Self-similar real trees defined as fixed-points and their geometric properties *

Nicolas Broutin[†] Henning Sulzbach[‡]

October 18, 2016

Abstract

We consider fixed-point equations for probability measures charging measured compact metric spaces that naturally yield continuum random trees. On the one hand, we study the existence, the uniqueness of the fixed-points and the convergence of the corresponding iterative schemes. On the other hand, we study the geometric properties of the random measured real trees that are fixed-points, in particular their fractal properties. We obtain bounds on the Minkowski and Hausdorff dimension, that are proved tight in a number of applications, including the very classical continuum random tree, but also for the dual trees of random recursive triangulations of the disk introduced by Curien and Le Gall [*Ann Probab*, vol. 39, 2011]. The method happens to be especially powerful to treat cases where the natural mass measure on the real tree only provides weak estimates on the Hausdorff dimension.

1 Introduction

Since the pioneering work of Aldous [4, 6] who introduced the Brownian continuum random tree as a scaling limit for uniformly random labelled trees, similar objects have been shown to play a crucial role in a number of limits of combinatorial problems that relate to computer science, physics or biology. These objects are real trees, or tree-like compact metric spaces (see Section 3 for a formal definition), and they are usually equipped with a probability measure that yields a notion of "mass". They quite naturally appear when studying asymptotic properties of discrete combinatorial or probabilistic objects that are intrinsically "branching" or recursive such as branching processes and fragmentation processes. More surprisingly, further prominent examples are that of random maps [21, 40, 44, 51] and of Liouville quantum gravity [27] that would a priori not be expected to relate to tree structures.

In a number of cases, these continuous objects, or more precisely their distributions, happen to satisfy a stochastic fixed-point equation; such fixed-point equations are often formulated in terms of the distribution of functions (later referred to as height functions) that encode the trees. One may think in particular of the Brownian continuum random tree [7], of trees that are dual to recursive triangulations of the disk [23], but also of the genealogies of self-similar fragmentations [36]. We will be more precise about the equations we consider shortly, but it is nonetheless informative to fix ideas: informally, a *distributional fixed-point* equation for random variable X taking values in some Polish space $\mathbb{S}$ of "objects" is an equation of the form

$$X \stackrel{d}{=} T((X_i)_{i\ge 1}, \Xi), \tag{1}$$

where $(X_i)_{i\ge 1}$ is a family of independent and identically distributed copies of X, T is a suitable map, and Ξ incorporates additional external randomness. (A precise formulation of such an equation for

*NB acknowledges the support of the grant ANR-14-CE25-0014 (ANR GRAAL). The research of HS was supported by the FSMP, reference: ANR-10-LABX-0098, and a Feodor Lynen Research Fellowship of the Alexander von Humboldt Foundation.

[†]Inria, 2 rue Simone Iff, CS 42112, 75589 Paris Cedex 12, France. Email: nicolas.broutin@inria.fr

[‡]McGill University, 3480 University Street, H3A 0E9 Montreal, QC, Canada. Email: henning.sulzbach@gmail.com

random metric spaces is more involved. See display (4) below.) The fact that natural limit objects satisfy equations such as the one in (1) raises many questions about the properties of such equations, and of their possible fixed-points:

(i) *Under which conditions does there exist a fixed-point ?*

(ii) *Under which conditions is this fixed-point unique ?*

(iii) *Can the fixed-point be obtained by some iterative procedure ?*

The answers to these questions of course depend on the space $\mathbb{S}$ that is considered, and some special care is needed in specifying this space.

Although there is a strong bond between the study of random real trees and that of random functions, the "random function" approach may sometimes appear as less intrinsic and somewhat artificial. Still, the construction of such encoding functions is always an important question since it "flattens" the metric space by giving a representation in terms of a continuous height function of a single real parameter. In particular, the "functional" point of view usually provides a nice point of entry because of the large collection of tools that are available in function spaces, but also because of the increased control on measurable maps in such spaces.

On the other hand, one of the striking features of random real trees that appear ubiquitous is their fractal nature, that one may quantify via various fractal dimensions giving an idea of the "amount of matter" within small balls in the intrinsic metric. Aside from the lack of embedding in a Euclidean space, this is similar to classical problems in stochastic geometry such as random Cantor sets and Koch curves ([33, Chapter 15]), the fractal percolation process [19] and various quantities related to Brownian motion and other Lévy processes. Among the most classical real trees one may cite the Lévy trees (including the Brownian continuum random trees and the stable trees), which are the scaling limits of rescaled Galton–Watson processes, and whose fractal properties have been investigated by Duquesne and Le Gall [29, 30] and Picard [48]. Another important example is that of the fragmentation trees, encoding certain self-similar fragmentation processes whose fractal properties have been studied by Haas and Miermont [36] and more recently by Stephenson [52]. In view of the recursive self-similarity of Equation (1), this raises an additional question about the geometry of the fixed-points:

(iv) *Can one quantify the fractal dimensions of the fixed-points ?*

Finally, observe that, for instance, the Brownian continuum random tree is binary, in the sense that the removal of any point disconnects the space into 1, 2 or 3 connected components with probability one (the number of connected components is called the degree of the point removed). This is to compare with the classical decomposition of the Brownian CRT into three pieces [3, 7]. Another example we have already mentioned, dual trees of recursive triangulations of the disk happen to have maximal degree three, while the natural fixed-point equation they satisfy only uses two pieces. These considerations raise yet another question about the geometry of solutions to equations such as (1):

(v) *Can one fully characterize the degrees of points in fixed-points from the equation ?*

Our aim in this paper is to provide answers to questions (i)–(v) in a general framework in which the limit objects are some classes of measured real trees. This framework allows for instance to deal with certain recalcitrant cases where the natural height function for the tree is not a "good" encoding, in the sense that its optimal Hölder exponent does not yield the fractal dimension of the metric space (we will be more precise shortly). At this point, let us mention that questions (i), (ii) and (iii) have recently been studied by Albenque and Goldschmidt [3] for the specific example where the fixed-point equation is the one described by Aldous in [7] and that is satisfied by the Brownian CRT. In passing, our results answer a question in [3] regarding point (iii) and the convergence to the (non-unique, but natural) fixed-point. Other applications of our results concern trees arising as scaling limits in the problem of recursive triangulations of the disk (see [23] and [17]), and but also other natural generalizations.

Organization of the paper. The paper is organized as follows: In Section 2, we first give the relevant background on the objects, metrics and spaces, and geometric properties we use in the document; we then introduce the precise setting for the recursive equations we consider, and the corresponding functional point of view. Section 3 is devoted to the statements of our main results; it also contains a sample of applications and an overview of the techniques we use. Section 4 contains the proofs of the results about existence and uniqueness of solutions to our recursive equations, as well the behaviour of iterative schemes. Section 5 contains the proofs of the geometric properties of the fixed-points. Finally, Section 6 is devoted to applications.

2 Settings and preliminaries

2.1 Spaces, metrics and convergence

Throughout the paper, we always assume that metric spaces are *compact*. General references on the topics that we are about to discuss include [18, 34, 35].

THE GROMOV–HAUSDORFF–PROKHOROV TOPOLOGY. For two compact metric spaces (X, d) and (X', d'), the *Gromov–Hausdorff distance* $\mathrm{d_{GH}}(X, X')$ is defined as

$$\mathrm{d_{GH}}((X, d), (X', d')) = \inf_{Z,\phi,\phi'} \mathrm{d_H^Z}(\phi(X), \phi'(X')), \tag{2}$$

where the infimum is taken over all compact metric spaces (Z, d^Z), and isometries $\phi : X \to Z$ and $\phi' : X' \to Z$. Here, $\mathrm{d_H^Z}$ denotes the Hausdorff distance in Z, that is

$$\mathrm{d_H^Z}(A, B) = \inf\{\varepsilon > 0 : A \subseteq B^\varepsilon \text{ and } B \subseteq A^\varepsilon\},$$

with $A^\varepsilon = \{x \in Z : d^Z(x, A) < \varepsilon\}$. If (X, d) and (X', d') are isometric, then $\mathrm{d_{GH}}((X, d), (X', d')) = 0$. The semi-metric $\mathrm{d_{GH}}$ induces a metric on the set $\mathbb{K}^{\mathrm{GH}}$ of isometry classes of compact metric spaces and turns this set into a Polish space, see, e.g. [39, Theorem 2.1].

A compact rooted (or pointed) measured metric space (X, d, μ, ρ) is a compact metric space (X, d) endowed with a probability measure μ and one distinguished point ρ. For two compact rooted measured metric spaces $\mathfrak{X} = (X, d, \mu, \rho)$ and $\mathfrak{X}' = (X', d', \mu', \rho')$ we define the *Gromov–Hausdorff–Prokhorov* distance by

$$\mathrm{d_{GHP}}(\mathfrak{X}, \mathfrak{X}') = \inf_{Z,\phi,\phi'} \left\{ d^Z(\phi(\rho), \phi'(\rho')) + \mathrm{d_H^Z}(\phi(X), \phi'(X')) + \mathrm{d_P^Z}(\phi_*(\mu), \phi'_*(\mu')) \right\}.$$

Here, the infimum is to be understood as in (2), $\phi_*(\mu)$ is the push-forward of μ, and $\mathrm{d_P^Z}$ denotes the Prokhorov metric on the set of probability measures on Z, that is

$$\mathrm{d_P^Z}(\mu, \nu) = \inf\{\varepsilon > 0 : \mu(A) \le \nu(A^\varepsilon) + \varepsilon \text{ and } \nu(A) \le \mu(A^\varepsilon) + \varepsilon \text{ for all measurable sets } A\}.$$

We call $\mathfrak{X}$ and $\mathfrak{X}'$ GHP-isometric if there exist an isometry ϕ between X and X' that maps ρ to ρ' and such that $\phi_*(\mu) = \mu'$. Clearly, if $\mathfrak{X}$ and $\mathfrak{X}'$ are GHP-isometric, then $\mathrm{d_{GHP}}(\mathfrak{X}, \mathfrak{X}') = 0$. The distance $\mathrm{d_{GHP}}$ induces a metric on the set $\mathbb{K}^{\mathrm{GHP}}$ of GHP-isometry classes of compact rooted measured metric spaces that turns it into a Polish space [1].

THE GROMOV–PROKHOROV TOPOLOGY. Analogously to the Gromov–Hausdorff–Prokhorov distance, for two compact rooted measured metric spaces $\mathfrak{X} = (X, d, \mu, \rho)$ and $\mathfrak{X}' = (X', d', \mu', \rho')$, we define

$$\mathrm{d_{GP}}(\mathfrak{X}, \mathfrak{X}') = \inf_{Z,\phi,\phi'} \left\{ d^Z(\phi(\rho), \phi'(\rho')) + \mathrm{d_P^Z}(\phi_*(\mu), \phi'_*(\mu')) \right\}.$$

We call $\mathfrak{X}, \mathfrak{X}'$ GP-isometric if $\mathrm{d_{GP}}(\mathfrak{X}, \mathfrak{X}') = 0$ which happens to be the case if and only if there exists an isometry from $\mathrm{supp}(\mu)$ to $\mathrm{supp}(\mu')$ that maps ρ to ρ'. Endowed with $\mathrm{d_{GP}}$, the set $\mathbb{K}^{\mathrm{GP}}$ of GP-isometry classes of compact rooted measured metric spaces becomes a Polish space [34]. In general, GP-equivalence classes in $\mathbb{K}^{\mathrm{GP}}$ contain spaces that are not GHP-isometric. But when both μ and μ' have full support, then $\mathfrak{X}$ and $\mathfrak{X}'$ are GP-isometric if and only if they are GHP-isometric. Thus, if we denote by $\mathbb{K}_{\mathrm{f}}^{\mathrm{GHP}}$ the set of GHP-isometry classes of compact rooted measured metric spaces satisfying

C1. $\mathrm{supp}(\mu) = X$,

then, there exists a natural bijection ι between the spaces $\mathbb{K}_{\mathrm{f}}^{\mathrm{GHP}}$ and $\mathbb{K}^{\mathrm{GP}}$. Note that, for any $\delta > 0$, the quantity

$$\kappa_\delta(\mathfrak{X}) = \inf\{\mu(\{y \in X : d(x, y) \leq \delta\}) : x \in X\},$$

only depends on the GHP-equivalence class of $\mathfrak{X}$. Moreover,

$$\mathbb{K}_{\mathrm{f}}^{\mathrm{GHP}} = \{\mathfrak{X} \in \mathbb{K}^{\mathrm{GHP}} : \kappa_\delta(\mathfrak{X}) > 0 \,\forall\, \delta > 0\}.$$

A straightforward application of the Portemanteau lemma shows that κ_δ is upper semi-continuous with respect to $\mathrm{d_{GHP}}$. Hence, it follows easily that $\mathbb{K}_{\mathrm{f}}^{\mathrm{GHP}}$ is a measurable set.

Clearly, the topology generated by $\mathrm{d_{GP}}$ is coarser than the one generated by $\mathrm{d_{GHP}}$ even if we restrict our attention to measures with full support: indeed, the map ι^{-1} is not continuous and the space $\mathbb{K}_{\mathrm{f}}^{\mathrm{GHP}}$ (considered as subspace of $\mathbb{K}^{\mathrm{GHP}}$) is not complete. However, it is important to note that, by [12, Corollary 5.6], $\mathbb{K}_{\mathrm{f}}^{\mathrm{GHP}}$ endowed with the relative topology generated by $\mathrm{d_{GHP}}$ is Polish. It follows from the Lusin–Souslin theorem, see e.g. [38, Theorem 15.1] that ι^{-1} is measurable. Hence, we can and will consider any random variable with values in $\mathbb{K}_{\mathrm{f}}^{\mathrm{GHP}}$ also as random variable in $\mathbb{K}^{\mathrm{GP}}$. (We thank Stephan Gufler for pointing out the short argument showing the measurability of ι^{-1}.)

We conclude the section by two short comments. First, with respect to some concepts treated in this paper, the assumption on compactness of metric spaces can be relaxed to local compactness, see, e.g. [1]. Second, in the context of technical results concerning the framework introduced, we use results from the works [26, 34, 35] which only treat the case of un-rooted compact measured metric spaces. Incorporating a root vertex only generates marginal modifications which we do not discuss in detail.

2.2 Real trees, continuum trees and recursive decompositions

In this paper, we are interested in a certain class of metric spaces that are tree-like.

REAL TREES. A metric space $(\mathcal{T}, d)$ is a called *real tree* if it has the following properties:

i) for every $x, y \in \mathcal{T}$ there exists a unique isometry $\varphi_{x,y} : [0, d(x, y)] \to \mathcal{T}$ with $\varphi_{x,y}(0) = x$ and $\varphi_{x,y}(d(x, y)) = y$, (we write $[\![x, y]\!] := \varphi_{x,y}([0, d(x, y)])$ for the *segment* between x and y in $\mathcal{T}$),

ii) if $q : [0, 1] \to \mathcal{T}$ is a continuous and injective map with $q(0) = x, q(1) = y$, then $q([0, 1]) = [\![x, y]\!]$.

We denote by $\mathbb{T}^{\mathrm{GH}}$ the closed subset of $\mathbb{K}^{\mathrm{GH}}$ consisting of isometry classes of compact real trees. For a compact real tree $(\mathcal{T}, d)$ and $x \in \mathcal{T}$, we denote by $\deg(x)$ the number of connected components of $\mathcal{T} \setminus \{x\}$. (By compactness, this number is at most countably infinite.) We call $x \in \mathcal{T}$ a *leaf* if $\deg(x) = 1$, and abbreviate $\mathscr{L}$ for the set of leaves. Note that $\mathscr{L}$ is totally disconnected. We call $x \in \mathcal{T}$ a *branch point* if $\deg(x) \geq 3$ and note that, again by compactness, the set of branch points $\mathscr{B}$ is at most countable. Finally, we let $\mathscr{S} = \mathcal{T} \setminus \mathscr{L}$ be the *skeleton* of $\mathcal{T}$. Then, $\mathscr{S}$ is dense in $\mathcal{T}$ and, unless we are in the trivial case for which $\mathcal{T} = \{\rho\}$, $\mathscr{S}$ is uncountable.

MEASURED AND CONTINUUM REAL TREES. A compact rooted measured real tree $\mathfrak{T} = (\mathcal{T}, d, \mu, \rho)$ is a compact real tree $(\mathcal{T}, d)$ endowed with a probability measure μ and a distinguished point $\rho \in \mathcal{T}$ called the root. We call $\|\mathcal{T}\| := \sup\{d(x, \rho) : x \in \mathcal{T}\}$ the *height* of $\mathfrak{T}$. By $\mathbb{T}^{\mathrm{GHP}} \subseteq \mathbb{K}^{\mathrm{GHP}}$ we denote the closed subset of GHP-isometry classes of compact rooted measured real trees. As we have already noted, there is a special interest in spaces carrying a measure with full support. Thus, we let $\mathbb{T}_{\mathrm{f}}^{\mathrm{GHP}} = \mathbb{K}_{\mathrm{f}}^{\mathrm{GHP}} \cap \mathbb{T}^{\mathrm{GHP}}$ and call elements in $\mathbb{T}_{\mathrm{f}}^{\mathrm{GHP}}$ *continuum real trees*. Note that both $\mathbb{K}_{\mathrm{f}}^{\mathrm{GHP}}$ and $\mathbb{T}_{\mathrm{f}}^{\mathrm{GHP}}$ are non-closed subsets of $\mathbb{K}^{\mathrm{GHP}}$. In the literature on continuum real trees, see, e.g. [3, 6], one often finds the following two additional conditions:

C2. μ has no atoms,

C3. $\mu(\mathscr{L}) = 1$,

Note that **C2** and **C3** imply $\mathscr{L}$ to be uncountable, and **C1** and **C3** imply $\mathscr{L}$ to be dense in $\mathcal{T}$. It turns out that all continuum trees playing a role in this paper satisfy both **C2** and **C3**. However, we emphasize the fact that we do not impose these conditions beforehand: they can be proved to hold as a non-trivial consequence of our setting.

REAL TREES ENCODED BY EXCURSIONS. One natural way to define real trees is via an encoding by continuous excursions (see e.g., [32, 39]). Let $\mathcal{C}$ be the space of continuous functions on $[0,1]$, which we always endow with the uniform norm $\|f\| = \sup_{t\in[0,1]} |f(t)|$. Let also $\mathcal{C}_{\text{ex}}$ denote the set of unit-length continuous excursions, that is the set of functions $f \in \mathcal{C}$ such that $f(0) = f(1) = 0$. For $f \in \mathcal{C}_{\text{ex}}$, define the semi-metric

$$d_f(x,y) := f(x) + f(y) - 2\inf\{f(s) : x \wedge y \leq s \leq x \vee y\}\,.$$

Let $\mathcal{T}_f = [0,1]/\!\sim$ where $x \sim y$ if and only if $d_f(x,y) = 0$. Then, the compact metric space $(\mathcal{T}_f, d_f)$ is a real tree, which we call the real tree encoded by f; we will also sometimes denote the continuous excursion f as being a *height process* for the real tree $\mathcal{T}_f$. We use $\mathscr{L}_f$, $\mathscr{B}_f$ and $\mathscr{S}_f$ for the sets of leaves, of branch points and for the skeleton of $\mathcal{T}_f$, respectively.

As noted in [39] (see also [28, Corollary 1.2]), for every compact real tree $(\mathcal{T}, d)$, there exists $f \in \mathcal{C}_{\text{ex}}$ such that $(\mathcal{T}, d)$ and $(\mathcal{T}_f, d_f)$ are isometric (hence, they represent the same element in $\mathbb{K}^{\text{GH}}$.) Two real trees $(\mathcal{T}_f, d_f)$ and $(\mathcal{T}_g, d_g)$, encoded by continuous excursions f and g respectively, are isometric, if, for instance, $f = g \circ \phi$ for $\phi : [0,1] \to [0,1]$ a continuous and strictly increasing function. In this case, we call the encoding excursions f and g *equivalent*. Of course, $(\mathcal{T}_f, d_f)$ and $(\mathcal{T}_g, d_g)$ may be isometric even if f and g are not equivalent.

The real tree $(\mathcal{T}_f, d_f)$ can be turned into a compact rooted measured real tree $\mathfrak{T}_f = (\mathcal{T}_f, d_f, \mu_f, \rho_f)$ using the push-forward measure $\mu_f := \text{Leb} \circ \pi_f^{-1}$, where Leb denotes the Lebesgue measure on $[0,1]$, and the root $\rho_f := \pi_f(0)$. Then, μ_f has full support, and thus $\mathfrak{T}_f$ satisfies **C1**. Denote by $\mathcal{C}^*_{\text{ex}}$ the set of continuous excursions which are not monotonic on any non-trivial interval. For $f \in \mathcal{C}^*_{\text{ex}}$, we can make the following observations,

- $\mathscr{L}_f \setminus \{\rho_f\} = \{x \in \mathcal{T}_f \setminus \{\rho_f\} : \#\pi_f^{-1}(x) = 1\}$,
- the set of local minima of f is dense in $[0,1]$,
- both $\mathscr{L}_f$ and $\mathscr{B}_f$ are dense in $\mathcal{T}_f$,
- $\mu_f(x) = 0$ for all $x \in \mathscr{L}_f$.

