_n^b = Y_n^b(i, j) = \sum_{|b'|=n, b \leq b'} M_{b'}^{1-1/\alpha}.$$

A practically identical proof then shows that for all $b \in B$, $Y_n^b \xrightarrow{\text{a.s.}} Y^b = Y^b(i, j)$, for random variables $(Y^b, b \in B)$ all a.s. satisfying $Y^b = Y^{b0} + Y^{b1}$. In this notation we have $Y(i, j) = Y^\emptyset(i, j)$. In particular, this allows us to define

$$\delta_C(i, i \wedge j) = \frac{2\alpha\Gamma(2 - \frac{1}{\alpha})}{\Gamma(3 - \frac{2}{\alpha})} \cdot Y^0(i, j), \quad \text{and} \quad \delta_C(j, i \wedge j) = \frac{2\alpha\Gamma(2 - \frac{1}{\alpha})}{\Gamma(3 - \frac{2}{\alpha})} \cdot Y^1(i, j). \quad (9)$$

The relation between the Y^b then implies that $\delta_C(i, j) = \delta_C(i, i \wedge j) + \delta_C(j, i \wedge j)$

4.4. Recovering (T, \mathcal{P}, U) from (C, \mathbb{N}, R)

Recall that a.s. for all distinct $i, j \in \mathbb{N}$, R^{ij} is a routing for i and j by Proposition 10. For distinct $i, j \in \mathbb{N}$, let $\delta_C(i, j)$, $\delta_C(i, i \wedge j)$ and $\delta_C(j, i \wedge j)$ be defined as in (8) and (9). Note that i and j have law ν , so by Proposition 15, $Y(i, j)$ is well-defined and $Y(i, j) \cdot 2\Gamma(2 - \frac{1}{\alpha})/\Gamma(3 - \frac{2}{\alpha})$ is $\widehat{\text{ML}}(1 - 1/\alpha)$ -distributed.

Theorem 16. *The following all hold almost surely.*

1. $d_T(U_i, p(U_i, U_j)) = \delta_C(i, i \wedge j)$ and $d_T(U_j, p(U_i, U_j)) = \delta_C(j, i \wedge j)$, and thus $d_T(U_i, U_j) = \delta_C(i, j)$.

2. Let $\pi(i, j) \in T$ be the unique element of $\llbracket U_i, U_j \rrbracket$ at distance $\delta_C(i, i \wedge j)$ from U_i , and let

$$\tau(i, j) = \int_{\llbracket \rho, i \wedge j \rrbracket} \frac{1}{\nu(Cz)} dz,$$

where the integral is with respect to the length measure on $\llbracket \rho, i \wedge j \rrbracket$. Then

$$(t(U_i, U_j), p(U_i, U_j)) = (\tau(i, j), \pi(i, j)).$$

Proof. For each $b \in B$, let

$$D_b = \frac{d_T(p_{b0}^{ij}, p_{b1}^{ij})}{M_b^{1-1/\alpha}}.$$

The reader may wish to glance at Figure 4 to refresh the definitions of the points p_{b0}^{ij} and p_{b1}^{ij} , and their relation to T_b^{ij} . (When consulting that figure, it may be useful to take $b = 11$, say, for concreteness. Also recall that the superscripts ij are omitted from the figure for legibility.)

For each $n \geq 0$, the trees $(T_b^{ij}, |b| = n)$ are rescaled α -stable trees and are conditionally independent given their masses. Moreover, the random variables $(\alpha^{-1} D_b, |b| = n)$ are i.i.d. and are $\widehat{\text{ML}}(1 - 1/\alpha)$ -distributed. These observations are consequences of Theorems 3 and 14.

Now note that, by (3), we have

$$\sum_{|b|=n} D_b M_b^{1-1/\alpha} = \sum_{|b|=n} d_T(p_{b0}^{ij}, p_{b1}^{ij}) = d_T(U_i, U_j),$$

for every n , and thus we trivially have

$$\lim_{n \rightarrow \infty} \sum_{|b|=n} D_b M_b^{1-1/\alpha} = d_T(U_i, U_j).$$

On the other hand, $(D_b, |b| = n)$ is independent of $(M_b, b \in B)$, so with $\mathcal{G}_n = \sigma((M_b, |b| \leq n))$ as in Proposition 15, we have

$$\mathbf{E}[d_T(U_i, U_j) \mid \mathcal{G}_n] = \mathbf{E}[D_\emptyset] \cdot \sum_{|b|=n} M_b^{1-1/\alpha} = \mathbf{E}[D_\emptyset] \cdot Y_n(i, j).$$

Taking n to infinity, it follows that

$$\mathbf{E}[d_T(U_i, U_j) \mid \mathcal{G}_\infty] \stackrel{\text{a.s.}}{=} \mathbf{E}[D_\emptyset] \cdot Y(i, j).$$

In view of (8) and the subsequent observation, it follows that

$$\mathbf{E}[D_\emptyset] = \frac{2\alpha\Gamma(2 - \frac{1}{\alpha})}{\Gamma(3 - \frac{2}{\alpha})}.$$

Thus, $d_T(U_i, U_j) \stackrel{\text{d}}{=} \mathbf{E}[D_\emptyset] \cdot Y(i, j)$, which implies that $d_T(U_i, U_j)$ is \mathcal{G}_∞ -measurable (see [29], Exercise 5.1.12). We thus have that $d_T(U_i, U_j) \stackrel{\text{a.s.}}{=} \mathbf{E}[D_\emptyset] \cdot Y(i, j)$. An essentially identical proof shows that $d_T(U_i, p(U_i, U_j)) \stackrel{\text{a.s.}}{=} \mathbf{E}[D_\emptyset] \cdot Y(i, i \wedge j)$ and that $d_T(U_j, p(U_i, U_j)) \stackrel{\text{a.s.}}{=} \mathbf{E}[D_\emptyset] \cdot Y(j, i \wedge j)$. This establishes the first claim of the theorem. In particular, this implies that $\pi(i, j)$ and $\tau(i, j)$ are well-defined.

For the second claim, by definition we have $\alpha(i \wedge j) = t(U_i, U_j)$, and (1) also gives

$$\alpha(i \wedge j) = \inf \left\{ t : \int_0^t \mu(T(U_i, r)) dr \geq d(\rho, i \wedge j) \right\}.$$

It is convenient to parameterise $\llbracket \rho, i \rrbracket$ by length; to this end, for $\gamma \in [0, \ell(U_i))$, write $z(\gamma)$ for the unique point $z \in \llbracket \rho, i \rrbracket$ with $d(\rho, z) = \gamma$. Then for all such γ we have

$$\int_0^{\alpha(z(\gamma))} \mu(T(U_i, r)) dr = \gamma,$$

from which it follows that

$$\alpha(z(\gamma)) = \int_0^\gamma \frac{1}{\mu(T(U_i, \alpha(z(y))))} dy.$$

Recall that we also have $\alpha(i \wedge j) = t(U_i, U_j)$. The result will thus follow if we can show that $\mu(T(U_i, \alpha(z))) = \nu(C_z)$ for $z \in \llbracket \rho, i \rrbracket$, by taking $\gamma = d(\rho, i \wedge j)$ so that $z(\gamma) = i \wedge j$. We have

$$\{j \in \mathbb{N} : j \in C_z\} = \{j \in \mathbb{N} : U_j \in T(U_i, \alpha(z))\}.$$

We also have

$$\mu(T(U_i, \alpha(z))) = \lim_{n \rightarrow \infty} \frac{1}{n} \#\{j \leq n : U_j \in T(U_i, \alpha(z))\}.$$

by the Glivenko-Cantelli theorem and

$$\nu(C_z) = \lim_{n \rightarrow \infty} \frac{1}{n} \#\{j \leq n : j \in C_z\}$$

by Proposition 9. This completes the proof. \square

Corollary 17. *The triple (T, U, \mathcal{P}) is measurable with respect to the triple (C, \mathbb{N}, R) .*

Proof. First, since U is a.s. dense in T , the collection of pairwise distances

$$(\delta_C(i, j), i, j \in \mathbb{N}) = (d_T(U_i, U_j), i, j \in \mathbb{N})$$

uniquely reconstructs (T, d_T) up to metric space isometry, and further reconstructs the sequence U of points of T . Next, since μ is the empirical measure of the collection U , this also reconstructs μ and thus reconstructs $T = (T, d_T, \mu)$ up to measured metric space isometry.