Note that there exist functions in $\mathcal{C}^*_{\text{ex}}$ which have no strict local minima on $(0,1)$ (see [49] for an explicit construction). Furthermore, for any $\varepsilon \in (0,1)$ and $x \in [0,1]$, one can construct a function $f \in \mathcal{C}^*_{\text{ex}}$ such that $\pi_f(x) \in \mathscr{B}_f$ and $\mu_f(\{\pi_f(x)\}) = 1 - \varepsilon$. By [28, Comment 1.1], for any compact rooted measured real tree $(\mathcal{T}, d, \mu, \rho)$ satisfying **C1 – C3**, there exists $f \in \mathcal{C}_{\text{ex}}$ such that $(\mathcal{T}_f, d_f, \mu_f, \rho_f)$ and $(\mathcal{T}, d, \mu, \rho)$ are GHP-isometric. Finally, it is well-known that the map $\tau : (\mathcal{C}_{\text{ex}}, \|\cdot\|) \to (\mathbb{T}_{\text{f}}^{\text{GHP}}, \text{d}_{\text{GHP}})$ such that $\tau(f) = \mathfrak{T}_f$ is continuous (even globally Lipschitz continuous by [1, Proposition 3.3]). Hence, for any random variable $X \in \mathcal{C}_{\text{ex}}$, the corresponding (GHP-equivalence class of the) compact rooted measured real tree $\mathfrak{T}_X$ is a random variable with values in $\mathbb{T}_{\text{f}}^{\text{GHP}}$. Similarly, for any random variable $\mathfrak{T}$ in $\mathbb{T}_{\text{f}}^{\text{GHP}}$, there exists a random variable $X \in \mathcal{C}_{\text{ex}}$, such that $\mathfrak{T}_X$ and $\mathfrak{T}$ are identically distributed. (See Lemma 14.)

2.3 Metric spaces described by recursive decompositions

We now introduce a general framework for random compact rooted measured metric spaces satisfying recursive distributional decompositions.

Let $K > 1$ and Γ be a rooted plane tree with K vertices. We call Γ the *structural tree* of the recursive decomposition, and accordingly, the decomposition of a space (or tree) will involve K subparts. Label the root of Γ by 1 and the remaining nodes in the *depth-first order*[1]. From now on, we will consider Γ as a tree on the set $[K] := \{1, 2, \ldots, K\}$. Furthermore, for $i \geq 2$, we denote by $\varpi_i \in \{1, 2, \ldots, i-1\}$ the label of the parent of node i. Next, fix $\alpha \in (0,1)$, and $\mathbf{r}, \mathbf{s} \in \Sigma_K := \{\mathbf{x} = (x_1, x_2, \ldots, x_K) \in (0,1)^K : x_1 + \ldots + x_K = 1\}$. We consider the following construction (see Section 1.4 in [3] for a related construction): given compact rooted measured metric spaces $(X_i, d_i, \mu_i, \rho_i)$, $i \in [K]$, construct a compact rooted measured metric space (X, d, μ, ρ) as follows:

i) Independently sample points $\eta_i \in X_i$, $i \in [K]$, according to the probability measures μ_i;

ii) Let X° denote the disjoint union of the X_i, $i = 1, \ldots, K$; let $\sim$ be the equivalence relation on X° in which $\rho_i \sim \eta_{\varpi_i}$, for $i = 2, \ldots, K$; define X as the quotient $X^\circ/\!\sim$ and write φ for the canonical injection from X° into X.

iii) Let d° be the maximal semi-metric on X° that is not greater than $r_i^\alpha d_i$ on X_i, and for which $d^\circ(x,y) = 0$ if $x \sim y$; define d as the metric induced on X by d°;

iv) Let μ° be the unique probability measure on X° that is compatible with $s_i \mu_i$ when restricted to X_i; define μ as the push-forward of μ° under φ.

v) Finally, let $\rho = \varphi(\rho_1)$ be the root.

Note that, because of the need to sample $\eta_1, \ldots, \eta_K$, the GHP-equivalence class of the resulting space $\mathfrak{X} = (X, d, \mu, \rho)$ is random. It is crucial to observe that its distribution only depends on the GHP-isometry classes of $\mathfrak{X}_1, \ldots, \mathfrak{X}_K$. Hence, denoting by $\mathcal{M}_1(\mathbb{K}^{\text{GHP}})$ the set of probability measures on $\mathbb{K}^{\text{GHP}}$, the map $\psi : (\mathbb{K}^{\text{GHP}})^K \times \Sigma_K^2 \to \mathcal{M}_1(\mathbb{K}^{\text{GHP}})$ (where Σ_K^2 incorporates the choice of $(\mathbf{r}, \mathbf{s})$) whose image is described by the construction above is well-defined. For probability measures τ on Σ_K^2 and $\aleph$ on $\mathbb{K}^{\text{GHP}}$, we define the annealed measure

$$\Psi(\aleph, \tau)(A) := \mathbf{E}\left[\psi(\mathfrak{X}_1, \ldots, \mathfrak{X}_K, \mathcal{R}, \mathcal{S})(A)\right], \quad A \subseteq \mathbb{K}^{\text{GHP}} \quad \text{measurable}, \tag{3}$$

where $\mathfrak{L}((\mathcal{R}, \mathcal{S})) = \tau$, $\mathfrak{L}(\mathfrak{X}_1) = \ldots = \mathfrak{L}(\mathfrak{X}_K) = \aleph$ and $(\mathcal{R}, \mathcal{S}), \mathfrak{X}_1, \ldots, \mathfrak{X}_K$ are independent.

Now given K, Γ, $\alpha \in (0,1)$ and a distribution τ on Σ_K^2, we are interested in laws on compact rooted measured metric spaces $\aleph \in \mathcal{M}_1(\mathbb{K}^{\text{GHP}})$ satisfying

$$\aleph = \Psi(\aleph, \tau). \tag{4}$$

Informally, equation (4) says the following: Starting with K independent copies of (isometry classes) of spaces distributed according to $\aleph$, the space obtained by the operation described in points i) – v) based on an independent vector $(\mathcal{R}, \mathcal{S})$ with distribution τ again has distribution $\aleph$. Therefore, we refer to (4) as a stochastic fixed-point equation at the level of compact rooted measured metric spaces. Note that, while we have introduced the map Ψ for distributions on the space of GHP-isometry classes of spaces, in the same way, Ψ can be defined relying on GP-isometry classes, and we will occasionally use Ψ in this sense.

Remark. Our aim is to study tree-like structures, but we emphasize the fact that, while the decomposition leading to (4) is clearly tree-like, it is not true that any distribution $\aleph$ on $\mathbb{K}^{\text{GHP}}$ that satisfies (4) charges only tree-like structures. In fact, there are distributions that are fixed-points and whose metric components are with probability one not real trees (see Proposition 3). It is however true that the support of the mass measure of such fixed-points is almost surely a real tree (this follows from our main result Theorem 1).

[1] One can see a rooted plane tree as a subset t of $\cup_{n\geq 0}\mathbb{N}^n$, such that (a) if a word $u \in t$ then all its prefixes also are in t and (b) if $ui \in t$ then also $uj \in t$ for $j = 1, \ldots, i-1$. The depth-first order on t is then order induced on t by the lexicographic order.

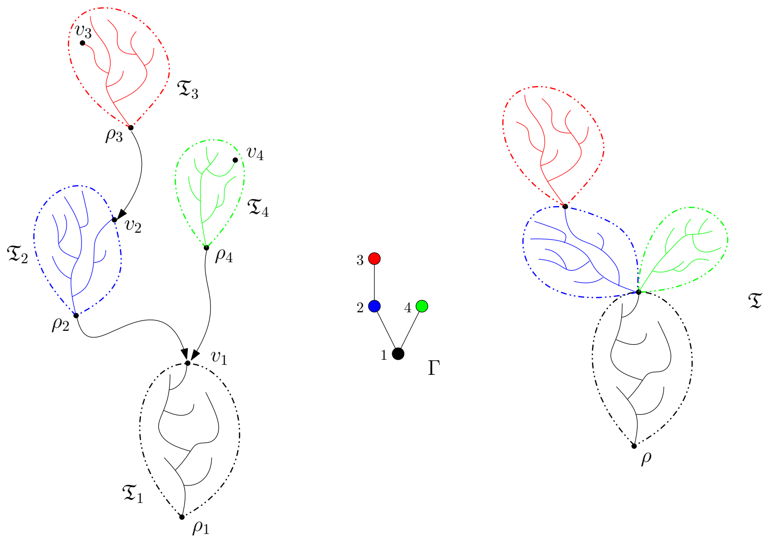

Figure 1: *The construction of $\mathfrak{X}$ with law $\psi(\mathfrak{X}_1, \mathfrak{X}_2, \mathfrak{X}_3, \mathfrak{X}_4, \mathcal{R}, \mathcal{S})$ with the structural tree Γ shown in the middle. For simplicity, we have not rescaled any of the distances. An excursion point of view is depicted in Figure 2. (For the sake of representation, we have chosen the spaces to be tree-like)*

Given K, Γ and $(\mathcal{R}, \mathcal{S})$, not every α is admissible for (4) to have non-trivial solutions. The parameter α is chosen as the unique value such that the depth of a point sampled according to μ has finite mean, see the discussion of (6) below. Here, this condition can be expressed as follows. Let Γ_i the set of nodes in the subtree of Γ rooted at i (including i itself). Define $\alpha \in (0,1)$ as the unique solution to

$$\mathbf{E}\left[\sum_{1 \le i \le K} \mathcal{R}_i^\alpha \mathbf{1}_{\{J \in \Gamma_i\}}\right] = 1 \quad \text{where} \quad \mathbf{P}(J = j \,|\, \mathcal{S}, \mathcal{R}) = \mathcal{S}_j, \quad 1 \le j \le K. \tag{5}$$

Such an α always exists by monotonicity and continuity in α of the expected value and the values for $\alpha \in \{0,1\}$. *From now on, unless specified otherwise, we will always assume that α has been chosen to satisfy* (5).

THE HEIGHT OF A RANDOM POINT. In a random compact rooted measured metric space $\mathfrak{X} = (X, d, \mu, \rho)$, heights of points sampled according to μ play an important role. For $1 \le i \le K$, write E_i for the set of nodes on the path from 1 to i in Γ (including 1, but excluding i). In our construction, with $\mathfrak{X}$ (or, rather its distribution) satisfying (4) and abbreviating $Y := d(\rho, \zeta)$ where ζ has distribution μ, we have

$$Y \overset{d}{=} \sum_{i=1}^{K} \mathbf{1}_{\{J=i\}} \left[\mathcal{R}_i^\alpha Y^{(i)} + \sum_{j \in E_i} \mathcal{R}_j^\alpha Y^{(j)}\right] \overset{d}{=} \sum_{i=1}^{K} \beta_i^\alpha Y^{(i)}, \quad \text{where} \quad \beta_i = \mathcal{R}_i \mathbf{1}_{\{J \in \Gamma_i\}}, \tag{6}$$

and $Y^{(1)}, \ldots, Y^{(K)}$ are distributed like Y and $\xi, (\mathcal{R}, \mathcal{S}), Y^{(1)}, \ldots, Y^{(K)}$ are independent. By Theorem 1 and Theorem 2 in [31], the distribution of $d(\rho, \zeta)$ is uniquely determined by (6) in the space of probability distributions on $[0, \infty)$ up to a multiplicative constant. In fact, these theorems further imply that, for any different choice of α, the random variable $d(\rho, \zeta)$ has either infinite mean, or, almost surely, $X = \{\rho\}$. Since $\sum_{i=1}^K \mathbf{E}[\beta_i^{2\alpha}] < 1$, a standard contraction argument (e.g. [50, Theorem 3]) shows that the distribution of $d(\rho, \zeta)$ is also determined by (6) in the space of distributions on $\mathbb{R}$ with finite variance, again up a to multiplicative constant. (In general, there exist further non-integrable solutions to (6), compare [10, Theorem 2.2].) Finally, we note that, from the last display, one can easily deduce that

$$\mathbf{P}\left(d(\rho, \zeta) > 0\right) \in \{0, 1\}. \tag{7}$$

EXAMPLES. **1)** The most celebrated example of a measured real tree that satisfies a fixed-point equation such as (4) is Aldous' Brownian continuum random tree [4, 6]. Its recursive structure has been investigated in [7], where it has been proved that it satisfies a fixed-point equation of the type (4), with $K = 3$, Γ is the tree on $\{1,2,3\}$ with 2 and 3 that are children of 1, $\mathcal{R} = \mathcal{S} = (\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3) \sim$ Dirichlet$(1/2, 1/2, 1/2)$, and in this case, $\alpha = 1/2$: starting with three independent Brownian continuum random trees $\mathfrak{T}_1$, $\mathfrak{T}_2$, and $\mathfrak{T}_3$, one again obtains a Brownian continuum random tree $\mathfrak{T}$ if one rescales the trees into $(\mathcal{T}_1, \mathcal{S}_1^{1/2} d_1, \mathcal{S}_1 \mu_1, \rho_1)$, $(\mathcal{T}_2, \mathcal{S}_2^{1/2} d_2, \mathcal{S}_2 \mu_2, \rho_2)$ and $(\mathcal{T}_3, \mathcal{S}_3^{1/2} d_3, \mathcal{S}_3 \mu_3, \rho_3)$, and then identifies the roots ρ_2, ρ_3 and a point sampled in $\mathcal{T}_1$ with distribution μ_1.

2) An instance has also appeared in the context of recursive triangulations of the disk [17, 23]. There, $K = 2$ (and thus Γ is the tree on $\{1,2\}$ where 2 is a child of 1), $\mathcal{S} = (\mathcal{S}_1, \mathcal{S}_2) \sim$ Dirichlet$(2,1)$, $\mathcal{R} = \mathcal{S}$ and α turns out to be given by $(\sqrt{17} - 3)/2$. In other words, starting with two independent trees $\mathcal{T}_{\mathscr{X}_1}, \mathcal{T}_{\mathscr{X}_2}$ distributed like $\mathcal{T}_{\mathscr{X}}$, this states that attaching $\mathcal{T}_{\mathscr{X}_2}$ rescaled by Δ_1^β at a uniformly random point of $\mathcal{T}_{\mathscr{X}_1}$ rescaled by Δ_2^β leads to tree with distribution $\mathcal{T}_{\mathscr{X}}$.

3) Another tree related to **2)** is given by $K = 2$, $\mathcal{S} = (\mathcal{S}_1, \mathcal{S}_2) \sim$ Dirichlet$(2,1)$, $\mathcal{R} \sim$ Dirichlet$(1,1)$ independent of $\mathcal{S}$, and for this case, one finds $\alpha = 1/3$. The tree has not been considered explicitly, but it appears in [17, 23] via one of its encoding processes.

2.4 Recursive decompositions: An excursion point of view

It is convenient to express the construction in Section 2.3 in terms of excursions. Let $K, \Gamma, \mathbf{r}, \mathbf{s}$ as in Section 2.3 and assume for now that $\alpha \in (0,1)$ is arbitrary. Each node of Γ will be assigned two intervals, except the leaves that will be assigned a single one. In the following, we write $\partial\Gamma$ for the set of leaves of Γ (the nodes i such that $\Gamma_i = \{i\}$), and Γ^o for $\Gamma \setminus \partial\Gamma$. Next, for $i \in [K]$, let $v_i = i + \#\{1 \le j < i : j \notin E_i \cup \partial\Gamma\}$ and define $V_i = \{v_i, v_i + 2|\Gamma_i| - \partial\Gamma_i - 1\}$. Observe that V_i contains a unique element if $i \in \partial\Gamma$, and two otherwise; furthermore, $(V_i)_{1 \le i \le K}$ forms a partition of $\{1, 2, \ldots, L\}$, where $L := 2K - |\partial\Gamma|$. For $\mathbf{u} = (u_i) \in (0,1)^K$, there is a unique decomposition of the unit interval into L half-open[2] intervals $I_1, \ldots, I_L$ such that,

- $s_i = \sum_{k \in V_i} \mathrm{Leb}(I_k)$ for all $i \in [K]$,
- $u_i = \mathrm{Leb}(I_{\min V_i})/s_i$ for all $i \in \Gamma^o$.

Now, define

$$\Lambda_i = \bigcup_{k \in V_i} I_k \qquad \text{and} \qquad \varphi_i : \Lambda_i \to [0,1] \tag{8}$$

as the unique function which is bijective, monotonically increasing and piecewise linear with constant slope. We can now define the operator $\Phi : \mathcal{C}_{\text{ex}}^K \times \Sigma_K^2 \times (0,1)^K \to \mathcal{C}_{\text{ex}}$, such that $g = \Phi(f_1, \ldots, f_k, \mathbf{r}, \mathbf{s}, \mathbf{u})$ is the unique excursion satisfying

$$g(x) - g(y) = r_{v_\ell}^\alpha \left[f_{v_\ell}(\varphi_{v_\ell}(x)) - f_{v_\ell}(\varphi_{v_\ell}(y)) \right], \tag{9}$$

for all $1 \le \ell \le L$ and $x, y \in I_\ell$. In other words, up to the scaling factor r_i^α in space, the function f_i is first fitted to an interval of length $s_i = \mathrm{Leb}(\Lambda_i)$ and then used on the set Λ_i. For an illustration see Figure 2 (and compare it with the corresponding version involving trees on Figure 1). By construction, for $f_1, \ldots, f_K \in \mathcal{C}_{\text{ex}}, r, s \in \Sigma_K$ and Ξ uniformly distributed on $(0,1)^K$, we have

$$\mathfrak{L}\left(\mathfrak{T}_{\Phi(f_1, \ldots, f_K, r, s, \Xi)}\right) = \psi(\mathfrak{T}_{f_1}, \ldots, \mathfrak{T}_{f_K}, r, s).$$

In particular, the distribution of the random compact rooted measured real tree $(\mathcal{T}_X, d_X, \mu_X, \rho_X)$ satisfies (4) if

$$X \stackrel{d}{=} \Phi\big(X^{(1)}, \ldots, X^{(K)}, \mathcal{R}, \mathcal{S}, \Xi\big), \tag{10}$$

[2] We call an interval half-open if it is of the form $(a, b]$ with $0 \le a \le b \le 1$ or $[0, a]$ with $0 \le a \le 1$.

where $X^{(1)}, \ldots, X^{(K)}$ are independent copies of X, independent of $(\mathcal{R}, \mathcal{S}, \Xi)$, Ξ and $(\mathcal{R}, \mathcal{S})$ being independent and $\Xi = (\xi_1, \ldots, \xi_K)$ being uniformly distributed on $(0,1)^K$; of course, in this case, α shall be chosen as in (5). The fixed-point equation in (10) can be expressed alternatively as

$$
\begin{aligned}
X(\cdot) &\stackrel{d}{=} \sum_{i=1}^{K} \mathbf{1}_{S_i}(\cdot) \left[\mathcal{R}_i^\alpha X^{(i)}(\varphi_i(\cdot)) + \sum_{j \in E_i} \mathcal{R}_j^\alpha X^{(j)}(\xi_j) \right] \\
&\stackrel{d}{=} \sum_{i=1}^{K} \mathcal{R}_i^\alpha \left[\mathbf{1}_{S_i}(\cdot) X^{(i)}(\varphi_i(\cdot)) + \sum_{j \in \Gamma_i \setminus \{i\}} \mathbf{1}_{S_j}(\cdot) X^{(i)}(\xi_i) \right].
\end{aligned}
$$

Let us note that, when asking for tree solutions to (4), the excursion point of view of the recursive decomposition is technically preferable since it grants access to random excursions (and their corresponding encoded real trees) as well as to nodes sampled independently according to the mass measure using the external randomness provided by Ξ. Distributional identities are used only in the final step of the formulation of the fixed-point equation (10) rather than arising intrinsically in the construction in Section 2.3 leading to a rigorous definition of the map Ψ in (4).

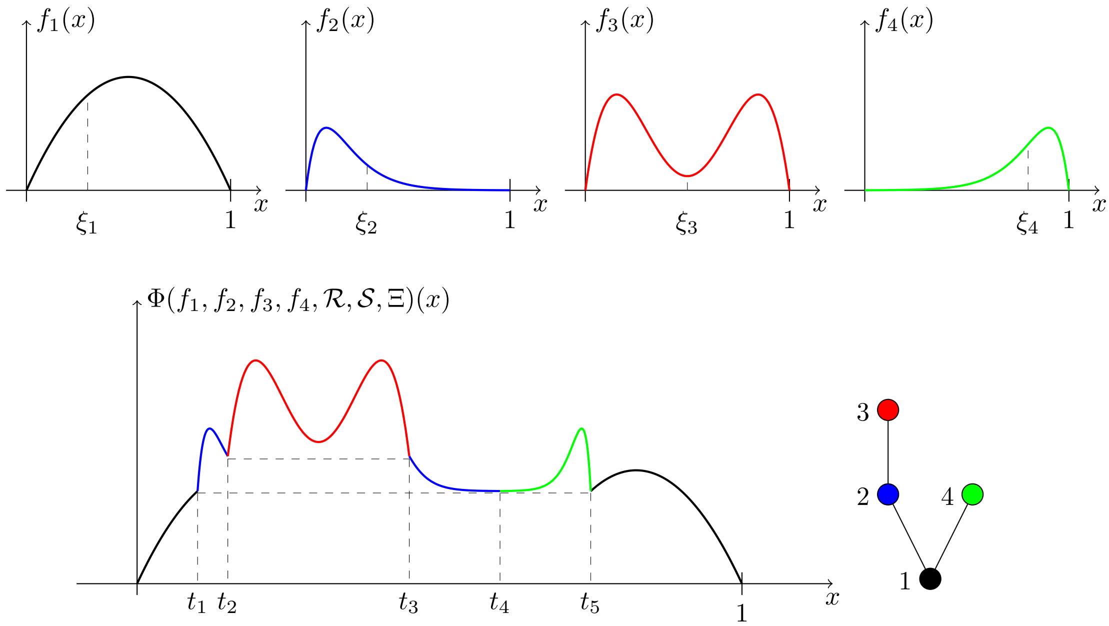

Figure 2: *An example of the functional construction of Section 2.4: Here $K = 4$, $L = 6$, the structural tree Γ is the tree shown on the bottom right. The functions f_1, f_2, f_3, f_4 are composed into $\Phi(f_1, f_2, f_3, f_4, \mathcal{R}, \mathcal{S}, \Xi)$. In order to keep the focus on the structure of the construction, we have not used scalings for the distances and the scaling $\mathcal{S} = (0.35, 0.20, 0.30, 0.15)$ for time. The corresponding point of view using trees is depicted in Figure 1.*

EXAMPLES. **1)** An identity of the kind in (10) holds for the Brownian excursion $\mathbf{e}$, and is of course intimately related to the corresponding decomposition of the Brownian continuum random tree. By [7, Corollary 3], we have

$$\mathbf{e}(\,\cdot\,) \stackrel{d}{=} \mathbf{1}_{[0,U_1)}(\,\cdot\,)\Delta_1^{1/2}\mathbf{e}^{(1)}\left(\frac{\cdot}{\Delta_1}\right) + \mathbf{1}_{[U_3,1]}(\,\cdot\,)\Delta_1^{1/2}\mathbf{e}^{(1)}\left(\frac{\cdot - \Delta_2 - \Delta_3}{\Delta_1}\right)$$
$$+ \mathbf{1}_{[U_1,U_2)}(\,\cdot\,)\left[\Delta_2^{1/2}\mathbf{e}^{(2)}\left(\frac{\cdot - U_1}{\Delta_2}\right) + \Delta_1^{1/2}\mathbf{e}^{(1)}(\xi)\right]$$
$$+ \mathbf{1}_{[U_2,U_3)}(\,\cdot\,)\left[\Delta_3^{1/2}\mathbf{e}^{(3)}\left(\frac{\cdot - U_2}{\Delta_3}\right) + \Delta_1^{1/2}\mathbf{e}^{(1)}(\xi)\right], \tag{11}$$

where $\mathbf{e}^{(1)}, \mathbf{e}^{(2)}, \mathbf{e}^{(3)}$ are distributed like $\mathbf{e}$, $(\Delta_1, \Delta_2, \Delta_3)$ is Dirichlet$(1/2, 1/2, 1/2)$ distributed, $U_1 = \xi\Delta_1$, $U_2 = U_1 + \Delta_2$, $U_3 = U_2 + \Delta_3$, ξ is uniformly distributed on $[0, 1]$ and $(\Delta_1, \Delta_2, \Delta_3)$, ξ, $\mathbf{e}^{(1)}, \mathbf{e}^{(2)}, \mathbf{e}^{(3)}$ are independent.

2) The functional version of the fixed-point equation appearing in Example **2)** of Section 2.3 concern a certain process $\mathscr{X}$ that satisfies the following identity on the space $\mathcal{C}$:

$$\mathscr{X}(\,\cdot\,) \stackrel{d}{=} \mathbf{1}_{[0,U_1)}(\,\cdot\,)\Delta_1^{\beta}\mathscr{X}^{(1)}\left(\frac{\cdot}{\Delta_1}\right) + \mathbf{1}_{[U_2,1]}(\,\cdot\,)\Delta_1^{\beta}\mathscr{X}^{(1)}\left(\frac{\cdot - \Delta_2}{\Delta_1}\right)$$
$$+ \mathbf{1}_{[U_1,U_2)}(\,\cdot\,)\left[\Delta_2^{\beta}\mathscr{X}^{(2)}\left(\frac{\cdot - U_1}{\Delta_2}\right) + \Delta_1^{\beta}\mathscr{X}^{(1)}(\xi)\right], \tag{12}$$

where $\beta = (\sqrt{17} - 3)/2$, (Δ_1, Δ_2) is Dirichlet$(2, 1)$ distributed, $U_1 = \xi\Delta_1, U_2 = U_1 + \Delta_2$, ξ is uniform on $[0, 1]$, and $\mathscr{X}^{(1)}$, $\mathscr{X}^{(2)}$ are distributed like $\mathscr{X}$, and $(\Delta_1, \Delta_2), \xi, \mathscr{X}^{(1)}, \mathscr{X}^{(2)}$ are independent.

3) Similarly to **2)** above, there is a functional version to Example **3)** of Section 2.3. The process $\mathscr{H}$ involved satisfies the following modified fixed-point equation in the space $\mathcal{C}$:

$$\mathscr{H}(\,\cdot\,) \stackrel{d}{=} \mathbf{1}_{[0,U_1)}(\,\cdot\,)W^{1/3}\mathscr{H}^{(1)}\left(\frac{\cdot}{\Delta_1}\right) + \mathbf{1}_{[U_2,1]}(\,\cdot\,)W^{1/3}\mathscr{H}^{(1)}\left(\frac{\cdot - \Delta_2}{\Delta_1}\right)$$
$$+ \mathbf{1}_{[U_1,U_2)}(\,\cdot\,)\left[(1 - W)^{1/3}\mathscr{H}^{(2)}\left(\frac{\cdot - U_1}{\Delta_2}\right) + W^{1/3}\mathscr{H}^{(1)}(\xi)\right]. \tag{13}$$

Here, (Δ_1, Δ_2) is Dirichlet$(2, 1)$ distributed, $U_1 = \xi\Delta_1, U_2 = U_1 + \Delta_2, \xi, W$ are uniformly distributed, $\mathscr{H}^{(1)}, \mathscr{H}^{(2)}$ are distributed like $\mathscr{H}$. Furthermore, $(\Delta_1, \Delta_2), \xi, W, \mathscr{H}^{(1)}, \mathscr{H}^{(2)}$ in (13).

We note that, in relation to Question (ii) in the introduction, the recursive equations in (11), (12) and (13) do not define the random processes $\mathbf{e}$, $\mathscr{X}$ or $\mathscr{H}$ uniquely (there is at least a free multiplicative scaling). It turns out that the multiplicative scaling is the only degree of freedom when considering solutions in the space $\mathcal{C}_{\text{ex}}$, see Theorem 1 below.

2.5 Fractal properties of metric spaces

Let (S, d) be a metric space. For $\delta > 0$ and a bounded and non-empty set B let $N_B(\delta)$ be the smallest number m such that there exist m balls or radius δ covering B. The Minkowski dimension (also box-counting dimension) emerges from a power-law behavior of $N_B(\delta)$ as $\delta \to 0$. More precisely, we define the *lower Minkowski dimension* $\underline{\dim}_{\mathrm{M}}(B)$ and the *upper Minkowski dimension* $\overline{\dim}_{\mathrm{M}}(B)$ as follows,

$$\underline{\dim}_{\mathrm{M}}(B) := \liminf_{\delta\to 0} \frac{\log N_B(\delta)}{-\log \delta}, \quad \overline{\dim}_{\mathrm{M}}(B) := \limsup_{\delta\to 0} \frac{\log N_B(\delta)}{-\log \delta}.$$

If both values coincide, we simply call $\dim_{\mathrm{M}}(B) := \underline{\dim}_{\mathrm{M}}(B)$ the *Minkowski dimension* of B.

The *Hausdorff dimension* of an arbitrary set $A \subseteq S$ is defined using the family of (outer) Hausdorff measures $(H^s)_{s>0}$ given by

$$H^s(A) := \lim_{\delta\to 0}\left\{\sum_{i\geq 1} |U_i|^s : A \subseteq \bigcup_{i\geq 1} U_i \text{ and } |U_i| \leq \delta \text{ for all } i \geq 1\right\}.$$

The Hausdorff dimension of A is now defined by

$$\dim_{\mathrm{H}}(A) := \inf\{s \geq 0 : H^s(A) = 0\},$$

where one should notice that $H^t(A) < \infty$ implies $H^s(A) = 0$ for $s > t$. We will need the following version of the mass distribution principle (see, e.g. [33, Proposition 4.9]). Let $B_r(x) := \{y \in S : d(x,y) < r\}$ denote the ball of radius r around x. Then, for a Borel-measurable set $A \subseteq S$, a finite measure ν on S with $\nu(A) > 0$ and $c > 0$, we have

$$\limsup_{r \to 0} \nu(B_r(x))/r^s \leq c \text{ for all } x \in A \quad \Rightarrow \quad \dim_{\mathrm{H}}(A) \geq s. \tag{14}$$

Lower and upper Minkowski dimension as well as Hausdorff dimension are invariant under isometries. Furthermore, for a bounded and non-empty set B, we have

$$\dim_{\mathrm{H}}(B) \leq \underline{\dim}_{\mathrm{M}}(B) \leq \overline{\dim}_{\mathrm{M}}(B).$$

Recall that a function $f : [0,1] \to \mathbb{R}$ is α-Hölder continuous with $0 < \alpha \leq 1$ (or Hölder with exponent α) if there exists a finite constant $K > 0$ such that, for all $0 \leq x, y \leq 1$,

$$|f(x) - f(y)| \leq K|x - y|^\alpha.$$

For $f \in \mathcal{C}_{\mathrm{ex}}$, we let α_f be the supremum over all Hölder exponents of excursions equivalent to f. Next, for $p > 0$, the p-variation $[f]_p(t), t \in [0,1]$, of a function $f \in \mathcal{C}$ is defined by

$$[f]_p(t) = \sup \left\{ \sum_{i=0}^{n-1} |f(t_{i+1}) - f(t_i)|^p : n \geq 1 \text{ and } 0 \leq t_1 \leq \ldots \leq t_n \leq t \right\}.$$

Note that $[f]_p$ is monotonically increasing and, if $[f]_p(1) < \infty$, then $[f]_p \in \mathcal{C}$. Let ω_f be the infimum over all $p > 0$ for which $[f]_p(1) < \infty$. It is easy to see that $\omega_f \leq \alpha_f^{-1}$. Furthermore, if f is no-where constant, the converse inequality follows from Theorem 3.1 in [20]. Combining this with Theorem 3.1 in Picard [48] shows that, for any $f \in \mathcal{C}_{\mathrm{ex}}$ which is no-where piecewise constant, we have

$$\overline{\dim}_{\mathrm{M}}(\mathcal{T}_f) = \frac{1}{\alpha_f}. \tag{15}$$

The last identity is used in Corollary 6. (Theorem 1 below shows that *every* non-trivial excursion X satisfying (10) is no-where constant.)

Remark. For a function $f \in \mathcal{C}$ (or rather its equivalence class defined as for excursions), one defines the lower and upper *divider dimensions* $\underline{\dim}_{\mathrm{D}}(f)$ and $\overline{\dim}_{\mathrm{D}}(f)$ analogously to the Minkowski dimension by replacing $N_B(\delta)$ by

$$M_\delta(f) = \min \left\{ M \geq 1 : \begin{array}{l} \exists 0 = t_0 \leq t_1 \leq \ldots \leq t_{M-1} = 1 \text{ such that} \\ |f(t_{i+1}) - f(t_i)| \leq \delta \text{ for all } 0 \leq i \leq M-1 \end{array} \right\}.$$

By Theorem 2.3 in [2], we have $\alpha_f^{-1} = \overline{\dim}_{\mathrm{D}}(f)$ if f is no-where constant. It is easy to see that $\overline{\dim}_{\mathrm{D}}(f) = \overline{\dim}_{\mathrm{M}}(\mathcal{T}_f)$ for $f \in \mathcal{C}_{\mathrm{ex}}$ and the same identity for the lower dimension. This yields an alternative proof of (15) which does not rely on the variations of paths and the results in [48] and [20].

3 Main results

3.1 Characterizations of solutions to (4) and (10)

Our first theorem clarifies the set of solutions to the fixed-point equations (4) and (10).

Theorem 1. *Let $0 < c < \infty$. Then,*

i) there exists a unique continuous excursion X (in distribution) that satisfies (10) and $\mathbf{E}[X(\xi)] = c$;

ii) for any random variable $\mathfrak{X} = (X, d, \mu, \rho)$ in $\mathbb{K}^{\text{GHP}}$ satisfying (4) with $\mathbf{E}[d(\rho, \zeta)] = c$, $\mathfrak{T}_X$ and $(supp(\mu), d, \mu, \rho)$ have the same distribution.

Furthermore, for all $m \geq 1$, we have $\mathbf{E}[\|X\|^m] < \infty$, and, almost surely,

iii) $X(s) > 0$ for all $s \in (0, 1)$;

iv) X is no-where monotonic.

Some comments are in order. First of all, as motivated by the formulation of point *ii)*, there can exist further solutions to (4) which are not almost surely continuum trees (even with values outside $\mathbb{T}^{\text{GHP}}$), see Proposition 3 below and the example discussed following Theorem 1.6 in [3]. Next, random compact rooted measured metric spaces (not only real trees) with values in $\mathbb{K}^{\text{GHP}}$ or $\mathbb{K}^{\text{GP}}$ solving (4) with $\mathbf{E}[d(\rho, \zeta)] = \infty$ do not exist. Similarly, any solution to (4) with $\mathbf{E}[d(\rho, \zeta)] = 0$ must be the trivial space $\mathfrak{X} = \{\rho\}$ almost surely. The analogous statements hold for solutions to (10).

Finally, characterizing the distribution of a random continuous function by a stochastic fixed-point equation typically requires moment assumptions, see, e.g. [45, Lemma 18]. The uniqueness property in the theorem follows from the restriction to non-negative excursions and the very special form of the fixed-point equation (10), where, for any fixed $t \in (0, 1)$, almost surely, the variations around t are governed by a *single* copy of the process. Let $\mathcal{D}$ be the set of functions $f : [0, 1] \to \mathbb{R}$ with right-continuous paths such that $\lim_{t \uparrow s} f(t)$ exists for all $s \in (0, 1]$. On $\mathcal{D}$, we consider the Skorokhod topology, see e.g. [15] for details. By $\mathcal{D}_{\text{ex}}$ we denote the set of non-negative functions $f \in \mathcal{D}$ with $f(0) = f(1) = 0$. Our proof shows that, up to a multiplicative constant, the excursion X is the unique solution to (10) (in distribution) in the space $\mathcal{D}_{\text{ex}}$. Furthermore, X is the unique solution to (10) (in distribution) in the space $\mathcal{D}$ (and hence, also in $\mathcal{C}$) satisfying $\mathbf{E}[\|X\|^{\mathfrak{m}}] < \infty$, where $\mathfrak{m} = 1 + \lfloor 1/\alpha \rfloor$; in particular, $\mathfrak{m} \geq 1/\alpha$. Note that,

$$\mathfrak{m} = \min \left\{ m \in \mathbb{N} : \sum_{i=1}^{K} \mathbf{E}[\mathcal{R}_i^{m\alpha}] < 1 \right\}. \tag{16}$$

Albenque and Goldschmidt [3] show that, in the case of the Brownian continuum random tree, the fixed-point in Example **1)** of Section 2.3 is attractive with respect to weak convergence on the space of probability measures on $\mathbb{K}^{\text{GP}}$. They also raise the question whether this was true with respect to the Gromov–Hausdorff–Prokhorov distance. Our next result confirms this under certain moment conditions. (Note however that the results in [3] are not directly comparable to ours since the trees there are *unrooted*.)

In the following, we let $\phi_{\text{GHP}} : \mathcal{M}_1(\mathbb{K}^{\text{GHP}}) \to \mathcal{M}_1(\mathbb{K}^{\text{GHP}})$ be the map that to $\aleph \in \mathcal{M}_1(\mathbb{K}^{\text{GHP}})$ associates $\phi_{\text{GHP}}(\aleph) = \Psi(\aleph, \tau)$, where $\tau = \mathfrak{L}(\mathcal{R}, \mathcal{S})$ and Ψ is the map defined in (3). Analogously, we define ϕ_{GP} based on $\mathbb{K}^{\text{GP}}$. For a measure ν on $\mathbb{K}^{\text{GP}}$ or $\mathbb{K}^{\text{GHP}}$, we write $\mathbf{E}_\nu$ for the expectation with respect to spaces sampled from ν.

Theorem 2. *Fix $c > 0$ and let X be the unique continuous excursion satisfying (10) with $\mathbf{E}[X(\xi)] = c$. Then, we have the following two statements:*

i) if ν is a probability measure on $\mathbb{K}^{\text{GP}}$ with $\mathbf{E}_\nu[d(\rho, \zeta)] = c$, then the sequence of distributions $(\phi_{\text{GP}}^n(\nu))_{n \geq 1}$ converges weakly to the law of $\mathfrak{T}_X$.

ii) if ν is a probability measure on $\mathbb{T}_{\text{f}}^{\text{GHP}}$ with $\mathbf{E}_\nu[d(\rho, \zeta)] = c$ and $\mathbf{E}_\nu[\|\mathcal{T}\|^{\mathfrak{m}}] < \infty$, then the sequence of distributions $(\phi_{\text{GHP}}^n(\nu))_{n \geq 1}$ converges weakly to the law of $\mathfrak{T}_X$.

Although the moment condition in Theorem 2 *ii)* is probably not optimal, one certainly needs some condition as demonstrated by the following proposition.

Proposition 3. *Let $c > 0$. Fix $K \geq 2$, Γ and (the distribution of) $\mathcal{S}$. Then, there exists an $\mathcal{R}$ such that*

i) there exists infinitely many mutually singular fixed-points of (4) on $\mathbb{K}^{\text{GHP}}$ such that $\mathbf{E}_\nu[d(\rho, \zeta)] = c$ and $\mathbf{E}_\nu[\|\mathcal{T}\|^{1/\alpha}] = \infty$, including some that are a.s. not real trees;

ii) there exists a probability measure ν on $\mathbb{T}_{\text{f}}^{\text{GHP}}$ such that $\mathbf{E}_\nu[d(\rho, \zeta)] = c$, $\mathbf{E}_\nu[\|\mathcal{T}\|^{1/\alpha}] = \infty$ and $(\phi_{\text{GHP}}^n(\nu))_{n \geq 1}$ does not converge weakly to the law of $\mathfrak{T}_X$.

3.2 Geometry, fractal dimensions and optimal Hölder exponents

It is informative to first present a heuristic argument that, at the very least, gives an idea of the value of the Minkowski dimension that one should expect. Consider $\mathfrak{T} = (\mathcal{T}, d, \mu, \rho)$ satisfying (4). In a covering of $\mathcal{T}$ by open balls, if one neglects the contribution of the balls that might intersect more than one subtree in the recursive decomposition, then the fixed-point equation (4) should imply that we approximately have

$$N_{\mathcal{T}}(\delta) \approx \sum_{i=1}^{K} N_{\mathcal{R}_i^\alpha \mathcal{T}_i}(\delta) = \sum_{i=1}^{K} N_{\mathcal{T}_i}(\mathcal{R}_i^{-\alpha}\delta). \tag{17}$$

In particular, if one roughly has $N_{\mathcal{T}}(\delta) \approx \delta^{-s}$ for some $s > 0$, then it should be the case that $1 = \sum_{i=1}^{K} \mathcal{R}_i^{\alpha s}$, and thus the constant s should be given by $s = \alpha^{-1}$. We now provide results that justify that this is indeed the case under some conditions that happen to be satisfied in most examples.

In the following theorem and subsequently, we use the generic random variable $\overline{\mathcal{R}}$ which is distributed like $\mathcal{R}_I$ with I uniformly chosen among $1, \dots, K$.

Theorem 4 (Upper bound on $\overline{\dim}_{\mathrm{M}}$)**.** *Assume that $\mathfrak{T} = (\mathcal{T}, d, \mu, \rho)$ satisfies (4) with values in $\mathbb{T}_{\mathrm{f}}^{\mathrm{GHP}}$. Moreover, assume that $\overline{\mathcal{R}}$ admits a density g on $[0,1]$ satisfying $g(t) \le Ct^{\gamma-1}$ for all $t \in [0,1]$ where $KC > \gamma > 0$. Then, almost surely,*

$$\overline{\dim}_{\mathrm{M}}(\mathcal{T}) \le (KC - \gamma)/\alpha.$$

Remark. Note that, if $\overline{\mathcal{R}}$ has density $(K-1)^{-1}t^{(2-K)/(K-1)}$, then $\overline{\dim}_{\mathrm{M}}(\mathcal{T}) \le \alpha^{-1}$ almost surely. Although it may seem arbitrary, the argument around (17) motivates this choice, and we will see it will turn out to be an important case later on. Finally, our proof shows that the bound in the theorem remains valid for any random variable $\mathfrak{T} \in \mathbb{T}^{\mathrm{GHP}}$ satisfying (4) if $\|\mathcal{T}\|$ has finite moments of all orders.