Finally note that, almost surely, every point $(p, t) \in \mathcal{P}$ separates some pair of points from the sequence U . In other words, every element of \mathcal{P} may be represented as $(p, t) = (p(U_i, U_j), t(U_i, U_j))$ for some $i, j \in \mathbb{N}$. It follows from the second statement of Theorem 16 that, almost surely, we may reconstruct \mathcal{P} from C and the routings R as

$$\mathcal{P} = \{(\pi(i, j), \tau(i, j)) : i, j \in \mathbb{N}, i \neq j\}.$$

\square

The proof of Corollary 17 describes a specific measurable map, which we now denote Φ , with the property that $\Phi(C, \mathbb{N}, R) \stackrel{\text{a.s.}}{=} (T, U, \mathcal{P})$. The map Φ is built using δ_C , the empirical measure, and the points $(\pi(i, j), \tau(i, j))$. Denoting the laws of the triples (C, \mathbb{N}, R) and (T, U, \mathcal{P}) by \mathcal{L} and \mathcal{M} , respectively, this immediately entails the following corollary.

Corollary 18. *Let (C', \mathbb{N}, R') be any random variable with law \mathcal{L} . Then $\Phi(C', \mathbb{N}, R')$ has law \mathcal{M} .*

5. Questions and perspectives

Though there are now several important cases in which the cut-trees and their reconstructions are well-understood, it remains to develop a fully general theory. The following list of questions provide some concrete avenues for research along these lines.

1. The results on empirical measures in Section 3.2 require compactness of the cut tree. In the case of α -stable trees considered in this work, compactness follows from existing results in the literature. More generally, though it is likely possible to prove compactness *ad hoc* for specific models, it would be interesting to develop general sufficient conditions for compactness of the cut tree of an \mathbb{R} -tree.
2. Are there cases other than those addressed by the current paper or by Broutin and Wang [19] where the cut tree has the same law as the original tree?
3. What conditions on the law of (T, U, \mathcal{P}) are sufficient to guarantee that the triple may almost surely be reconstructed from (C, \mathbb{N}, R) ?
4. For a given triple (T, U, \mathcal{P}) even if the cut tree C is not compact, it may be that in some cases the images of the points in u define an “empirical measure” on C . When does this occur?
5. The distributional identities that this paper is about are “annealed” in that one averages over the realization of the tree, the sampled points and the cuts. It would also be interesting to study the above properties (compactness, sampled points dense etc) for a fixed tree.

In the case of α -stable trees, there are also interesting unanswered questions; here are two which we find worthy of study, one quite concrete and the other rather vague.

6. What is the law of the cut-tree of an α -stable tree if the driving Poisson process has intensity $\lambda \otimes dt$, where λ is the length measure? In other words, what happens if cuts fall uniformly on the skeleton rather than at branch points?
7. The map Φ almost surely reconstructs (T, U, \mathcal{P}) from (C, \mathbb{N}, R) . Is Φ stable under small perturbations of (C, \mathbb{N}, R) ? To formalise such a statement, one would need to define a more robust reconstruction map F , presumably extending the definition of Φ . Having found an appropriate generalisation, the question is then whether F has the property that if $(C_k, \mathbb{N}, R_k) \xrightarrow{\text{a.s.}} (C, \mathbb{N}, R)$ as $k \rightarrow \infty$ then $F(C_k, \mathbb{N}, R_k) \xrightarrow{\text{a.s.}} F(C, \mathbb{N}, R)$.

A stability statement such as the second one would allow one to deduce distributional information about a random tree from information about its cut tree. The next and final section of the paper describes a concrete situation in which this would be useful: a model of discrete random trees with a complicated law, but for which a tree obtained by the discrete version of the reconstruction map has a simple and explicit description.

5.1. A stationary tree aggregation process

The discrete process of reconstruction described in Section 1.3 arises in a somewhat different setting, which provides an additional motivation for its study. (This arose in discussions of the third author with Edward Crane, Nic Freeman, James Martin, Bálint Tóth and Dominic Yeo.) We describe a rooted tree-valued process which grows until it becomes infinite, and is then “burnt” back to the root. This is intended to model a mean-field forest fire process (see [24, 51]), viewed from the perspective of a particular vertex, but the precise details of this interpretation are unnecessary here.

Fix a probability distribution \mathcal{W} on the set of rooted trees and consider a rooted tree-valued Markov process $(T_0(t), \rho)$, which evolves as follows. Start from a single vertex $T_0(0) = \rho$, the root. At any subsequent time t , given that the current state is $T_0(t) = (T, \rho)$, at rate given by the number of vertices of T , sample a rooted tree (T', r) from \mathcal{W} and an independent random vertex v from T and attach r to v by an edge, and root the resulting tree at ρ . It is possible for the jump-times $J_1 < J_2 < \dots$ of this process to accumulate (i.e. $J_m \rightarrow J_\infty$ as $m \rightarrow \infty$ for some $J_\infty < \infty$), in which case we kill it.