We now turn to the Hausdorff dimension.

Theorem 5 (Lower bound on $\dim_{\mathrm{H}}$)**.** *Assume $\mathfrak{T} = (\mathcal{T}, d, \mu, \rho)$ satisfies (4) with values in $\mathbb{T}_{\mathrm{f}}^{\mathrm{GHP}}$, $\mathbf{P}(\|\mathcal{T}\| > 0) > 0$ and, for some $\delta > 0$, $\mathbf{E}[\overline{\mathcal{R}}^{-\delta}] < \infty$. Then*

$$\dim_{\mathrm{H}}(\mathcal{T}) \ge \alpha^{-1}.$$

Corollary 6 (Optimal Hölder exponents)**.** *Suppose that X satisfies (10), that $\mathbf{P}\left(\|X\| > 0\right) > 0$ and that the conditions of Theorems 4 are satisfied with $C = \gamma$ and $\gamma = (K-1)^{-1}$. Then, almost surely:*

(a) $\dim_{\mathrm{H}}(\mathcal{T}_X) = \underline{\dim}_{\mathrm{M}}(\mathcal{T}_X) = \overline{\dim}_{\mathrm{M}}(\mathcal{T}_X) = \alpha^{-1}$;

(b) for any $\gamma < \alpha$, there exists a process $\tilde{X}$ which is equivalent to X and has γ-Hölder continuous paths.

(c) for any $\gamma > \alpha$ and $\tilde{X} \in \mathcal{C}_{ex}$, on the event $\{\mathcal{T}_{\tilde{X}} = \mathcal{T}_X\}$, $\tilde{X}$ has γ-Hölder continuous paths with probability zero.

Finally, we have the following results about the degrees in $\mathfrak{T}_X$. Let $\mathscr{D}(\mathcal{T}_X)$ be the (random) set of degrees of $\mathcal{T}_X$. Let also $\mathscr{D}(\Gamma) = \{1 + \#\{j : \varpi_j = i\} : 1 \le i \le K\}$; then, by the following proposition, with probability one, $\mathscr{D}(\Gamma)$ is the set of degrees of the connect points $\varphi(\eta_i)$, $1 \le i \le K$ in $\mathcal{T}$. Observe that $1 \in \mathscr{D}(\Gamma)$, but that it is possible that 2 is not an element of $\mathscr{D}(\Gamma)$. ($\mathscr{D}(\Gamma)$ is also the set of degrees in the tree obtained from Γ by connecting an additional node to its root.)

Proposition 7. *Let $0 < c < \infty$ and X be the unique solution (in distribution) of (4) in Theorem 1 with $\mathbf{E}\left[X(\xi)\right] = c$. Then, with probability one,*

i) the root ρ_X is a leaf;

ii) a point sampled from μ_X is a leaf;

iii) μ_X has no atoms;

iv) the set $\mathscr{D}(\mathcal{T}_X)$ of degrees of points in $\mathcal{T}_X$ is fully determined by $\mathscr{D}(\Gamma)$:

$$\mathscr{D}(\mathcal{T}_X) = \begin{cases} \{1,2,3\} & \text{if } \mathscr{D}(\Gamma) = \{1,2\} \\ \mathscr{D}(\Gamma) \cup \{2\} & \text{otherwise}. \end{cases}$$

3.3 A taste of applications

All applications in this work are discussed in detail in Section 6; they cover in particular generalizations of the trees dual to laminations of the disk. Here, we only state immediate consequences for the three trees encoded by the functions $\mathscr{Z}$, $\mathscr{H}$ and $\mathbf{e}$ that we have used as examples earlier. Observe that, the structural tree Γ is fixed if $K = 2$, while for $K \geq 3$, in order to describe it, it suffices to specify the parents $\varpi_3, \dots, \varpi_K$.

Corollary 8. *Up to a multiplicative constant for the distance function, we have (uniqueness being understood in distribution):*

(a) The Brownian Continuum random tree $(\mathcal{T}_\mathbf{e}, d_\mathbf{e}, \mu_\mathbf{e}, \pi_\mathbf{e}(0))$ is the unique random rooted continuum tree satisfying (4) with $K = 3$, $\vartheta_3 = 1$, $\alpha = 1/2$, $\mathcal{S} = \mathcal{R} = (\Delta_1, \Delta_2, \Delta_3)$ with $\Delta_1, \Delta_2, \Delta_3$ as in (11).

(b) The tree $(\mathcal{T}_\mathscr{Z}, d_\mathscr{Z}, \mu_\mathscr{Z}, \pi_\mathscr{Z}(0))$ is the unique random rooted continuum tree satisfying (4) with $K = 2$, $\mathcal{S} = \mathcal{R} = (\Delta_1, \Delta_2)$, with Δ_1, Δ_2 as in (12) and $\alpha = (\sqrt{17} - 3)/2$.

(c) The tree $(\mathcal{T}_\mathscr{H}, d_\mathscr{H}, \mu_\mathscr{H}, \pi_\mathscr{H}(0))$ is the unique random rooted continuum tree satisfying (4) with $K = 2$, $\alpha = 1/3$, $\mathcal{S} = (\Delta_1, \Delta_2)$, $\mathcal{R} = (W, (1 - W))$ with Δ_1, Δ_2, W as in (13).

Furthermore, these uniqueness results also hold in the space $\mathbb{K}_\mathrm{f}^\mathrm{GHP}$ of full-support metric measure spaces.

The assertion concerning the Brownian continuum random tree (in the unrooted set-up and using a slightly different definition of continuum trees) in Corollary 8 is the main result in [3]. Similarly, the claim for the process $\mathscr{Z}$ in Corollary 9 has already been given in [17] in the space $\mathcal{C}$ under additional moment assumptions. In terms of processes, we have the following

Corollary 9. *Up to a multiplicative constant, we have (uniqueness being understood in the sense of distributions):*

(a) The Brownian excursion $\mathbf{e}$ is the unique continuous excursion satisfying (11).

(b) The process $\mathscr{Z}$ is the unique continuous excursion satisfying (12).

(c) The process $\mathscr{H}$ is the unique continuous excursion satisfying (13).

We now formulate the implications of results on the fractal dimensions of $\mathcal{T}_\mathscr{Z}$, $\mathcal{T}_\mathscr{H}$ and $\mathcal{T}_\mathbf{e}$. Note that the fractal dimension of the continuum random tree has already been established in [5] (see also [29] for the more general case of Lévy trees) and the Minkowski dimension of $\mathcal{T}_\mathscr{Z}$ in [17].

Corollary 10. *Almost surely, for the processes $\mathbf{e}$, $\mathscr{Z}$ and $\mathscr{H}$ satisfying (11), (12) and (13),*

(a) $\dim_\mathrm{M}(\mathcal{T}_\mathbf{e}) = \dim_\mathrm{H}(\mathcal{T}_\mathbf{e}) = 2$.

(b) $\dim_\mathrm{M}(\mathcal{T}_\mathscr{Z}) = \dim_\mathrm{H}(\mathcal{T}_\mathscr{Z}) = \beta^{-1}$, *where* $\beta = (\sqrt{17} - 3)/2$,

(c) $\dim_\mathrm{M}(\mathcal{T}_\mathscr{H}) = \dim_\mathrm{H}(\mathcal{T}_\mathscr{H}) = 3$.

3.4 Overview of the main techniques

Most of our proofs rely on an expansion of the fixed-point equation (10). For fixed-point equations describing real-valued distributions, this idea is classical, compare, e.g. [9, Section 2.3]. For the Brownian continuum random tree $\mathcal{T}_\mathbf{e}$, it was used both in [22] as well as, more recently, in [3].

Let $\Theta = \bigcup_{n \geq 0} [K]^n$ be the complete infinite K-ary tree and, for each $n \geq 0$, denote by $\Theta_n = [K]^n \subseteq \Theta$ the set of vertices on level n in Θ. Next, let

$$\{(\mathcal{R}^\vartheta, \mathcal{S}^\vartheta) : \vartheta \in \Theta\}, \qquad \{\Xi^\vartheta = (\xi_1^\vartheta, \dots, \xi_K^\vartheta) : \vartheta \in \Theta\},$$

be two independent sets of independent random variables where each $(\mathcal{R}^\vartheta, \mathcal{S}^\vartheta)$ has the distribution of $(\mathcal{R}, \mathcal{S})$ and each Ξ^ϑ is uniformly distributed on $(0,1)^K$. The components of $\mathcal{R}^\vartheta$ and $\mathcal{S}^\vartheta$ are assigned to the edges out of ϑ as follows: for the edge e_i^ϑ between ϑ and ϑi, define $\mathcal{R}(e_i^\vartheta) = \mathcal{R}_i^\vartheta$ and $\mathcal{S}(e_i^\vartheta) = \mathcal{S}_i^\vartheta$.

These edge-weights then induce values for the vertices that we define multiplicatively; each node $\vartheta \in \Theta$ is assigned a *length* $\mathcal{L}(\vartheta)$ and a *rescaling factor for distances* $\mathcal{V}(\vartheta)$ which are given by

$$\mathcal{V}(\vartheta) = \prod_{e \in \pi_\vartheta} \mathcal{R}(e) \qquad \text{and} \qquad \mathcal{L}(\vartheta) = \prod_{e \in \pi_\vartheta} \mathcal{S}(e), \tag{18}$$

where π_ϑ denotes the set of edges on the path from the root $\emptyset$ to ϑ. We can think of Θ as providing the parameters that are required by the recursive decomposition. In Section 4.2, we will see that, for any $0 < c < \infty$, one can construct a family of random variables $\{\mathcal{X}^\vartheta : \vartheta \in \Theta\}$, such that

$$\mathcal{X}^\vartheta = \Phi\big(\mathcal{X}^{\vartheta 1}, \ldots, \mathcal{X}^{\vartheta K}, \mathcal{R}^\vartheta, \mathcal{S}^\vartheta, \Xi^\vartheta\big), \tag{19}$$

where, for all $\vartheta \in \Theta$, the distribution of $\mathcal{X}^\vartheta$ does not depend on ϑ, $\mathcal{X}^\vartheta$ is measurable with respect to $\{\mathcal{R}^{\vartheta\sigma}, \mathcal{S}^{\vartheta\sigma}, \Xi^{\vartheta\sigma} : \sigma \in \Theta\}$, and $\mathbf{E}[\mathcal{X}^\vartheta(\xi)] = c$. In particular, the fixed-point corresponding to $\mathcal{X}^\vartheta$ is *endogenous*, see, e.g., [9, Definition 7].

It should be clear that, for any $\vartheta \in \Theta$ and $n \geq 1$, the unit interval is decomposed in half-open intervals such that, for $\sigma \in \Theta_n$, $\mathcal{X}^{\vartheta\sigma}$ multiplied by $(\mathcal{V}(\vartheta\sigma)/\mathcal{V}(\vartheta))^\alpha$ governs the behaviour of the process $\mathcal{X}^\vartheta$ on a set Λ^ϑ_σ of Lebesgue measure $\mathcal{L}(\vartheta\sigma)/\mathcal{L}(\vartheta)$ composed of a subset of these intervals. Let us give a precise formulation of this decomposition: first, for all $\vartheta \in \Theta$ and $1 \leq j \leq K$, let the set Λ^ϑ_j and the function φ^ϑ_j be defined as S_j and φ_j in (8) using the vector $(\mathcal{S}^\vartheta, \Xi^\vartheta)$. Then, given Λ^ϑ_σ and φ^ϑ_σ for $\sigma \in \Theta_n, n \geq 1$, for $1 \leq j \leq K$, let

$$\Lambda^\vartheta_{\sigma j} = \big(\varphi^\vartheta_\sigma\big)^{-1}(\Lambda^\sigma_j),$$

and let $\varphi^\vartheta_{\sigma j}$ be the unique piece-wise linear bijective and increasing function mapping $\Lambda^\vartheta_{\sigma j}$ to $[0,1]$. Throughout the paper, we write $\Lambda_\vartheta = \Lambda^\emptyset_\vartheta, \vartheta \in \Theta$ and $\mathcal{X} = \mathcal{X}^\emptyset$ for the quantities at the root of Θ.

Remark. Our construction also leads to a corresponding decomposition of the tree $\mathfrak{T}_{\mathcal{X}^\vartheta}$ into subtrees GHP-isometric to $\mathfrak{T}_{\mathcal{X}^{\vartheta\sigma}}, \sigma \in \Theta_n$ for any $n \geq 1$. Note that a similar direct construction of almost surely coupled isometry classes of random compact rooted measured metric spaces (or real trees) using the map Ψ is not available since Ψ operates only on the set of probability measures.

3.5 Outline of the proofs

The remainder of the paper is organized as follows: In Section 4, we prove Theorems 1 *i), ii)*, 2 and Proposition 3. Here, Section 4.1 is devoted to showing that, up to a scaling constant, there exists at most one solution to (4) in $\mathbb{K}^{\mathrm{GHP}}_{\mathrm{f}}$ (or $\mathbb{K}^{\mathrm{GP}}$). In Section 4.2 this solution is constructed together with a unique continuous excursion satisfying (10). Parts *i)* and *ii)* of Theorem 1 are proved at the end of this section. Section 4.3 contains the proofs of Theorem 2 and Proposition 3.

In Section 5 we present proofs to the remaining statements in Section 3. Here, in Section 5.1, we start with the verification of Theorem 4. Then, in Section 5.2, we give the proof of the lower bound on the Hausdorff dimension in Theorem 5. In Section 5.3, we discuss the proofs of Proposition 7 as well as the remaining parts *iii)* and *iv)* of Theorem 1. Corollary 6 is discussed in Section 5.4 where we also study promising candidates for encoding functions with good path properties.

Finally, Section 6 is dedicated to applications. We discuss the results formulated in Section 3.3 for the processes $\mathbf{e}$, $\mathscr{Z}$ and $\mathscr{H}$ in detail in Sections 6.1–6.3. Section 6.4 contains new results concerning a generalization of the lamination model [17, 23].

4 Proofs of the uniqueness results

4.1 Uniqueness of the encoding function

The proof of Theorem 1 consists of two steps. First, we show that there exists at most one compact rooted measured metric space with full support satisfying (4) in Proposition 12. Second, as in [17], we use a variant of the functional contraction method to construct a solution to (10) that has finite absolute moments. We indicate how to obtain the statements *i)* and *ii)* in Theorem 1 from the next two results right after proving Proposition 13 below.

For $k \in \mathbb{N}$, let $\mathbb{M}_k = \{(m_{ij})_{0 \le i,j \le k} : m_{ij} = m_{ji} \ge 0\}$ be the set of symmetric $(k+1)$ by $(k+1)$ matrices with non-negative entries. We also write $\mathbb{M}_{\mathbb{N}_0}$ for the set of infinite dimensional matrices satisfying these properties. For a rooted measured compact space $\mathfrak{X} = (X, d, \mu, \rho)$, let $\zeta_1, \zeta_2, \dots$ be independent realizations of the measure μ on X. Furthermore, set $\zeta_0 = \rho$. Observe that the distribution of the random infinite matrix $\mathfrak{D}_{\mathfrak{X}} = (d(\zeta_i, \zeta_j))_{i,j \ge 0}$ does not depend on the representative of the Gromov–Hausdorff–Prokhorov (or Gromov–Prokhorov) isometry class of $\mathfrak{X}$. Hence, we can define the *distance matrix* $\nu^{\mathfrak{X}} \in \mathcal{M}_1(\mathbb{M}_{\mathbb{N}_0})$ for elements $\mathfrak{X}$ of $\mathbb{K}^{\mathrm{GHP}}$ or $\mathbb{K}^{\mathrm{GP}}$. The importance of $\nu^{\mathfrak{X}}$ is highlighted by the following well-known proposition. For GP-isometry classes, part *i)* is typically referred to as Gromov's reconstruction theorem, see Gromov [35, Section 3 1/2] or Vershik [54, Theorem 4]. *ii)* and *iii)* (again in the GP case) are Theorem 5 in [26] (see also [41, Corollary 2.8]). For GHP-isometry classes, these results immediately follow from the bimeasurability of ι discussed at the end of Section 2.1.

Proposition 11. *We have the following results:*

i) Let $\mathfrak{X}_1, \mathfrak{X}_2 \in \mathbb{K}_{\mathrm{f}}^{\mathrm{GHP}}$ (or $\mathbb{K}^{\mathrm{GP}}$). Then, $\nu^{\mathfrak{X}_1} = \nu^{\mathfrak{X}_2}$ if and only if $\mathfrak{X}_1 = \mathfrak{X}_2$.

ii) The distribution of a random variable $\mathfrak{X} \in \mathbb{K}_{\mathrm{f}}^{\mathrm{GHP}}$ (or $\mathbb{K}^{\mathrm{GP}}$) is uniquely determined by $\mathbf{E}[\nu^{\mathfrak{X}}]$, the intensity measure of (the random measure) $\nu^{\mathfrak{X}}$:

$$\mathbf{E}[\nu^{\mathfrak{X}}](A) := \mathbf{E}[\nu^{\mathfrak{X}}(A)], \quad A \subseteq \mathbb{M}_{\mathbb{N}_0} \quad \textit{measurable}.$$

iii) For random variables $\mathfrak{X}, \mathfrak{X}_n, n \ge 1$ with values in $\mathbb{K}^{\mathrm{GP}}$, we have $\mathfrak{X}_n \to \mathfrak{X}$ in distribution as $n \to \infty$ if and only if $\mathbf{E}[\nu^{\mathfrak{X}_n}] \to \mathbf{E}[\nu^{\mathfrak{X}}]$ weakly.

Remark. By [34, Corollary 3.1], a sequence of random variables $\mathfrak{X}_n, n \ge 1$ in $\mathbb{K}^{\mathrm{GP}}$ converges in distribution if and only if $\mathbf{E}[\nu^{\mathfrak{X}_n}]$ converges weakly and the family of distributions $\mathfrak{L}(\mathfrak{X}_n), n \ge 1$ is tight. As discussed in Remark 2.7 in [26] (and repeated here for the reader's convenience), Proposition 11 *iii)* says that the tightness condition is equivalent to the fact that the weak limit of $\mathbf{E}[\nu^{\mathfrak{X}_n}], n \ge 1$ can be obtained as annealed distance matrix of some random variable $\mathfrak{X} \in \mathbb{K}^{\mathrm{GP}}$.

With Proposition 11 at hand, we can now compare the distributions of the distance matrix of two solutions of the fixed-point equation (4). Here, and subsequently, for $\mathfrak{X} \in \mathbb{K}^{\mathrm{GHP}}$ (or $\mathbb{K}^{\mathrm{GP}}$), we write $\nu_n^{\mathfrak{X}}$ for the distribution of the distance matrix with respect to finitely many points $\zeta_0, \zeta_1, \dots, \zeta_n$. ($\nu_n^{\mathfrak{X}}$ is an element of $\mathcal{M}_1(\mathbb{M}_n)$.)

Proposition 12. *Let $\mathfrak{X} = (X, d, \mu, \rho)$ and $\mathfrak{U} = (U, d', \mu', \rho')$ be $\mathbb{K}^{\mathrm{GHP}}$ (or $\mathbb{K}^{\mathrm{GP}}$)-valued random variables satisfying (4) (in distribution) and $\mathbf{E}[d(\rho, \zeta)] = \mathbf{E}[d'(\rho', \zeta')]$ where ζ (resp. ζ') is chosen on X (resp. U) according to μ (resp. μ'). Then, $\mathbf{E}[\nu^{\mathfrak{X}}] = \mathbf{E}[\nu^{\mathfrak{U}}]$.*

Proof. For $n \ge 1$, let D_n (resp. D'_n) be a random variable on $\mathbb{M}_n$ with distribution $\mathbf{E}[\nu_n^{\mathfrak{X}}]$ (resp. $\mathbf{E}[\nu_n^{\mathfrak{U}}]$). We have already seen that D_1 and D'_1 are identically distributed in the course of (6). Our aim is to show that, for all $n \ge 2$, D_n satisfies a stochastic fixed-point which is known to admit at most one solution. To this end, we proceed by induction on n. It then follows immediately that $\mathbf{E}[\nu_n^{\mathfrak{X}}] = \mathbf{E}[\nu_n^{\mathfrak{U}}]$ for all $n \ge 1$ which proves the assertion. Let $A \subseteq \mathbb{M}_n$ be a measurable set. Abbreviating $\tau = \mathfrak{L}(\mathcal{R}, \mathcal{S})$ and $\aleph = \mathfrak{L}(\mathfrak{X})$, we deduce from (4) that

$$\begin{aligned}
\mathbf{P}\left(D_n \in A\right) &= \mathbf{E}[\nu_n^{\psi(\mathfrak{X}_1, \dots, \mathfrak{X}_K, \mathcal{R}, \mathcal{S})}(A)] \\
&= \int_{\Sigma_K^2} d\tau(r, s) \int_{(\mathbb{K}^{\mathrm{GHP}})^K} d\aleph^{\otimes k}(x_1, \dots, x_k) \int_{\mathbb{K}^{\mathrm{GHP}}} \nu_n^{x}(A) d\psi(x_1, \dots, x_k, r, s)(x) \\
&= \int_{\Sigma_K^2} d\tau(r, s) \int_{(\mathbb{K}^{\mathrm{GHP}})^K} d\aleph^{\otimes k}(x_1, \dots, x_k) \mathbf{E}[\nu_n^{\mathfrak{X}^*(x_1, \dots, x_k, r, s)}](A) \qquad (20)
\end{aligned}$$

where $\mathfrak{X}^*(x_1, \ldots, x_k, r, s)$ has distribution $\psi(x_1, \ldots, x_k, r, s)$. Fix $r, s, x_1, \ldots, x_K$ and let $\bar{x}_1, \ldots, \bar{x}_K$ be arbitrary representatives of $x_1, \ldots, x_K$. For $k \geq 1, 1 \leq i \leq K$, write $Y_{\bar{x}_i}^{(k)}$ for a generic random variable with distribution $\nu_k^{\bar{x}_i}$ (which, of course, only depends on x_i rather than $\bar{x}_i$).