As long as $\mathbf{E}[J_\infty] < \infty$, it is standard from renewal theory that one can create a stationary version of this process by the following procedure. First generate a size-biased version J_∞^* of J_∞ . Given that $J_\infty^* = t$, generate a path $((T_0^{(t)}(s), \rho), 0 \leq s < t)$ which has the same law as $((T_0(s), \rho), 0 \leq s < J_\infty)$ conditioned on

$J_\infty = t$. Then finally take an independent uniform $[0, 1]$ random variable U and define $T(s) = T_0^{(t)}(Ut + s)$ for $s < (1 - U)t$. For $s \geq (1 - U)t$, simply concatenate independent copies of $((T_0(s), \rho), 0 \leq s < J_\infty)$ onto the end to yield a path $((T(s), \rho), s \geq 0)$.

It turns out that there is a unique law \mathcal{W} on the “environment” of rooted trees that we aggregate onto $T(t)$ such that $\mathbf{E}[J_\infty] < \infty$ and also $T(0) \sim \mathcal{W}$ (since $(T(t), t \geq 0)$ is stationary, this is also the law of $T(t)$ for any $t > 0$). This law is awkward to describe fully, but it has the property that if $(T, \rho) \sim \mathcal{W}$ then

$$\mathbf{P}\{|T| = k\} = \frac{2}{k} \binom{2k-2}{k-1} 4^{-k}.$$

Moreover, conditionally on $|T| = k$, if v is picked uniformly from among the k vertices of T then (T, v) has the same distribution as (T, ρ) (i.e. T is invariant under random re-rooting).

Much easier to describe is the distribution of the *genealogical tree* $G(t)$ associated with $T(t)$ via the aggregation process. This is an analogue of the cut-tree, where rather than thinking about edge-removal causing fragmentation we have edge-addition causing coalescence. For this it is useful to imagine an enriched version of the above process, in which the edges of the sampled trees are also marked with “arrival times”. The correct distribution for these marks may be deduced from the construction of $T(t)$.

The genealogical tree $G(t)$ is a binary tree whose leaves correspond to vertices of $T(t)$ and whose internal vertices correspond to edges of $T(t)$. The root of $G(t)$ corresponds to the most recent edge to have appeared in $T(t)$. The two subtrees hanging off the internal vertex corresponding to an edge e are the genealogical trees of the two clusters which were joined together by e .

The stationarity of $T(t)$ induces stationarity for $G(t)$ and, in particular, for all t , $G(t)$ has the law of a critical binary Galton–Watson tree. Indeed, for a given G with $k - 1$ internal nodes and k leaves, we have

$$\mathbf{P}\{G(t) = G\} = \left(\frac{1}{2}\right)^{2k-1}$$

and there are

$$\frac{1}{k} \binom{2k-2}{k-1}$$

such trees G .

How does one obtain the tree T from its genealogical tree G ? Once again we need to mark the internal vertices of G with the labels of the edges to which they correspond, after which we perform the reconstruction precisely as described in Section 1.3 for the cut-tree. Moreover, because of the re-rooting invariance of a tree sampled according to \mathcal{W} , it turns out that the two end-points of the edge marking a particular internal vertex of G are uniformly distributed among the leaves of G in the two subtrees of G hanging off that internal vertex.

Conditional on having k leaves, $G(t)$ converges in distribution in the Gromov–Hausdorff–Prokhorov sense to a constant times the Brownian CRT, once its edge-lengths are rescaled by $k^{-1/2}$ and it is endowed with the uniform measure [43, 52]. The law of the signposts in $G(t)$ is uniform on the relevant subtrees, which is precisely the discrete analogue of the law of the signposts in the Brownian CRT. It is then natural to conjecture that, conditional on $|T(t)| = k$, a rescaled version of $T(t)$ also converges in distribution to the Brownian CRT. There are at least two proofs of this fact due to Edward Crane [23]; if an appropriately defined reconstruction map were known to be stable, this would provide a computation-free proof of the same result.

Acknowledgements

C.G.’s research was supported in part by EPSRC grant EP/J019496/1 and in part by EPSRC fellowship EP/N004833/1. L.A.B’s research was supported in part by an NSERC Discovery Grant and in part by the Leverhulme Trust Visiting Professorship. We are very grateful to Edward Crane for sharing the results of [23] with us. We would like to warmly thank the two referees for their careful reading of and insightful comments on an earlier version, which greatly improved the paper.