Furthermore, let $(J_i)_{i \geq 1}$ be a sequence of i.i.d. random variables on $\{1, \ldots, K\}$, with $\mathbf{P}(J_i = j) = s_j$ for all $1 \leq j \leq K$. For $i \geq 1$, sample ζ_i independently according to the measure μ_{J_i} on $\bar{x}_i$. Independently, for $1 \leq i \leq K$, sample η_i on $\bar{x}_i$ according to μ_i. Write d for the distance on the tree $\bar{x}$ constructed with the help of $\Gamma, \alpha, r, s, \bar{x}_1, \ldots, \bar{x}_K, \eta_1, \ldots, \eta_K$. (Note that we do not interpret $\bar{x}$ as a random variable.) By construction, the matrix $W_n = (w(i,j))_{0 \leq i,j \leq n}$ where $w(i,j) = d(\zeta_i, \zeta_j)$ has distribution $\mathbf{E}[\nu_n^{\mathfrak{X}^*(x_1, \ldots, x_k, r, s)}]$.

The idea now consists in decomposing W_n by looking in which of the subspaces the random points $\zeta_1, \ldots, \zeta_n$ fall. Recall that the root $\rho = \zeta_0$ is the (image of the) root ρ_1 of $\bar{x}_1$. We start with a few simple cases as a warm up:
i) If $J_l = 1$, for all $l = 1, \ldots, n$, then the distances between points $\zeta_0 = \rho, \zeta_1, \ldots, \zeta_n$ are all measured within $\bar{x}_1$ and conditional on this event, W_n is distributed as $r_1^\alpha Y_{\bar{x}_1}^{(n)}$.
ii) If there is some $i \neq 1$ such that $\#\{1 \leq l \leq n : J_l = i\} = n$, then the distances between $\zeta_1, \ldots, \zeta_n$ all are measured within $\bar{x}_i$. The distance from ζ_i to ζ_0 involves all subspaces on the path $\pi(1,i)$ between i and 1 (including 1, but excluding i) in the structural tree Γ. Furthermore, $d(\zeta_0, \zeta_l)$ can be decomposed as $d(\rho_i, \zeta_l)$ plus the sum of the distances $d(\rho_p, \eta_p)$, for $p \in \pi(1,i)$. It follows that, conditional on this event and on i, the matrix W_n is distributed as

$$r_i^\alpha Y_{\bar{x}_i}^{(n)} + F_0\left(\sum_{p \in \pi(1,i)} r_p^\alpha Y_{\bar{x}_p}^{(1)} \right),$$

where $Y_{\bar{x}_i}^{(n)}, Y_{\bar{x}_p}^{(1)}$, $p \in \pi(1,i)$, are independent and $F_0 : \mathbb{M}_2 \to \mathbb{M}_{n+1}$ is the linear operator that copies the (only) non-zero entry of the input matrix to all the non-diagonal entries.
iii) The situation is very similar if there exists $i \in \{1, \ldots, n\}$ such that $J_i = j$, while $J_l = 1$, for $l = 0, 1, \ldots, i-1, i+1, \ldots, n$. Then, distances between the ζ_j, $j \neq i$, only concern $\bar{x}_1$ while the distances between ζ_i and elements of $\{\zeta_0, \ldots, \zeta_{i-1}, \zeta_{i+1}, \zeta_n\}$ all involve $\bar{x}_p$, for $p \in \pi(i,1)$. Furthermore, when studying these distances, we might think of η_1 as ζ_j^*, to replace ζ_j in $\bar{x}_1$. Let $F_j : \mathbb{M}_2 \to \mathbb{M}_{n+1}$ denote the operator copying the non-zero entry of the input matrix to all the entries (k,l) for $k \neq j + l \mod (n+1)$. Then, the same argument as in *ii)* yields that, conditional on this event and on i, the matrix W_n is distributed as

$$r_1^\alpha Y_{\bar{x}_1}^{(n)} + F_j\left(\sum_{p \in \pi(i,1)} r_p^\alpha Y_{\bar{x}_p}^{(1)} \right),$$

where $Y_{\bar{x}_1}^{(n)}, Y_{\bar{x}_p}^{(1)}$, $p \in \pi(1,i)$, are independent.
iv) In general, for $1 \leq i \leq K$, let $L_i = \{1 \leq j \leq k : J_j = i\}$ be the collection of test points falling in $\bar{x}_i$, and write $\ell_i = \#L_i$. Let $M = \max\{\ell_i : 1 \leq i \leq K\}$. The cases *i)* and *ii)* above have treated the situations where $M = n$ and $M = n-1$, respectively. So we can now assume that $M \leq n-2$. As we have already observed before, for every i, and any point x_i in $\bar{t}_i$ and $x_j \in \bar{t}_j$, $j \neq i$, the segment between x_i and x_j in $\bar{t}$ exits the (image of) $\bar{t}_i$ at either ρ_i or η_i. In any case, the number of random points in the subspaces to consider (in addition to the root) never exceeds $n-1$ (the η_i plus at most $n-2$ of the original ones). More precisely, let

$$\ell_i^* = \begin{cases} \ell_i + 1 & \text{if } \ell_j > 0 \text{ for some } j \in \Gamma_i \setminus \{i\} \\ \ell_i & \text{otherwise}. \end{cases}$$

Conditional on the sets L_i, $1 \leq i \leq K$, in distribution, W_n can be expressed as a deterministic linear operator involving rescaled independent copies of $Y_{\bar{t}_i}^{(k)}, 1 \leq i \leq K$ for $k \in \{\ell_i^* : 1 \leq i \leq K\}$ whose precise expression has little importance.

Recall that the distribution of $Y^{(k)}_{\bar{x}_i}$ is equal to $\nu^{x_i}_k$. Thus, using the observations in *i), ii), iii)* and *iv)*, and performing the integration in (20), we obtain

$$D_n \stackrel{d}{=} A_n D_n + B_n \tag{21}$$

for a non-negative random variable A_n and a linear operator B_n where $(A_n, B_n), D_n$ are independent, and B_n is composed of $\mathcal{R}, \mathcal{S}$ and independent copies of certain $D_1, \ldots, D_{n-1}$. A fixed-point equation of this kind is called a perpetuity. Note that $|A_n| < 1$ almost surely. It follows from classical results on perpetuities, e.g. from [55, Theorem 1.5], that this fixed-point equation has at most one solution in distribution. This finishes the proof. $\square$

4.2 Construction of a solution

It is here that we construct the process family of processes $(\mathcal{X}^\vartheta)_{\vartheta\in\Theta}$ mentioned in Section 3.4 and that plays a central role in a number of proofs later on. The following proposition is essentially a generalization of Theorem 6 and Theorem 17 in [17]. We keep the presentation rather compact and refer to [17] for more details on technical points. Let $\mathcal{M}_1(\mathcal{C}_{\text{ex}})$ be the set of probability measures on $\mathcal{C}_{\text{ex}}$. For $c > 0$, let

$$\mathcal{M}^c = \left\{ \mu \in \mathcal{M}_1(\mathcal{C}_{\text{ex}}) : \iint_0^1 x(t)dt\mu(dx) = c \text{ and } \int \|x\|^m \mu(dx) < \infty, \text{ for all } m \geq 1 \right\}.$$

Fix $c > 0$, $f \in \mathcal{C}_{\text{ex}}$ with $\int_0^1 f(t)dt = c$ and let $Q^\vartheta_0 = f, \vartheta \in \Theta$. Recursively, for $n \geq 1$ and $\vartheta \in \Theta$, define

$$Q^\vartheta_n = \Phi(Q^{\vartheta 1}_{n-1}, \ldots, Q^{\vartheta K}_{n-1}, \mathcal{R}^\vartheta, \mathcal{S}^\vartheta, \Xi^\vartheta). \tag{22}$$

Proposition 13. *For any $\vartheta \in \Theta$, almost surely, the sequence Q^ϑ_n defined in (22) converges uniformly to a process $\mathcal{X}^\vartheta$. For any $\vartheta \in \Theta$, we have, almost surely,*

$$\mathcal{X}^\vartheta = \Phi(\mathcal{X}^{\vartheta 1}, \ldots, \mathcal{X}^{\vartheta K}, \mathcal{R}^\vartheta, \mathcal{S}^\vartheta, \Xi^\vartheta). \tag{23}$$

Furthermore, $\mathfrak{L}(\mathcal{X}^\vartheta)$ is the unique solution to (10) in the set $\mathcal{M}^c$.

Remark. The family $\{\mathcal{X}^\vartheta : \vartheta \in \Theta\}$ depends on the function f in a weak way. The proof of the proposition shows the following: For $\mu \in \mathcal{M}^c$, let $\{\mathcal{Y}^\vartheta : \vartheta \in \Theta\}$ be a family of random processes with $\mathfrak{L}(\mathcal{Y}^\vartheta) = \mu$ for all $\vartheta \in \Theta$. Then, the sequence Q^ϑ_n initiated with $\{\mathcal{Y}^\vartheta : \vartheta \in \Theta\}$ converges almost surely uniformly to $\mathcal{X}^\vartheta$.

Proof of Proposition 13. For $n \geq 0$, by definition, almost surely,

$$\begin{aligned} Q^\vartheta_{n+1}(t) - Q^\vartheta_n(t) = \sum_{i=1}^K \mathbf{1}_{\Lambda^\vartheta_i}(t) \Bigg[&(\mathcal{R}^\vartheta_i)^\alpha \left[Q^{\vartheta i}_n(\varphi^\vartheta_i(t)) - Q^{\vartheta i}_{n-1}(\varphi^\vartheta_i(t)) \right] \\ &+ \sum_{j \in E_i} (\mathcal{R}^\vartheta_j)^\alpha \left[Q^{\vartheta j}_n(\xi^\vartheta_j) - Q^{\vartheta j}_{n-1}(\xi^\vartheta_j) \right] \Bigg]. \end{aligned} \tag{24}$$

Recalling the choice of α in (5), it follows that $\mathbf{E}[Q^\vartheta_n(\xi)] = c$ by induction on $n \geq 0$. Next, from (22), again by induction on n, it follows that the sequences $(Q^\vartheta_n)_{n\geq 0}, \vartheta \in \Theta$, are identically distributed. Abbreviate $\Delta Q^\vartheta_n := Q^\vartheta_{n+1} - Q^\vartheta_n$. From the last display, again by induction on n, one verifies that, for $n \geq 0$ fixed, the random variables $\Delta Q^\vartheta_n, \vartheta \in \Theta$, are identically distributed. Recall the definition of β_i in (6). From (24), it follows that

$$\begin{aligned} \mathbf{E}[\Delta Q^\vartheta_n(\xi)^2] &= \sum_{i=1}^K \mathbf{E}[\beta_i^{2\alpha}] \mathbf{E}[\Delta Q^\vartheta_{n-1}(\xi)^2] \\ &\quad + \sum_{i=1}^K \sum_{j_1 \neq j_2 \in E_i \cup \{i\}} \mathbf{E}\left[\mathbf{1}_{\Lambda^\vartheta_i}(\xi) (\mathcal{R}^\vartheta_{j_1})^\alpha (\mathcal{R}^\vartheta_{j_2})^\alpha \Delta Q^{\vartheta j_1}_{n-1}(\xi^\vartheta_{j_1}) \Delta Q^{\vartheta j_2}_{n-1}(\xi^\vartheta_{j_2}) \right]. \end{aligned}$$

Given $\xi \in \Lambda_i^\vartheta$ and the values of $\mathcal{R}_{j_1}^\vartheta$ and $\mathcal{R}_{j_2}^\vartheta$, the random variables $\Delta Q_{n-1}^{\vartheta j_1}(\xi_{j_1}^\vartheta)$ and $\Delta Q_{n-1}^{\vartheta j_2}(\xi_{j_2}^\vartheta)$ are zero-mean independent random variables. Thus, the second term in the last display vanishes. It follows that

$$\mathbf{E}[\Delta Q_n^\vartheta(\xi)^2] \leq \sum_{i=1}^K \mathbf{E}[\beta_i^{2\alpha}]\mathbf{E}[\Delta Q_{n-1}^\vartheta(\xi)^2] \leq C\left(\sum_{i=1}^K \mathbf{E}[\beta_i^{2\alpha}]\right)^n,$$

for some $C > 0$ recalling that $\sum_{i=1}^K \mathbf{E}[\beta_i^{2\alpha}] < 1$.

Next, we aim at showing that, for all $m \geq 1$, we have $\mathbf{E}[|\Delta Q_n^\vartheta(\xi)|^m] \leq C_m q_m^n$ for some constants $C_m > 0$ and $0 < q_m < 1$. The last display verifies this claim for $m = 1, 2$. Assume it is true for all $\ell \leq m-1$ and let $C_* = \max(C_1, \ldots, C_{m-1})$, $q_* = \max(q_1, \ldots, q_{m-1}) < 1$. Then, again from (24), we deduce that

$$\mathbf{E}[|\Delta Q_n^\vartheta(\xi)|^m] \leq \sum_{i=1}^K \mathbf{E}[\beta_i^{m\alpha}]\mathbf{E}[|\Delta Q_{n-1}^\vartheta(\xi)|^m] + \sum_{i=1}^K \sum_{j_1,\ldots,j_m} \mathbf{E}\left[\mathbf{1}_{\Lambda_i^\vartheta}(\xi)\prod_{k=1}^m |\Delta Q_{n-1}^{\vartheta j_k}(\xi_{j_k}^\vartheta)|\right], \quad (25)$$

where the sum ranges over all $j_1, \ldots, j_m \in E_i \cup \{i\}$ which are not all identical. Now, similarly to the argument above, on the event that $\xi \in \Lambda_i^\vartheta$, the random variables in the product on the right hand side of the last display are independent for different values of j_k. Thus, using the induction hypothesis, we deduce

$$\mathbf{E}[|\Delta Q_n^\vartheta(\xi)|^m] \leq \sum_{i=1}^K \mathbf{E}[\beta_i^{m\alpha}]\mathbf{E}[|\Delta Q_{n-1}^\vartheta(\xi)|^m] + \sum_{i=1}^K \sum_{j_1,\ldots,j_m} (C_* q_*^{(n-1)})^{\#\{j_1,j_2,\ldots,j_m\}}, \quad (26)$$

with $j_1, \ldots, j_m$ as in the sum in (25). From here, a simple induction over n shows that $\mathbf{E}[|\Delta Q_n^\vartheta(\xi)|^m]$ decays exponentially in n.

The decay exponential decay at uniform point can then be used to prove uniform exponential decay: Let $m \geq 1$ be enough large such that $\sum_{i=1}^K \mathbf{E}[\mathcal{R}_i^{m\alpha}] < 1$. Define $C_{**} = \max(C_1, \ldots, C_m)$ and $q_{**} = \max(q_1, \ldots, q_m)$. Then, it follows from (24) that

$$\mathbf{E}[\|\Delta Q_n^\vartheta\|^m] \leq \sum_{i=1}^K \mathbf{E}[\mathcal{R}_i^{m\alpha}]\mathbf{E}[\|\Delta Q_{n-1}^\vartheta\|^m] + \sum_{i=1}^K \sum_{j_1,\ldots,j_m} \mathbf{E}[\|\Delta Q_n^\vartheta\|^{\ell_i}] C_{**}^{\ell_i} q_{**}^{(m-\ell_i)(n-1)},$$

with $j_1, \ldots, j_m$ as in the sum in (26) where $\ell_i = \ell_i(j_1, j_2, \ldots, j_m)$ denotes the number of indices equal to i. Using the bound $\mathbf{E}[\|\Delta Q_n^\vartheta\|^\ell] \leq \mathbf{E}[\|\Delta Q_n^\vartheta\|^m]^{\ell/m}$, another induction on n shows that $\mathbf{E}[\|\Delta Q_n^\vartheta\|^m]$ decays exponentially since $\sum_{i=1}^K \mathbf{E}[\mathcal{R}_i^{m\alpha}] < 1$. From there, standard arguments (see, e.g. the proof Theorem 6 in [17]) imply that, almost surely, Q_n^ϑ converges uniformly and, writing $\mathcal{X}^\vartheta$ for its limit, the identity in (23) holds. Furthermore, the random variable $\|\mathcal{X}^\vartheta\|$ has finite polynomial moments of all orders.

Let $\mu \in \mathcal{M}^c$ and define a family of random processes W_n^ϑ analogously to Q_n^ϑ in (22) where W_n^ϑ is initiated by a set of independent random variables $\{W_0^\vartheta, \vartheta \in \Theta\}$ having distribution μ. The same arguments as above show that $(W_n^\vartheta)_{n\geq 0}$ converges uniformly to a solution W^ϑ of (10). Furthermore, as before, one can show that, first, $\mathbf{E}[(Q_n^\vartheta(\xi) - W_n^\vartheta(\xi))^2] \to 0$ exponentially fast and then $\mathbf{E}[\|Q_n^\vartheta - W_n^\vartheta\|^m] \to 0$ for all $m \geq 1$. Hence, $W^\vartheta = \mathcal{X}^\vartheta$ almost surely. This implies the uniqueness statement in the proposition as well as the claim in the remark. □

Proof of Theorem 1 i) and ii). ii) is an immediate consequence of Proposition 12 and Proposition 11 *ii)* since $\mathbf{E}\left[\nu^{\mathfrak{X}}\right]$ remains invariant upon replacing X by the support of the measure. The uniqueness claim for the process X in Theorem 1 *i)* can be deduced as follows: Let $\mathcal{X}$ be the process constructed in Proposition 13, and assume that Y is a continuous excursion satisfying (10) with $\mathbf{E}[Y(\xi)] = \mathbf{E}[\mathcal{X}(\xi)]$. Then, by Proposition 12, $\mathbf{E}[\nu_n^{\mathfrak{T}_\mathcal{X}}] = \mathbf{E}[\nu_n^{\mathfrak{T}_Y}]$ for all $n \geq 1$. Let $f_n : \mathbb{M}_n \to \mathbb{R}$ be defined by $f_n(\mathbf{m}) = \sup\{m_{0,i} : 0 \leq i \leq n\}$, where $\mathbf{m} = (m_{i,j})_{0\leq i,j\leq n}$. Clearly, as $n \to \infty$, weakly,

$$f_n\big(\mathbf{E}[\nu_n^{\mathfrak{T}_\mathcal{X}}]\big) \to \mathfrak{L}(\|\mathcal{X}\|), \quad \text{and} \quad f_n\big(\mathbf{E}[\nu_n^{\mathfrak{T}_Y}]\big) \to \mathfrak{L}(\|Y\|).$$

Hence, $\|Y\|$ must be distributed as $\|\mathcal{X}\|$, and in particular, by Proposition 13, $\|Y\|$ must have finite moments of all orders. The uniqueness under the finite moment condition in Proposition 13 then implies that $\mathfrak{L}(Y) = \mathfrak{L}(\mathcal{X})$. $\square$

4.3 Attractiveness of the fixed-points of (4)

In this section, we prove Theorem 2 and construct the counter-examples of Proposition 3. We start with the following lemma that provides a height function representation of random elements in $\mathbb{T}_{\mathrm{f}}^{\mathrm{GHP}}$.

Lemma 14. *Let $\mathfrak{T}$ be a random variable on $\mathbb{T}_{\mathrm{f}}^{\mathrm{GHP}}$. Then, there exists a random variable X on $\mathcal{C}_{ex}$ such that $\mathfrak{T}_X$ and $\mathfrak{T}$ are identically distributed.*

Proof. Let $\mathfrak{T} \in \mathbb{T}_{\mathrm{f}}^{\mathrm{GHP}}$ fixed. Then, for any representative t of the equivalence class $\mathfrak{T}$, by the construction using the shuffling order on the branch points of t in [28], there exists a random variable $x(t)$ on $\mathcal{C}_{\mathrm{ex}}$ such that $\mathfrak{T}_{x(t)} = \mathfrak{T}$. The distribution of $x(t)$ does not depend on the choice of t. Hence, there exists a map $\sigma : \mathbb{T}_{\mathrm{f}}^{\mathrm{GHP}} \to \mathcal{M}_1(\mathcal{C}_{\mathrm{ex}})$ with $\tau_*(\sigma(\mathfrak{T})) = \delta_{\mathfrak{T}}$. It is easy to check that, for random $\mathfrak{T}$, a random variable X with law $\mathbf{E}[\sigma(\mathfrak{T})]$ (the annealed version of $\sigma(\mathfrak{T})$) has the desired property that $\mathfrak{T}_X$ is distributed like $\mathfrak{T}$. $\square$

Proof of Theorem 2. ii) This follows easily from the proof of Proposition 13. To see this, let ν^* be a law on $\mathcal{C}_{\mathrm{ex}}$ such that the (isometry class of the) tree encoded by a realization of ν^* has distribution ν. Then, $\phi^n(\nu)$ is equal to the distribution of the real tree encoded by $Q_n^{\emptyset}$ from the proof of Proposition 13 initiated by independent realizations of ν^* at level n of Θ. As we have seen in this proof, Q_n^{ϑ} converges almost surely to a solution of (4) (It is here where we need the moment assumption on $\|\mathcal{T}\|$.). This proves the claim.

i) For $m, n \geq 1$, write $\tilde{D}_n^{(m)}$ for a generic realization of the annealed distance matrix $\mathbf{E}[\nu_n^{\mathfrak{T}(m)}]$ where $\mathfrak{T}(m)$ has distribution $\phi^m(\nu)$. Recall that, for $n \geq 1$, the random distance matrix $D_n = (d(\zeta_i, \zeta_j))_{0 \leq i,j \leq n}$ of the unique solution to (10) satisfies the fixed-point equation (21), which we recall here for convenience

$$D_n \stackrel{d}{=} A_n D_n + B_n. \tag{27}$$

Here, given the set of random variables $\mathcal{R}, \mathcal{S}, (J_i)_{1 \leq i \leq K}$, in distribution, A_n is composed of $\mathcal{R}, \mathcal{S}, (J_i)_{1 \leq i \leq K}$, and B_n can be expressed as a linear combination of a family of independent copies of $D_1, \ldots, D_{n-1}$ whose coefficients depend only on $\mathcal{R}, \mathcal{S}, (J_i)_{1 \leq i \leq K}$. Furthermore, the proof of (21) also shows that, almost surely, A_n and all coefficients in B_n take values in $[0, 1)$. Since $\phi^{m+1} = \phi \circ \phi^m$, the arguments from the proof of Proposition 12 show that

$$\tilde{D}_n^{(m+1)} \stackrel{d}{=} A_n \tilde{D}_n^{(m)} + B_n^{(m)},$$

where (A_n, B_n^m) and $\tilde{D}_n^{(m)}$ are independent and $(A_n, B_n^{(m)})$ can be expressed as in (27) upon replacing any copy of D_i in B_n by $\tilde{D}_i^{(m)}$. (Of course, we maintain the independence between $\mathcal{R}, \mathcal{S}, (J_i)_{1 \leq i \leq K}$ and the copies of $\tilde{D}_1^{(m)}, \ldots, \tilde{D}_{n-1}^{(m)}$.)

The remainder of the proof consist in showing that, in distribution (and in mean), for all $n \geq 1$, we have $\tilde{D}_n^{(m)} \to D_n$ as $m \to \infty$. By Proposition 11 (iii), this implies the assertion. Our proof relies on a contraction argument. In the following, we identify $\mathbb{M}_n$ with $\mathbb{R}^{\binom{n+1}{2}}$. For $p \in \mathbb{N}, p \geq 1$, let $\mathcal{M}_1^1(\mathbb{R}^p)$ be the set of probability measures on $\mathbb{R}^p$ whose Euclidean norm is integrable. Recall the *Wasserstein* distance on $\mathcal{M}_1^1(\mathbb{R}^p)$ defined by, for $\mu_1, \mu_2 \in \mathcal{M}_1^1(\mathbb{R}^p)$,

$$\ell_1(\mu_1, \mu_2) = \inf\{\mathbf{E}[\|X - Y\|] : \mathfrak{L}(X) = \mu_1, \mathfrak{L}(Y) = \mu_2\}.$$

We abbreviate $\ell_1(X, Y) := \ell_1(\mathfrak{L}(X), \mathfrak{L}(Y))$ for random variables X, Y in $\mathbb{R}^p$. We proceed by induction on $n \geq 1$, and assume that, for all $1 \leq i < n$, $\ell_1(\tilde{D}_i^{(m)}, D_i) \to 0$ as $m \to \infty$. Clearly, since the random coefficients appearing in B_n and $B_n^{(m)}$ are bounded, it follows that $\ell_1(B_n^{(m)}, B_n) \to 0$ as $m \to \infty$. A standard contraction argument shows that

$$\ell_1(\tilde{D}_n^{(m+1)}, D_n) \leq \mathbf{E}[A_n]\ell_1(\tilde{D}_n^{(m)}, D_n) + \ell_1(B_n^{(m)}, B_n).$$

From here, since $\mathbf{E}[A_n] < 1$ and $\ell_1(B_n^{(m)}, B_n) \to 0$, it follows easily that $\ell_1(\tilde{D}_n^{(m)}, D_n) \to 0$ as $m \to \infty$. The assertion now follows by induction on n upon establishing the base case $n = 1$. Here, for the sake of convenience, in the rest of the proof, we use the notation $\tilde{D}_1^{(m)}, D_1$ when referring to the entry (1,2) in the 2 by 2 distance matrix. Distributional convergence $\tilde{D}_1^{(m)} \to D_1$ as $m \to \infty$ follows immediately from Theorem 2 b) in [31]. Furthermore, $\mathbf{E}[\tilde{D}_1^{(m)}] = \mathbf{E}[D_1]$ for all $m \geq 1$. It is well-known that convergence in ℓ_1 for non-negative random variables is equivalent to distributional convergence together with convergence of the mean, see, e.g. [14, Lemma 8.3]. This concludes the proof. □

Proof of Theorem 3. i) The example we provide generalizes the one by Albenque and Goldschmidt [3] in the special case of Example **1**). Take $\mathcal{R} = \mathcal{S}$. Let ν_c be the unique law solving of (4) in $\mathbb{T}_{\mathrm{f}}^{\mathrm{GHP}}$ with $\mathbf{E}[d_c(\rho_c, \zeta_c)] = c$, where ζ_c is sampled according to μ_c on $\mathcal{T}_c$ and $(\mathcal{T}_c, d_c, \mu_c, \rho_c)$ has distribution ν_c. Such a solution exists by Theorem 1. The idea is to construct another random compact rooted measured real tree in $\mathbb{T}^{\mathrm{GHP}}$ by appending massless hair to a tree sampled from ν_c.

Choose an integer $\kappa \geq 1$, and fix the distribution of $\mathcal{S}$. Let $\mathfrak{X} \in \mathbb{K}^{\mathrm{GHP}}$ and choose a representative (X, d, μ, ρ) of $\mathfrak{X}$. Let $\mathcal{P}$ be a Poisson point process with intensity measure $\mu \otimes x^{-(1/\alpha+1)}dx$ on $X \times [0, \infty)$. For every point $(u, x) \in \mathcal{P}$, take κ disjoint segments of length x, glue all of them at the point $u \in X$, each by one extremity. Clearly, only finitely many of the appended segments have length more than ϵ, for any $\epsilon > 0$, and the resulting space is almost surely compact. The distribution of the isometry class of the resulting compact rooted measured space does not depend on the choice of representative. Hence, this operation defines a map $\psi_1 : \mathbb{K}^{\mathrm{GHP}} \to \mathcal{M}_1(\mathbb{K}^{\mathrm{GHP}})$, and a version at the level of measures $\Psi_1 : \mathcal{M}_1(\mathbb{K}^{\mathrm{GHP}}) \to \mathcal{M}_1(\mathbb{K}^{\mathrm{GHP}})$ that is then defined by

$$\Psi_1(\upsilon)(A) = \mathbf{E}[\psi_1(Y)(A)], \qquad \text{with} \qquad \mathfrak{L}(Y) = \upsilon.$$

The important observation is that Ψ and Ψ_1 commute. In other words, for a probability measure υ on $\mathbb{K}^{\mathrm{GHP}}$, we have

$$\Psi(\Psi_1(\upsilon), \mathfrak{L}(\mathcal{S}, \mathcal{S})) = \Psi_1(\Psi(\upsilon, \mathfrak{L}(\mathcal{S}, \mathcal{S}))).$$

It follows immediately, that for any fixed-point ν of (4), the measure $\Psi_1(\nu)$ also solves (4). In particular, $\Psi_1(\nu_c)$ is such a fixed-point and almost surely concentrated on $\mathbb{T}^{\mathrm{GHP}} \setminus \mathbb{T}_{\mathrm{f}}^{\mathrm{GHP}}$.

Let $\mathfrak{T} = (\mathcal{T}, d, \mu, \rho)$ be a random variable with distribution $\Psi_1(\rho_c)$. The height $\|\mathcal{T}\|$ is at least as large as the length of the longest segment, and so for $h > 0$ we have

$$\begin{aligned}\mathbf{P}(\|\mathcal{T}\| \geq h) &\geq \mathbf{P}(\mathrm{Po}(\textstyle\int_h^\infty x^{-(1/\alpha+1)}dx) \geq 1)\\ &= 1 - \exp(-\alpha h^{-1/\alpha}),\end{aligned}$$

where $\mathrm{Po}(\lambda)$ denotes a Poisson random variable with parameter λ. Since $\mathfrak{m} \geq 1/\alpha$ (see above (16)), and it follows readily that $\mathbf{E}[\|\mathcal{T}\|^{\mathfrak{m}}] = \infty$.

Furthermore, for different values of κ the corresponding laws are mutually singular. Finally, note that there is nothing specific to massless *segments* in the argument. In particular, we can replace the segments by loops, or equivalently identify the extremities of all the segments together when $\kappa \geq 2$. This proves that some fixed-points are real trees with probability zero.

ii) The only reason why the above example does not also prove the second claim is that, in the tree with distribution $\Psi_1(\nu_c)$, almost surely, the measure does not have full support. To resolve that, for any fixed compact measured rooted space, we construct a variant of ψ_1 where only the distance and the measure change. Here, we restrict ourselves to the subspace $\overline{\mathbb{K}}^{\text{GHP}}$ of isometry classes satisfying $\mathbf{E}[d(\rho,\zeta)] < \infty$ when ζ is sampled according to μ on X in the space (X,d,μ,ρ). First, we start with the same compact rooted space (X,d,μ,ρ) constructed from $\mathfrak{X}^\circ = (X^\circ, d^\circ, \mu^\circ, \rho^\circ)$ with the help of a Poisson process as above. Subsequently, we focus on the case in which $\kappa = 1$. Choose $\mu_{\mathcal{P}}$ on X as follows. We first set $\mu_{\mathcal{P}}(X^\circ) = 0$. Then, for $(u,x) \in \mathcal{P}$, we associate a total $\mu_{\mathcal{P}}$ mass $\min\{x,1\}^{1/\alpha+1}$; we distribute this mass along the segment of length x with density $\alpha/(1+\alpha) \cdot s^{1/\alpha}\mathbf{1}_{\{0\le s\le x\}}ds$ if $x \le 1$, and with density $e^{-s}(1-e^{-x})^{-1}\mathbf{1}_{\{0\le s\le x\}}ds$ if $x > 1$. Then,

$$\mathbf{E}[\mu_{\mathcal{P}}(X)] = \int_0^\infty \min\{x,1\}^{1/\alpha+1}x^{-(1/\alpha+1)}dx = \int_0^1 dx + \int_1^\infty x^{-(1/\alpha+1)}dx = 1+\alpha < \infty\,,$$

and it follows that $\mu_{\mathcal{P}}(X) < \infty$ almost surely. We let μ^* be the unique probability measure on X that is proportional to $\mu+\mu_{\mathcal{P}}$. For a random point ζ^* sampled according to μ^*, the expected distance $\mathbf{E}[d(\rho,\zeta^*)]$ is finite. Indeed,

$$\mathbf{E}[d(\rho,\zeta^*)] \le \mathbf{E}\left[\int d(\rho,u)(\mu+\mu_{\mathcal{P}})(du)\right] \le \mathbf{E}[d(\rho,\zeta)] + \mathbf{E}\left[\int d(\rho,u)\mu_{\mathcal{P}}(du)\right]\,,$$

and

$$\mathbf{E}\left[\int d(\rho,u)\mu_{\mathcal{P}}(du)\right] \le \|X\| + 1 + \sup_{x\ge 1}\int_0^x e^{-s}(1-e^{-x})^{-1}ds = \|X\|+2.$$

Thus, since $\mathfrak{X} \in \overline{\mathbb{K}}^{\text{GHP}}$, we have $\lambda^* := \mathbf{E}[d(\rho,\zeta^*)] < \infty$, and so we can rescale the metric and define $d^*(\cdot,\cdot) = d(\cdot,\cdot)\times c/\lambda^*$. The distribution of the GHP-isometry class of $\mathfrak{X}^* = (X,d^*,\mu^*,\rho)$ satisfying $\mathbf{E}[d(\rho,\zeta^*)] = c$ obtained in this way only depends on the GHP-equivalence class of $\mathfrak{X}^\circ$. Thus, we can define the corresponding maps $\psi_2 : \overline{\mathbb{K}}^{\text{GHP}} \to \mathcal{M}_1(\overline{\mathbb{K}}^{\text{GHP}})$ and $\Psi_2 : \mathcal{M}_1(\overline{\mathbb{K}}^{\text{GHP}}) \to \mathcal{M}_1(\overline{\mathbb{K}}^{\text{GHP}})$ as above. Note that, as opposed to ψ_1 and Ψ_1, we now have $\psi_2(\overline{\mathbb{K}}^{\text{GHP}}_{\text{f}}) \subseteq \mathcal{M}_1(\overline{\mathbb{K}}^{\text{GHP}}_{\text{f}})$ and $\Psi_2(\mathcal{M}_1(\overline{\mathbb{K}}^{\text{GHP}}_{\text{f}})) \subseteq \mathcal{M}_1(\mathbb{K}^{\text{GHP}}_{\text{f}})$. It remains only to check that, if $\nu^* = \Psi_2(\nu_c)$, then $(\phi^n_{\text{GHP}}(\nu^*))_{n\ge1}$ does not converge to the law of $\mathfrak{T}_X$ where X solves (4) with $\mathbf{E}[X(\xi)] = c$. For this, we prove that in any potential limit of $(\phi^n_{\text{GHP}}(\nu^*))_{n\ge1}$ the measure does not have full support. To see this, it suffices to observe that the maximum diameter of one of the rescaled copies $\mathfrak{T}^*_i$, $i = 1,\dots,K^n$, of $\mathfrak{T}$ used to construct $\phi^n_{\text{GHP}}(\nu^*)$ does not tend to zero with n, while the maximum mass of one of these does tend to zero with n; The claim about the mass is plainly obvious. Now, to see the that maximal diameter does not tend to zero, just observe that up to a fixed constant factor the metric part of $\mathfrak{T}^*$ is that of $\mathfrak{T}$, and that $\mathfrak{T}$ is a fixed-point of (4). Thus, for every $n\ge1$, the collection of the length of the hair in the union of $\mathfrak{T}^*_i$, $1\le i\le K^n$, has the same distribution as the one in $\mathfrak{T}$; in particular, the maximal length does not tend to zero, and since the corresponding segment lies within of some $\mathfrak{T}^*_i$, $1\le i\le K^n$, this completes the proof. □

5 Proofs of the geometric properties

5.1 The upper Minkowski dimension: Proof of Theorem 4

In the following, for the sake of clarity, we write aX for the metric space (X,ad), where a constant $a>0$ and a metric space (X,d). The approach here is to turn the heuristic presented in (17) into a rigorous argument. In this direction, observe that, for any compact real tree $(\mathcal{T},d)$ and $a>0$, we have $N_{a\mathcal{T}}(x) = N_{\mathcal{T}}(x/a)$. Second, for any deterministic compact rooted measured real trees $\mathfrak{T}_1,\dots,\mathfrak{T}_K$ where $\mathfrak{T}_i = (\mathcal{T}_i,d_i,\mu_i,\rho_i)$, and $r,s\in\Sigma_K, 0<\alpha<1$, write $\mathfrak{T} = (\mathcal{T},d,\mu,\rho)$ for the *deterministic* tree obtained by the construction in Section 2.3 with glue points $\eta_i, i\in[K]$ chosen arbitrarily but fixed. Then, we have the deterministic bounds

$$N_{\mathcal{T}}(x) \le \sum_{i=1}^K N_{r_i^\alpha\mathcal{T}_i}(x) \le N_{\mathcal{T}}(x) + K.$$

The lower bound follows from the fact that a union of coverings of a decomposition of a set also forms a covering of the set. Recall that, for $i = 2, \ldots, K$, the root ρ_i of $\mathcal{T}_i$ is identified with a point η_{ϖ_i} of $\mathcal{T}_{\varpi_i}$. As a consequence, any open covering of the tree $\mathcal{T}$ can be extended into a collection of open coverings of the K subtrees upon adding at most K balls of radius x among those centered at some of the ρ_i, $i = 2, \ldots, K$ or at the η_j, $j = 1, \ldots, K$. This explains the upper bound.

Next we consider the case where the tree $\mathfrak{T} = (\mathcal{T}, d, \mu, \rho)$ satisfies the fixed-point equation (4) and we let $v(x) = \mathbf{E}[N_{\mathcal{T}}(x)]$. We assume for now that $v(x) < \infty$ for every $x > 0$ (Lemma 16 below proves that it is indeed the case). Then, using (10) and the two observations above, we have the bounds

$$v(x) \le \sum_{i=1}^{K} \mathbf{E}[v(\mathcal{R}_i^{-\alpha} x)] = K\mathbf{E}[v(\overline{\mathcal{R}}^{-\alpha} x)] \le v(x) + K. \tag{28}$$

Thus, $|v(x) - K\mathbf{E}[v(\overline{\mathcal{R}}^{-\alpha} x)]| \le K$ and one may first consider the approximate equation

$$\tilde{v}(x) = K\mathbf{E}[\tilde{v}(\overline{\mathcal{R}}^{-\alpha} x)]. \tag{29}$$

Recall that g denotes the density of the random variable $\overline{\mathcal{R}}$. If $g(t) = Ct^{\gamma-1}$, then a direct computation shows that all solutions $\tilde{v}$ of (29) which are bounded on any compact interval of $(0, \infty)$ are multiples of the function $x \mapsto x^{(\gamma-KC)/\alpha}$. Subsequently, our task will be to prove that these solutions to the approximate equation provide upper bounds for $v(x)$ as $x \to 0$. This is the statement of the following proposition.

Proposition 15. *Under the conditions of Theorem 4, as $x \to 0$, we have $\mathbf{E}[N_{\mathcal{T}}(x)] = O(x^{(\gamma-CK)/\alpha})$.*

Taking Proposition 15 for granted for now, we immediately provide the straightforward proof of Theorem 4.

Proof of Theorem 4. By Proposition 15, we have $\mathbf{E}[N_{\mathcal{T}}(x)] = O(x^{(\gamma-CK)/\alpha})$. Using Markov's inequality and the Borel–Cantelli lemma, this implies that, for any $\varrho > (CK-\gamma)/\alpha$, we have $N_{\mathcal{T}}(2^{-n})2^{-\varrho n} \to 0$ almost surely. Since

$$N_{\mathcal{T}}(2^{-n-1}) \le N_{\mathcal{T}}(2^{-n}) \le N_{\mathcal{T}}(2^{-n+1})$$

it follows $N_{\mathcal{T}}(x)x^{\varrho} \to 0$ with probability one. This concludes the proof. □

The first step in the proof of Proposition 15 is to show that $\mathbf{E}[N_{\mathcal{T}}(x)] < \infty$ for every $x > 0$.

Lemma 16. *Under the assumptions of Theorem 4, for all $x > 0$, the quantity $\mathbf{E}[N_{\mathcal{T}}(x)]$ is finite.*

Proof. Fix $0 < x < 1$. It suffices to show that $\sum_{n\ge1} K^n \mathbf{P}\left(N_{\mathcal{T}}(x) \ge K^n\right)$ is finite. Consider the representation of $\mathcal{T} = \mathcal{T}_{\mathcal{X}}$ in terms of K^n subtrees based on the expansion up to level n of Θ (see Section 3.4): if the K^n subtrees all have height smaller than x, then $\mathcal{T}$ can be covered by the union of balls of size x centered at the roots of the subtrees. In particular, in such a case, we have $N_{\mathcal{T}}(x) \le K^n$.

Let $(\overline{\mathcal{R}}_n)_{n\ge1}$ be a sequence of independent random variables distributed like $\overline{\mathcal{R}}$ and $E_n = -\log \overline{\mathcal{R}}_n$.