References

- [1] R. Abraham and J.-F. Delmas. The forest associated with the record process on a Lévy tree. *Stochastic Process. Appl.*, 123(9):3497–3517, 2013.
- [2] R. Abraham, J.-F. Delmas, and P. Hoscheit. A note on the Gromov-Hausdorff-Prokhorov distance between (locally) compact metric measure spaces. *Electron. J. Probab.*, 18:no. 14, 21 pp. (electronic), 2013.
- [3] L. Addario-Berry, N. Broutin, and C. Holmgren. Cutting down trees with a Markov chainsaw. *Ann. Appl. Probab.*, 24(6):2297–2339, 2014.
- [4] M. Albenque and C. Goldschmidt. The Brownian continuum random tree as the unique solution to a fixed point equation. *Electron. Commun. Probab.*, 20:no. 61, 14 pp. (electronic), 2015.
- [5] D. Aldous. The continuum random tree. I. *Ann. Probab.*, 19(1):1–28, 1991.
- [6] D. Aldous. The continuum random tree II: an overview. In *Stochastic Analysis*. Cambridge University Press, 1991.
- [7] D. Aldous. The continuum random tree III. *Ann. Probab.*, 21(1):248–289, 1993.
- [8] D. Aldous. Recursive self-similarity for random trees, random triangulations and Brownian excursion. *Ann. Probab.*, 22(2):527–545, 1994.
- [9] D. Aldous and J. Pitman. The standard additive coalescent. *Ann. Probab.*, 26(4):1703–1726, 1998.
- [10] D. Aldous and J. Pitman. A family of random trees with random edge lengths. *Random Structures Algorithms*, 15(2):176–195, 1999.
- [11] D. Aldous and J. Pitman. Inhomogeneous continuum random trees and the entrance boundary of the additive coalescent. *Probab. Theory Related Fields*, 118(4):455–482, 2000.
- [12] D. J. Aldous and A. Bandyopadhyay. A survey of max-type recursive distributional equations. *Ann. Appl. Probab.*, 15(2):1047–1110, 2005.
- [13] E. Baur and J. Bertoin. Cutting edges at random in large recursive trees. In *Stochastic analysis and applications 2014*, volume 100 of *Springer Proc. Math. Stat.*, pages 51–76. Springer, Cham, 2014.
- [14] J. Bertoin. Self-similar fragmentations. *Ann. Inst. Henri Poincaré Probab. Stat.*, 38(3):319–340, 2002.
- [15] J. Bertoin. Fires on trees. *Ann. Inst. Henri Poincaré Probab. Stat.*, 48(4):909–921, 2012.
- [16] J. Bertoin. The cut-tree of large recursive trees. *Ann. Inst. Henri Poincaré Probab. Stat.*, 51(2):478–488, 2015.
- [17] J. Bertoin and G. Miermont. The cut-tree of large Galton-Watson trees and the Brownian CRT. *Ann. Appl. Probab.*, 23(4):1469–1493, 2013.
- [18] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular variation*, volume 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1989.
- [19] N. Broutin and M. Wang. Cutting down \mathbf{p} -trees and inhomogeneous continuum random trees. *Bernoulli*, 23(4A):2380–2433, 2017.
- [20] N. Broutin and M. Wang. Reversing the cut tree of the Brownian continuum random tree. *Electron. J. Probab.*, 22:no. 80, 23 pp. (electronic), 2017.
- [21] D. Burago, Y. Burago, and S. Ivanov. *A course in metric geometry*, volume 33 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001.
- [22] M. Camarri and J. Pitman. Limit distributions and random trees derived from the birthday problem with unequal probabilities. *Electron. J. Probab.*, 5:no. 2, 18 pp. (electronic), 2000.
- [23] E. Crane. Cluster growth in the steady state of the Ráth-Tóth forest fire model. In preparation, 2016+.
- [24] E. Crane, N. Freeman, and B. Tóth. Cluster growth in the dynamical Erdős-Rényi process with forest fires. *Electron. J. Probab.*, 20:no. 101, 33 pp. (electronic), 2015.
- [25] D. Dieuleveut. The vertex-cut-tree of Galton-Watson trees converging to a stable tree. *Ann. Appl. Probab.*, 25(4):2215–2262, 2015.
- [26] M. Drmota, A. Iksanov, M. Moehle, and U. Roesler. A limiting distribution for the number of cuts needed to isolate the root of a random recursive tree. *Random Structures Algorithms*, 34(3):319–336, 2009.
- [27] T. Duquesne and J.-F. Le Gall. *Random trees, Lévy processes and spatial branching processes*, volume 281, volume 281 of *Asterisque*. SMF, 2002.
- [28] T. Duquesne and J.-F. Le Gall. Probabilistic and fractal aspects of Lévy trees. *Probab. Theory Related Fields*, 131(4):553–603, 2005.
- [29] R. Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.
- [30] R. Durrett and T. M. Liggett. Fixed points of the smoothing transformation. *Z. Wahrsch. Verw. Gebiete*, 64(3):275–301, 1983.
- [31] S. N. Evans, J. Pitman, and A. Winter. Rayleigh processes, real trees, and root growth with re-grafting. *Probab. Theory Related Fields*, 134(1):81–126, 2006.
- [32] S. N. Evans and A. Winter. Subtree prune and regraft: a reversible real tree-valued Markov process. *Ann. Probab.*, 34(3):918–961, 2006.
- [33] J. A. Fill, N. Kapur, and A. Panholzer. Destruction of very simple trees. *Algorithmica*, 46(3-4):345–366, 2006.
- [34] K. Fukaya. Collapsing of Riemannian manifolds and eigenvalues of Laplace operator. *Invent. Math.*, 87(3):517–547, 1987.
- [35] C. Goldschmidt and B. Haas. A line-breaking construction of the stable trees. *Electron. J. Probab.*, 20:no. 16, 24 pp. (electronic), 2015.
- [36] A. Greven, P. Pfaffelhuber, and A. Winter. Convergence in distribution of random metric measure spaces (Λ -coalescent measure trees). *Probab. Theory Related Fields*, 145(1):285–322, 2009.
- [37] M. Gromov. *Metric structures for Riemannian and non-Riemannian spaces*. Modern Birkhäuser Classics. Birkhäuser