Then, the argument above implies that, with $c > 0$ and for all $n \geq (-\log x)/c$, we have

$$\begin{aligned}
\mathbf{P}\left(N_{\mathcal{T}}(x) \geq K^n\right) &\leq \mathbf{P}\left(\bigcup_{i_1,\dots,i_n \in [K]} \left\{\mathcal{R}^{\emptyset}_{i_1}\mathcal{R}^{i_1}_{i_2}\cdots\mathcal{R}^{i_1 i_2 \dots i_{n-1}}_{i_n}\|\mathcal{T}^{i_1\dots i_n}\| \geq x\right\}\right) \\
&\leq K^n \mathbf{P}\left(\prod_{i=1}^n \overline{\mathcal{R}}_i \cdot \|\mathcal{T}\| \geq x\right) \\
&\leq K^n \left[\mathbf{P}\left(e^{cn}\prod_{i=1}^n \overline{\mathcal{R}}_i \geq x\right) + \mathbf{P}\left(\|\mathcal{T}\| \geq e^{cn}\right)\right] \\
&\leq K^n \left[\mathbf{P}\left(\sum_{i=1}^n E_i \leq cn - \log x\right) + \mathbf{E}[\|\mathcal{T}\|^k]e^{-ckn}\right] \\
&\leq K^n \left[\mathbf{P}\left(\sum_{i=1}^n E_i \leq 2cn\right) + \mathbf{E}[\|\mathcal{T}\|^k]e^{-ckn}\right].
\end{aligned}$$

Now, for the sum of n terms not to exceed $2cn$, at least $\lfloor n/2 \rfloor$ of the summands shall be at most $4c$ and it follows that

$$\begin{aligned}
\mathbf{P}\left(N_{\mathcal{T}}(x) \geq K^n\right) &\leq K^n \binom{n}{\lfloor n/2 \rfloor}\mathbf{P}(E_1 \leq 4c)^{n/2} + \mathbf{E}[\|\mathcal{T}\|^k]K^n e^{-ckn} \\
&\leq (KC)^n (4c)^{(\gamma/2)n} + \mathbf{E}[\|\mathcal{T}\|^k]K^n e^{-ckn},
\end{aligned}$$

for all $k \geq 1$, some constant $C > 0$. Here, in the final step, we have used Stirling's formula and the fact that

$$\mathbf{P}\left(E_1 \leq 4c\right) = \mathbf{P}\left(\overline{\mathcal{R}} \geq e^{-4c}\right) \leq \mathbf{P}\left(\overline{\mathcal{R}} \geq 1 - 4c\right) \leq C_1(1 - (1-4c)^{\gamma}) \leq C_2(4c)^{\gamma},$$

for some universal constants $C_1, C_2 > 0$.

Since $\|\mathcal{T}\|$ has finite moments, first choosing $c > 0$ small enough and then k large enough shows that, for any $\varepsilon > 0$, there exists $C' > 0$ such that $\mathbf{P}(N_{\mathcal{T}}(x) \geq K^n) \leq C'\varepsilon^n$ still provided that $n \geq (-\log x)/c$. Choosing $\varepsilon \in (0, 1/K)$ shows that $\sum_{n \geq 1} K^n \mathbf{P}(N_{\mathcal{T}}(x) \geq K^n) < \infty$, and completes the proof. □

Proof of Proposition 15. As above, let $v(x) = \mathbf{E}[N_{\mathcal{T}}(x)]$. By definition and the previous lemma, v is finite and non-increasing. Furthermore, by the dominated convergence theorem, v is càdlàg. Thus, v is continuous almost everywhere and it suffices to consider the behavior at points of continuity; so from now on, we assume that x is a point of continuity of v. Using the observation in (28) and the assumed bound on the density g of $\overline{\mathcal{R}}$, we obtain

$$\begin{aligned}
v(x) &\leq K\mathbf{E}[v(\overline{\mathcal{R}}^{-\alpha}x)] \\
&= Kx^{1/\alpha}\int_x^{\infty} t^{-1/\alpha - 1}v(t)g(x^{1/\alpha}t^{-1/\alpha})dt \\
&\leq KC\alpha^{-1}x^{\gamma/\alpha}\int_x^{\infty} t^{-\gamma/\alpha-1}v(t)dt. \qquad (30)
\end{aligned}$$

We continue following the ideas in the proof of Gronwall's lemma. Let $r(x) = \int_x^{\infty} t^{-\gamma/\alpha-1}v(t)dt$. Then,

$$r'(x) = -x^{-\gamma/\alpha-1}v(x) \geq -CK\alpha^{-1}x^{-1}r(x).$$

We have $\alpha > 0$ and we now suppose that $x \leq \alpha$. Let $I(x) = CK\alpha^{-1}(\log x - \log \alpha) \leq 0$ and note that $I'(x) = CK\alpha^{-1}x^{-1}$. Then,

$$\frac{d}{dx}(r(x)e^{I(x)}) = r'(x)e^{I(x)} + r(x)I'(x)e^{I(x)} \geq 0.$$

Integrating yields

$$r(\alpha) - r(x)e^{I(x)} = \int_x^\alpha \frac{d}{dy}(r(y)e^{I(y)})dy \geq 0$$

Re-arranging the terms, we obtain

$$r(x) \leq r(\alpha)\alpha^{CK\alpha^{-1}} x^{-CK/\alpha}.$$

Using (30), it follows immediately that, as $x \to 0$, $v(x) = O(x^{(\gamma - CK)/\alpha})$. □

5.2 Lower bound on the Hausdorff dimension: Proof of Theorem 5

In order to find good lower bounds on the Hausdorff dimension of a metric space (S, d), one attempts to construct measures μ on its Borel σ-algebra such that $\mu(A)$ roughly behaves like $d(A)^s$ where $d(A)$ denotes the diameter of a set A and s is as high as possible. This can be quantified by the following expression relying on Frostman's Lemma: for a compact space (S, d), we have

$$\dim_{\mathrm{H}}(S) = \sup_\mu \sup \left\{ s \geq 0 : \sup_{x \in S, r > 0} r^{-s} \mu(B_r(x)) < \infty \right\},$$

where the first supremum ranges over all probability measures on (S, d) and $B_r(x)$ is denotes the ball of radius r around x in S.

For a real trees encoded by a function $f \in \mathcal{C}_{\mathrm{ex}}$, natural candidates for μ are push-forward measures $\mu_f^* := \mu^* \circ \pi_f^{-1}$ of measures μ^* on $[0,1]$ under the projection $\pi_f : [0,1] \to \mathcal{T}_f$. In our setting, in the case $\mathcal{R} = \mathcal{S}$ which covers both examples in (11) and (12), it is intuitive that the Lebesgue measure on $[0,1]$, or, equivalently, the canonical measure on $\mathcal{T}$, leads to an efficient choice. But when scaling factors in time and space are independent such as in example (13), one first has to find an appropriate time-change of the unit interval in order to re-correlate the masses of fragments in the tree with the extent of distances in the corresponding subtrees. Such a time change is constructed in Proposition 17. Furthermore, the measures constructed on the unit interval in this context are typically random; thus, one is led to build a pair of both the measure and the tree simultaneously. It is this situation in which the almost sure construction in Section 4.2 turns out especially useful.

In the context of the next proposition, we denote by Φ_1 the extension of the map Φ defined in (9) with $\alpha = 1$ to the space $\mathcal{C}_1 = \{f \in \mathcal{C} : f \geq 0, f(0) = 0, f(1) = 1\}$. Furthermore, for $i \in \Gamma^o$ let $\Lambda_i^{\vartheta-}$ and $\Lambda_i^{\vartheta+}$ be the two half-open intervals forming Λ_i^ϑ where $\inf \Lambda_i^{\vartheta-} \leq \sup \Lambda_i^{\vartheta+}$.

Proposition 17. *Almost surely, for any $\vartheta \in \Theta$, there exists a probability measure μ^ϑ on $[0,1] = \Lambda_\emptyset^\vartheta$, such that, for every $\sigma \in \Theta$,*

$$\mu^\vartheta(\Lambda_\sigma^\vartheta) = \mathcal{V}(\vartheta\sigma)/\mathcal{V}(\vartheta). \tag{31}$$

For the distribution function τ^ϑ of μ^ϑ, we have

$$\tau^\vartheta = \Phi_1(\tau^{\vartheta 1}, \ldots, \tau^{\vartheta K}, \mathcal{R}^\vartheta, \mathcal{S}^\vartheta, \Xi^\vartheta). \tag{32}$$

In other words,

$$\tau^\vartheta(\cdot) = \sum_{i=1}^K \mathbf{1}_{\Lambda_i^\vartheta}(\,\cdot\,) \left[\mathcal{R}_i^\vartheta \tau^{\vartheta i}(\varphi_i^\vartheta(\,\cdot\,)) + \sum_{j \in E_i} \mathcal{R}_j^\vartheta \tau^{\vartheta j}(\xi_j^\vartheta) + \mathbf{1}_{\Lambda_i^{\vartheta+}}(\,\cdot\,) \sum_{j \in \Gamma_i \setminus \{i\}} \mathcal{R}_j^\vartheta \right]. \tag{33}$$

Furthermore, in distribution, τ^ϑ is the unique continuous distribution function on $[0,1]$ satisfying (32). τ^ϑ is measurable with respect to $\{\mathcal{R}^{\vartheta\sigma}, \mathcal{S}^{\vartheta\sigma}, \Xi^{\vartheta\sigma} : \sigma \in \Theta\}$, and $\mathbf{E}[\tau^\vartheta(\xi)] = 1/2$.

Remark. Note that there is no free multiplicative scaling for the set of solutions to (32). This is because we restrict ourselves to bounded continuous solutions and the fixed-point equation (32) (or, equivalently (33)) contains an additive term. This is also reminiscent of the fact that $\tau^\vartheta(1) = 1$.

Proof. Let μ_n^ϑ be the unique probability measure on $[0, 1]$ which corresponds to the mass distribution on the partition $\{\Lambda_\sigma^\vartheta : \sigma \in \Theta_n\}$ such that $\mu_n^\vartheta(\Lambda_\sigma^\vartheta) = \mathcal{V}(\vartheta\sigma)/\mathcal{V}(\vartheta)$, and μ_n^ϑ has constant density on each of the sets Λ_σ^ϑ, $\sigma \in \Theta_n$. This construction is consistent in the sense that, for $m \geq n$ and $\sigma \in \Theta_n$, we have

$$\mu_m^\vartheta(\Lambda_\sigma^\vartheta) = \mu_n^\vartheta(\Lambda_\sigma^\vartheta). \tag{34}$$

The measure μ^ϑ is constructed as the almost sure limit of the sequence of random measures $(\mu_n^\vartheta)_{n\geq 0}$. Denote by τ_n^ϑ the distribution function of μ_n^ϑ. By construction, almost surely,

$$\tau_n^\vartheta = \Phi_1(\tau_{n-1}^{\vartheta 1}, \dots, \tau_{n-1}^{\vartheta K}, \mathcal{R}^\vartheta, \mathcal{S}^\vartheta, \Xi^\vartheta).$$

Note that $\{(\tau_n^\vartheta) : \vartheta \in \Theta\}$ is a family of identically distributed random variables. The uniform convergence of τ_n^ϑ is shown analogously to convergence of Q_n^ϑ in Proposition 13 and we omit the details. Here, it is important to note that, since $\tau_0^\vartheta(t) = t$ for all $t \in [0, 1]$, one can verify inductively that $\mathbf{E}[\tau_n^\vartheta(\xi)] = 1/2$ for all $n \geq 1$. The relevant constant that assumes a value smaller than one, and that yields convergence of τ_n^ϑ at a uniformly chosen point is $\sum_{i=1}^K \mathbf{E}[\beta_i^2]$. By the uniform convergence, the limit τ^ϑ satisfies (32) and is a continuous distribution function. The induced measure μ^ϑ satisfies the desired properties. $\square$

Finally, the lower bound on the Hausdorff dimension relies on the tail behavior of the depth of a uniformly chosen point and the distance between two uniformly chosen points in $\mathcal{T}$.

Lemma 18. *Assume that $\mathfrak{T} = (\mathcal{T}, d, \mu, \rho)$ satisfies (4), that $\mathbf{P}(\|\mathcal{T}\| > 0)$, and that $\mathbf{E}[\overline{\mathcal{R}}^{-s}] < \infty$ for some $s > 0$. Furthermore, let ζ, ζ' be two vertices sampled independently according to μ. Then, there exists $\delta > 0$, such that,*

$$\mathbf{P}(d(\rho, \zeta) < r) = O(r^\delta), \quad \textit{and} \quad \mathbf{P}(d(\zeta, \zeta') < r) = O(r^\delta), \quad r \downarrow 0.$$

Remark. Note that although in most classical examples, we have invariance by rerooting at a random point distributed according to the mass measure, this may not be the case and the claims for $d(\rho, \zeta)$ and for $d(\zeta, \zeta')$ are in general both necessary.

Proof. Write $Y = d(\rho, \zeta)$. From Theorem 1 *iii)*, we know that $Y > 0$ almost surely. We use a version of the complete K-ary tree Θ as introduced in Section 3.4, where each node $\vartheta \in \Theta$ is assigned an additional pair of random variables $(J^\vartheta, Y^\vartheta)$, Y^ϑ being a copy of Y, independent of all remaining quantities and $\mathbf{P}\left(J^\vartheta = i \,|\, \mathcal{S}^\vartheta\right) = \mathcal{S}_i^\vartheta$. Similarly to the construction of the process Q_n^ϑ in the proof of Proposition 13, for all $n \geq 0$, we now define a set of random variables $\{Y_n^\vartheta : \vartheta \in \Theta\}$. First, let $Y_0^\vartheta = Y^\vartheta$ and, by induction, define

$$Y_n^\vartheta = \sum_{i=1}^K \mathbf{1}_{\{J^\vartheta = i\}} \left[(\mathcal{R}_i^\vartheta)^\alpha Y_{n-1}^{\vartheta i} + \sum_{j \in E_i} (\mathcal{R}_j^\vartheta)^\alpha Y_{n-1}^{\vartheta j} \right].$$

By (6), for all $n \geq 0$ and all $\vartheta \in \Theta$, Y_n^ϑ is distributed as Y.

Let $m_n = \min\{\mathcal{L}_\vartheta : |\vartheta| = n\}$ be the minimum Lebesgue measure of a set Λ_ϑ for $\vartheta \in \Theta_n$, and

$$\nu = \mathfrak{L}\left(1 + |E_J|\right).$$

Note that $\nu(\{0\}) = 0$ and $\nu(\{1\}) < 1$. Then, $Y_n^\emptyset$ is the sum of a number of independent copies of Y (this number of copies is hereafter denoted by N_n) each of them multiplied by a random factor not smaller than m_n^α. N_n is the distributed as the number of leaves in the family of a discrete-time branching

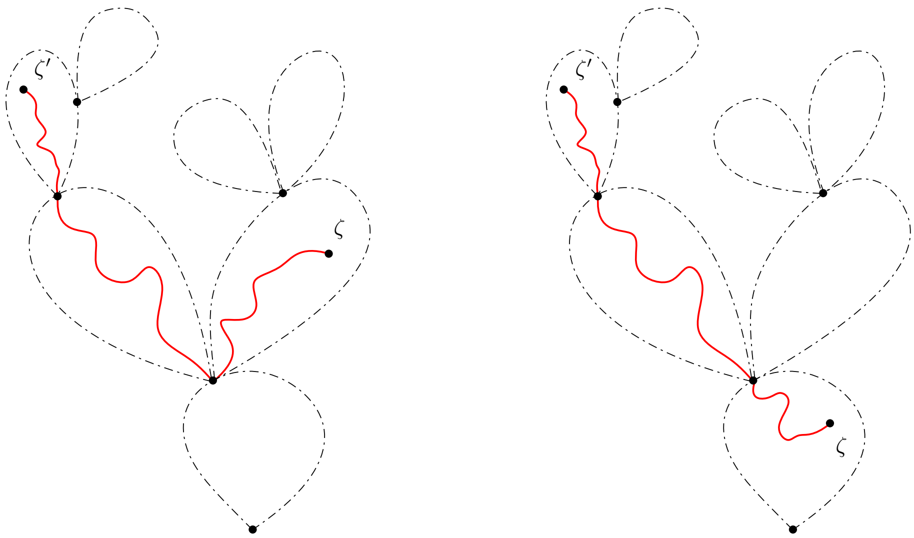

Figure 3: *The different cases in the recursion for the distance between two random nodes ζ and ζ' (the easier case where both random points fall within the same subtree is not illustrated).*

process with offspring distribution ν in generation n. Summarizing, with a sequence of independent copies $\tilde{Y}_1, \tilde{Y}_2, \ldots$ of Y and $c > 0$, we have for every $n \geq 1$,

$$\mathbf{P}\left(Y \leq r\right) = \mathbf{P}(Y_n^{\emptyset} \leq r) \leq \mathbf{P}\left(\sum_{i=1}^{N_n} e^{-c\alpha n}\tilde{Y}_i < r\right) + \mathbf{P}\left(m_n < e^{-cn}\right).$$

The following rough bound on N_n is sufficient: For $0 < \gamma < \nu([1,\infty))$, we have

$$\mathbf{P}\left(N_n < \lceil \gamma n \rceil\right) \;\leq\; \mathbf{P}\left(\mathrm{Bin}(n, \nu([1,\infty))) < \lceil \gamma n \rceil\right) \;\leq\; Ce^{\varepsilon n},$$

for some $C > 0$ and $\varepsilon > 0$ both depending on γ. Next, choose $s > 0$ such that $\mathbf{E}[\overline{\mathcal{R}}^{-s}] < \infty$. Then,

$$\mathbf{P}\left(m_n < e^{-cn}\right) \leq K^n \mathbf{P}\left(\prod_{i=1}^{n} \overline{\mathcal{R}}_i < e^{-cn}\right) \leq e^{\left(\log\left(K\mathbf{E}[\overline{\mathcal{R}}^{-s}]\right) - cs\right)n}.$$

Choosing c sufficiently large, the right-hand side is $O(e^{-\delta n})$ for some $\delta > 0$. Hence, with $n = \lceil C \log(1/r) \rceil$,

$$\begin{aligned}\mathbf{P}\left(Y \leq r\right) &\leq \mathbf{P}\left(\sum_{i=1}^{\lceil \gamma n \rceil} e^{-c\alpha n}\tilde{Y}_i \leq r\right) + O(e^{-\min(\delta,\varepsilon)n}) \\ &\leq r^{-C\gamma \log \mathbf{P}(Y \leq r^{1-c\alpha C})} + O(r^{C\min(\delta,\varepsilon)}).\end{aligned}$$

Choosing $C < (\alpha c)^{-1}$ yields the assertion for $d(\rho, \zeta)$.