- Boston, Inc., Boston, MA, English edition, 2007. Based on the 1981 French original, With appendices by M. Katz, P. Pansu and S. Semmes, Translated from the French by Sean Michael Bates.
- [38] B. Haas and G. Miermont. Scaling limits of Markov branching trees with applications to Galton-Watson and random unordered trees. *Ann. Probab.*, 40(6):2589–2666, 2012.
 - [39] B. Haas, J. Pitman, and M. Winkel. Spinal partitions and invariance under re-rooting of continuum random trees. *Ann. Probab.*, 37(4):1381–1411, 2009.
 - [40] C. Holmgren. Random records and cuttings in binary search trees. *Combin. Probab. Comput.*, 19(3):391–424, 2010.
 - [41] A. Iksanov and M. Möhle. A probabilistic proof of a weak limit law for the number of cuts needed to isolate the root of a random recursive tree. *Electron. Comm. Probab.*, 12:28–35, 2007.
 - [42] S. Janson. Random cutting and records in deterministic and random trees. *Random Structures Algorithms*, 29(2):139–179, 2006.
 - [43] I. Kortchemski. Invariance principles for Galton-Watson trees conditioned on the number of leaves. *Stochastic Process. Appl.*, 122(9):3126–3172, 2012.
 - [44] A. Meir and J. W. Moon. Cutting down random trees. *J. Austral. Math. Soc.*, 11:313–324, 1970.
 - [45] A. Meir and J. W. Moon. Cutting down recursive trees. *Math. Biosci.*, 21:173–181, 1974.
 - [46] G. Miermont. Self-similar fragmentations derived from the stable tree. I. Splitting at heights. *Probab. Theory Related Fields*, 127(3):423–454, 2003.
 - [47] G. Miermont. Self-similar fragmentations derived from the stable tree. II. Splitting at nodes. *Probab. Theory Related Fields*, 131(3):341–375, 2005.
 - [48] G. Miermont. Tessellations of random maps of arbitrary genus. *Ann. Sci. Éc. Norm. Supér. (4)*, 42(5):725–781, 2009.
 - [49] A. Panholzer. Cutting down very simple trees. *Quaest. Math.*, 29(2):211–227, 2006.
 - [50] J. Pitman. *Combinatorial stochastic processes*, volume 1875 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2006. Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002, With a foreword by Jean Picard.
 - [51] B. Ráth and B. Tóth. Erdős-Rényi random graphs + forest fires = self-organized criticality. *Electron. J. Probab.*, 14:no. 45, 1290–1327, 2009.
 - [52] D. Rizzolo. Scaling limits of Markov branching trees and Galton-Watson trees conditioned on the number of vertices with out-degree in a given set. *Ann. Inst. Henri Poincaré Probab. Stat.*, 51(2):512–532, 2015.