Let us now move on to the case of $D = d(\zeta, \zeta')$. A decomposition of this random variable has already been discussed in the course of the proof of Proposition 12: Let ζ, ζ' be chosen according to the mass measure on $\mathfrak{T}_{\mathcal{X}}$. With $\mathfrak{T}_i := \mathfrak{T}_{\mathcal{X}^i}$, it follows that either ζ and ζ' both lie in some $\mathcal{T}_i$, in which case the path entirely lies in the corresponding subtree; or ζ and ζ' lie in $\mathcal{T}_i$ and $\mathcal{T}_j$ for distinct i, j. In this latter case, assuming without loss of generality that $i < j$, the structure of the decomposition depends on whether $i \in E_j$ or not. Let $q := q(i,j) := \max(E_i \cap E_j)$. Recall that $\mathbf{P}(J = i \mid \mathcal{S}) = \mathcal{S}_i$. With random variables J_1, J_2 that are independent copies of J given $\mathcal{S}$, we have (see Figure 3)

$$
\begin{aligned}
D \stackrel{d}{=} & \sum_{i=1}^{K} \mathbf{1}_{\{J_1=J_2=i\}} \mathcal{R}_i^{\alpha} D \\
& + \sum_{i<j\leq K} \mathbf{1}_{\{J_1=i,J_2=j \text{ or } J_1=j,J_2=i\}} \mathbf{1}_{\{i\notin E_j\}} \left(\sum_{k:q<k\in E_i} \mathcal{R}_k^{\alpha} Y^{(k)} + \sum_{l:q<l\in E_j} \mathcal{R}_l^{\alpha} Y^{(l)} \right) \\
& + \sum_{i<j\leq K} \mathbf{1}_{\{J_1=i,J_2=j \text{ or } J_1=j,J_2=i\}} \mathbf{1}_{\{i\in E_j\}} \left(\sum_{k:i<k\in E_j} \mathcal{R}_k^{\alpha} Y^{(k)} + \mathcal{R}_i^{\alpha} D \right),
\end{aligned} \tag{35}
$$

where $D, Y^{(1)}, \ldots, Y^{(K)}, (\mathcal{R}, \mathcal{S})$ are independent and $Y^{(1)}, \ldots, Y^{(K)}$ are distributed like Y. From [55, Theorem 1.5], it follows that the fixed-point equation characterizes the distribution of D. (Of course, this is the same argument we used in the proof of Proposition 12.) Moreover, the law of D can be obtained by iteration as follows: Let $D_0 := 1$, and, inductively define D_n by its distribution:

$$
\begin{aligned}
\mathfrak{L}(D_n) = \mathfrak{L}\Bigg(& \sum_{i=1}^{K} \mathbf{1}_{\{J_1=J_2=i\}} \mathcal{R}_i^{\alpha} D_{n-1} \\
& + \sum_{i<j\leq K} \mathbf{1}_{\{J_1=i,J_2=j \text{ or } J_1=j,J_2=i\}} \mathbf{1}_{\{i\notin E_j\}} \left(\sum_{k:q<k\in E_i} \mathcal{R}_k^{\alpha} Y^{(k)} + \sum_{l:q<l\in E_j} \mathcal{R}_l^{\alpha} Y^{(l)}) \right) \\
& + \sum_{i<j\leq K} \mathbf{1}_{\{J_1=i,J_2=j \text{ or } J_1=j,J_2=i\}} \mathbf{1}_{\{i\in E_j\}} \left(\sum_{i<k\in E_j} \mathcal{R}_k^{\alpha} Y^{(k)} + \mathcal{R}_i^{\alpha} D_{n-1} \right) \Bigg),
\end{aligned}
$$

with conditions as in (35). Then, again by [55, Theorem 1.5], we have $D_n \to D$ in distribution. Next, by the last display, we have in a stochastic sense,

$$
D_n \geq \sum_{i=1}^{K} \mathbf{1}_{\{J_1=J_2=i\}} \mathcal{R}_i^{\alpha} D_{n-1} + \sum_{i<j\leq K} \mathbf{1}_{\{J_1=i,J_2=j \text{ or } J_1=j,J_2=i\}} \mathcal{R}_i^{\alpha} \left(\mathbf{1}_{\{i\in E_j\}} D_{n-1} + \mathbf{1}_{\{i\notin E_j\}} Y^{(i)} \right).
$$

By the tail bound on Y we have just proved, for all $\delta > 0$ sufficiently small, we have $\mathbf{E}[Y^{-\delta}] < \infty$. Moreover, by the dominated convergence theorem, $\mathbf{E}[\mathbf{1}_A \mathcal{R}_i^{-s}] \to \mathbf{P}(A)$ for any event A as $s \to 0$. Summing up, it follows from the last display that, for sufficiently small s, there exists $C > 0$ and $0 < p < 1$ (both depending on s) such that

$$
\mathbf{E}[D_n^{-s}] \leq p\mathbf{E}[D_{n-1}^{-s}] + C.
$$

Hence, by induction on n, it is easy to see that $\sup_{n\geq 0} \mathbf{E}[D_n^{-s}] < \infty$. Thus, $\mathbf{E}[D^{-\delta}] < \infty$ for all sufficiently small δ which completes the proof. □

Remark. 1) The trees $\mathfrak{T}_{\mathbf{e}}, \mathfrak{T}_{\mathscr{L}}, \mathfrak{T}_{\mathscr{H}}$ are invariant under re-rooting at a random point, and in these cases, the distributions of $d(\rho, \zeta)$ and $d(\zeta, \zeta')$ are identical.
2) The random variable $e(\xi)$ has the Rayleigh distribution, thus $\mathbf{P}\left(\mathbf{e}(\xi) \leq t\right) = 1 - e^{-t^2/8} = t^2/8 + o(t^2)$ as $t \to 0$.
3) Recall from page 21 in [17], that

$$
\mathscr{H}(\xi) \cdot E^{1/3} \stackrel{d}{=} E
$$

where E has the standard exponential distribution and $\mathscr{H}, \xi, E$ are independent. By analyzing the moment sequence, one observes that $\mathscr{H}(\xi)$ has the Mittag–Leffler distribution with parameter $1/3$. In other words, $\mathscr{H}(\xi)$ is distributed like $S^{-1/3}$ where $S > 0$ is stable with $\log \mathbf{E}\left[\exp(-\lambda S)\right] = -\lambda^{1/3}, \lambda > 0$. From the results in [37, Section 6], in particular Example 6.17, it follows that $\mathscr{H}(\xi)$ has a continuous density f on $[0, \infty)$ with $f(0) = 1/\Gamma(2/3)$. In particular $\mathbf{P}\left(\mathscr{H}(\xi) \leq t\right) = t/\Gamma(2/3) + o(t)$ as $t \to 0$. (We thank Svante Janson for drawing our attention to Mittag–Leffler distributions in the study of $\mathscr{H}(\xi)$.)

Before moving on to the proof of Theorem 5, we need to introduce additional terminology about the expansion in (19). Let $\mathfrak{T}^\vartheta = \mathfrak{T}_{\mathcal{X}^\vartheta}$ for all $\vartheta \in \Theta$, and $\mathfrak{T} = \mathfrak{T}^\emptyset$. The structural tree Γ describes the way the K trees $\mathfrak{T}^1, \ldots, \mathfrak{T}^K$ are arranged in $\mathfrak{T}$. Similarly, we introduce a structural tree for the decomposition at level $n \geq 1$ as follows: Let Γ^n be the plane tree on K^n nodes labelled with elements of Θ_n that describes the adjacencies between the subtrees at level n of the decomposition if is carried out up to level n. We think of Γ^n as rooted at $1 \ldots 1$. Observe that Γ^n is now a random object, since the adjacency relations depend on the random points used to glue the trees. For a node $\vartheta \in \Gamma^n$, let $M'(\vartheta)$ denote the set of its children. By construction, if $\vartheta = \vartheta_1\vartheta_2 \ldots \vartheta_n$, then $M'(\vartheta)$ consists of nodes of the form $\vartheta_1 \ldots \vartheta_\ell \gamma 1 \ldots 1$ with $0 \leq \ell < n$ and $\gamma \in \Gamma_{\vartheta_{\ell+1}}$ (that is γ is a child of $\vartheta_{\ell+1}$ in Γ). By $M(\vartheta) \subseteq M'(\vartheta)$ we denote the subset of children of ϑ where, for any $0 \leq \ell < n$, if $\vartheta_1 \ldots \vartheta_\ell \gamma 1 \ldots 1 \in M'(\vartheta)$ for some γ, we keep only that child with minimal γ. Then, all the trees corresponding to nodes in $M(\vartheta)$ are glued on the tree corresponding to ϑ at points that are distinct with probability one, $M(\vartheta)$ is a maximal set with this property, and $\#M(\vartheta) \leq n$. (Here, we use that the mass measure has no atoms, see the proof of Proposition 7 below.) In other words, $\#M(\vartheta)$ counts the number of distinct exit points of the tree $\mathfrak{T}^\vartheta$, that is the points distinct from the root where geodesics may leave the set $\mathcal{T}^\vartheta$. See Figure 4 for an illustration of the construction and of the sets $M(\vartheta)$.

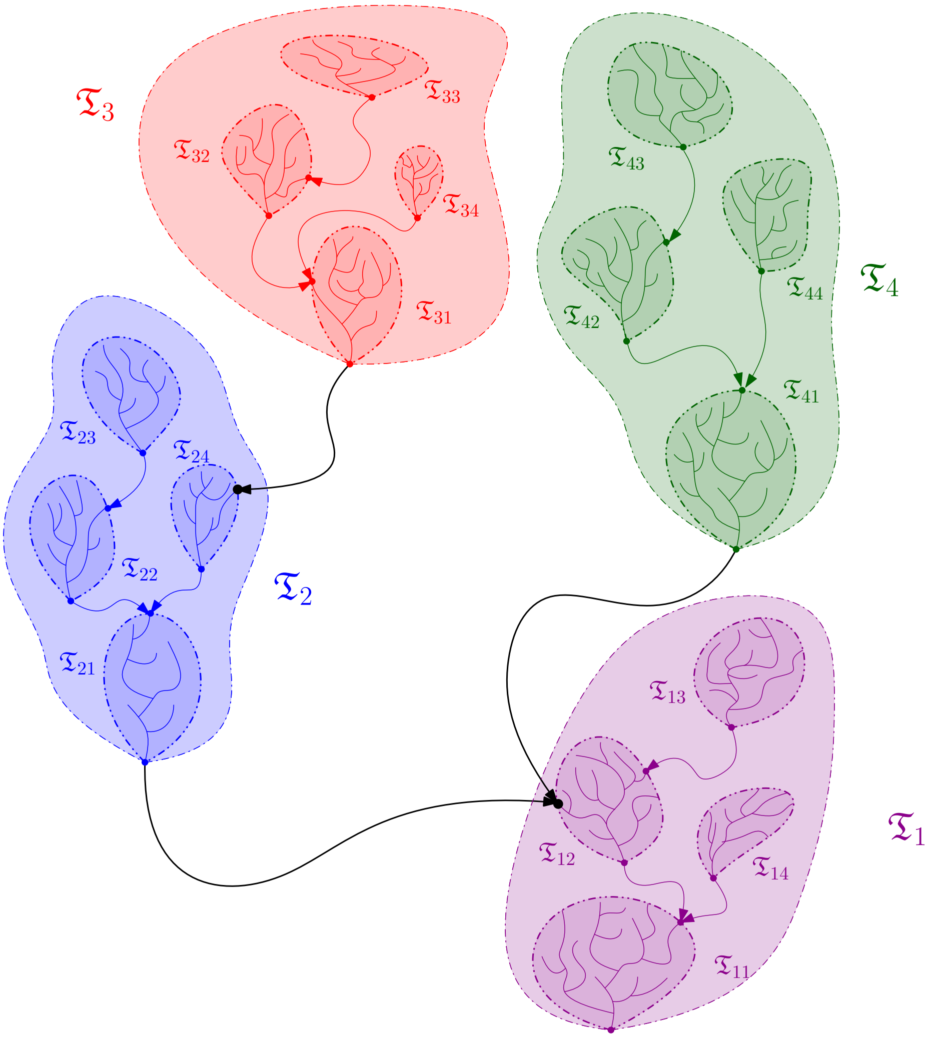

Figure 4: *The construction of $\mathfrak{T}$ from the first two levels of Θ. To illustrate the definition of $M(\vartheta)$, observe that $M'(12) = \{13, 21, 41\}$, $M(12) = \{13, 21\}$ while $M'(24) = M(24) = \{31\}$.*

Proof of Theorem 5. Fix $n \in \mathbb{N}$ and $\gamma < 1/\alpha$. We show that $\dim_{\mathrm{H}}(\mathfrak{T}) \geq \gamma$ on a set of measure one. Since the Hausdorff dimension only depends on the isometry class of $\mathfrak{T}$, this yields the assertion. For $x \in \mathcal{T}$, let $\vartheta(x)$ be the node in Θ_n with $x \in \mathcal{T}^{\vartheta(x)}$ in the decomposition of $\mathcal{T}$ at level $n \geq 1$. We abbreviate[3] $\mathfrak{T}(x) = \mathfrak{T}^{\vartheta(x)}$. Furthermore, let $H_x = d(x, \rho^{\vartheta(x)})$ be the height of x in $\mathcal{T}(x)$ and $E_x = d(x, \mathcal{T} \setminus \mathcal{T}(x))$ the distance to exit $\mathcal{T}(x)$ from x. With $M(x) := M(\vartheta(x))$, for all $x \in \mathcal{T}$, we have

$$E_x = \begin{cases} \min_{\sigma \in M(x)} d(x, \mathcal{T}^\sigma) & \text{if } \vartheta(x) = 1 \dots 1 \\ \min_{\sigma \in M(x)} d(x, \mathcal{T}^\sigma) \wedge H_x & \text{if } \vartheta(x) \neq 1 \dots 1 \,. \end{cases}$$

Recall that $B_r(x) = \{y \in \mathcal{T} : d(x,y) < r\}$ for $x \in \mathcal{T}$, $r > 0$. Obviously, for any $x \in \mathcal{T}$ and $r > 0$, we have $B_r(x) \subseteq \mathcal{T}(x)$ or $E_x \leq r$. Thus, for any (possibly random) probability measure λ on $[0,1]$, and with ζ chosen on $\mathcal{T}$ according to $\bar{\mu} := \lambda \circ \pi_{\mathcal{X}}^{-1}$,

$$\mathbf{P}(\bar{\mu}(B_r(\zeta)) > r^\gamma) \leq \mathbf{P}(\bar{\mu}(\mathcal{T}(\zeta)) > r^\gamma) + \mathbf{P}(E_\zeta \leq r). \tag{36}$$

For any $\vartheta \in \Theta_n$, we denote by $\sigma_0(\vartheta), \dots, \sigma_{n-1}(\vartheta)$ the potential elements of $M(\vartheta)$ where, seen as words on $\{1, \dots, K\}$, σ_ℓ and ϑ have a common prefix of length ℓ. Then, abbreviating $\sigma_\ell := \sigma_\ell(\vartheta(\zeta))$,

$$\mathbf{P}(E_\zeta \leq r) \leq \mathbf{P}\left(H_\zeta \leq r\right) + \mathbf{P}\left(d(\zeta, \mathcal{T}^{\sigma_i}) \leq r, \sigma_i \in M(\zeta) \text{ for some } 0 \leq i \leq n-1\right). \tag{37}$$

Let ζ', ζ'' be independent copies of ζ, independent of all remaining quantities. Then, for $i \in \{0, 1, \dots, n-1\}$, we have

$$\begin{aligned} \mathbf{P}\left(d(\zeta, \mathcal{T}^{\sigma_i}) \leq r, \sigma_i \in M(\zeta)\right) &= \mathbf{P}(\mathcal{V}(\vartheta(\zeta))^\alpha \cdot d(\zeta', \zeta'') \leq r, \sigma_i \in M(\zeta)) \\ &\leq \mathbf{P}(\mathcal{V}(\vartheta(\zeta))^\alpha \cdot d(\zeta', \zeta'') \leq r) \end{aligned}$$

Similarly, H_ζ is distributed like $\mathcal{V}(\vartheta(\zeta))^\alpha \cdot d(\rho, \zeta')$. Let $\eta \in (0,1)$ be a parameter to be chosen later. Applying the union bound on the right-hand side of (37) yields

$$\mathbf{P}(E_\zeta \leq r) \leq n\left\{\mathbf{P}(\mathcal{V}(\vartheta(\zeta)) \leq r^{\eta/\alpha}) + \mathbf{P}(d(\rho,\zeta) \leq r^{1-\eta}) + \mathbf{P}(d(\rho,\zeta) \leq r^{1-\eta})\right\}.$$

We can still choose the measure $\bar{\mu}$ in (36): we set $\bar{\mu} = \mu^\emptyset \circ \pi_{\mathcal{X}}^{-1}$ with $\mu^\emptyset$ as in Proposition 17. Then, $\bar{\mu}(\mathcal{T}(\zeta)) = \mathcal{V}(\vartheta(\zeta))$. Hence, $\mathbf{P}(\bar{\mu}(\mathcal{T}(\zeta)) > r^\gamma) = \mathbf{P}\left(\mathcal{V}(\vartheta(\zeta)) > r^\gamma\right)$. Combining the bounds and using Lemma 18, we see that there exists universal constants $\varepsilon_1 > 0$ and $C > 0$ such that

$$\mathbf{P}(\bar{\mu}(B_r(\zeta)) > r^\gamma) \leq \mathbf{P}\left(\mathcal{V}(\vartheta(\zeta)) > r^\gamma\right) + Cn\left\{\mathbf{P}(\mathcal{V}(\vartheta(\zeta)) \leq r^{\eta/\alpha}) + r^{\varepsilon_1(1-\eta)}\right\}.$$

We now choose the parameters. First, let $\eta \in (\gamma\alpha, 1)$. Then choose $\delta \in (\gamma, \eta/\alpha)$. Finally, let $n = n(r) = \lfloor -\delta \log r \rfloor$. Note that $\mathcal{V}(\vartheta(\zeta))$ is distributed like the product of n independent copies of the random variable $\sum_{i=1}^K \mathbf{1}_{\{J=i\}} \mathcal{R}_i$. The tail bound on $\overline{\mathcal{R}}$ implies that this random variable has exponential moments. Hence, by Cramér's theorem for large deviations, there exist $C_2, \varepsilon_2 > 0$ (depending on the remaining parameters but not on r), such that $\mathbf{P}(\mathcal{V}(\vartheta(\zeta)) > r^\gamma) \leq C_2 r^{\varepsilon_2}$ and $\mathbf{P}(\mathcal{V}(\vartheta(\zeta)) \leq r^{\eta/\alpha}) \leq C_2 r^{\varepsilon_2}$. Summarizing, there exists $C > 0$ (which may depend on all parameters but not on r), such that, for $r < 1$,

$$\mathbf{P}(\bar{\mu}(B_r(\zeta)) > r^\gamma) \leq -C \log r (r^{\varepsilon_2} + r^{\varepsilon_1(1-\eta)}).$$

It follows that for $r_n = 2^{-n}$,

$$\sum_{n \geq 1} \mathbf{P}(\bar{\mu}(B_{r_n}(\zeta)) > r_n^\gamma) < \infty.$$

Hence, by the Borel–Cantelli lemma, almost surely,

$$\limsup_{r \to 0} \bar{\mu}(B_r(\zeta))/r^\gamma \leq 2.$$

Thus, denoting $A = \{x \in \mathcal{T} : \limsup_{r \to 0} \bar{\mu}(B_r(x))/r^\gamma \leq 2\}$, we have $1 = \mathbf{P}\left(\zeta \in A\right) = \mathbf{E}[\bar{\mu}(A)]$ implying $\bar{\mu}(A) = 1$ almost surely. From (14), it follows that, almost surely, $\dim_{\mathrm{H}}(\mathcal{T}) \geq \dim_{\mathrm{H}}(A) \geq \gamma$ which completes the proof. $\square$

[3] When we do not introduce the components of $\mathfrak{T}$ with a given decoration explicitly, we always suppose that they would carry the same decoration; for instance $(\mathcal{T}^*, d^*, \mu^*, \rho^*)$ are the components of $\mathfrak{T}^*$.

5.3 Degrees and properties of the encoding: Proof of Proposition 7

The proofs of the missing parts of Theorem 1 rely on the dynamics governing the number of points in the set $M(\vartheta)$. The following lemma is straightforward from the construction, since the exit points counted by $\#M(\vartheta)$ are chosen on $\mathfrak{T}^\vartheta$ according to the mass measure. Recall that, for $\vartheta \in \Theta$, $\mathcal{L}(\vartheta)$ is the Lebesgue measure of the set Λ_ϑ (see Section 3.4).

Lemma 19. *i) Let $\varepsilon_1, \varepsilon_2, \ldots \in [K]$ and $\vartheta_n = \varepsilon_1 \ldots \varepsilon_n \in \Theta_n$. Then, the sequence $(\#M(\vartheta_n), \mathcal{L}(\vartheta_n)), n \geq 0$ is a homogeneous Markov chain on $\mathbb{N}_0 \times [0,1]$ starting at $(0,1)$, whose evolution can be described as follows: given $(\#M(\vartheta_n), \mathcal{L}(\vartheta_n))$, we have*

$$(\#M(\vartheta_{n+1}), \mathcal{L}(\vartheta_{n+1}))) = (\mathbf{1}_{\{\varepsilon_{n+1}\in\Gamma^o\}} + \mathrm{Bin}\big(\#M(\vartheta_n), \mathcal{S}^{\vartheta_n}_{\varepsilon_{n+1}}\big), \mathcal{L}(\vartheta_n)\cdot \mathcal{S}^{\vartheta_n}_{\varepsilon_{n+1}}).$$

ii) Let ξ be uniformly distributed on $[0,1]$, independent of all remaining quantities and, for $n \geq 0$, let $\tilde\vartheta \in \Theta_n$ be the unique node with $\xi \in \Lambda_{\tilde\vartheta}$. Define $(\tilde M_n, \tilde{\mathcal{L}}_n) := (\#M(\tilde\vartheta_n), \mathcal{L}(\tilde\vartheta_n))$. Then, the sequence $(\tilde M_n, \tilde{\mathcal{L}}_n), n \geq 0$ is a homogeneous Markov chain on $\mathbb{N}_0 \times [0,1]$ starting at $(0,1)$, whose evolution can be described as follows: given $(\tilde M_n, \tilde{\mathcal{L}}_n)$, we have

$$(\tilde M_{n+1}, \tilde{\mathcal{L}}_{n+1}) = (\mathbf{1}_{\{J_{n+1}\in\Gamma^o\}} + \mathrm{Bin}(\tilde M_n, \mathcal{S}^{n+1}_{J_{n+1}}), \tilde{\mathcal{L}}_n \cdot \mathcal{S}^{n+1}_{J_{n+1}}),$$

where $(\mathcal{S}^n)_{n\geq 0}$ is a family of i.i.d. copies of $\mathcal{S}$ and $\mathbf{P}(J_{n+1} = i \,|\, \mathcal{S}^{n+1}) = \mathcal{S}^{n+1}_i$.

Proof of Proposition 7 i), ii) and iii). We start with the proof of *i)*. Denote by Z the zero-set of $\mathcal{X}$, that is the set of points $s \in [0,1]$ for which $\mathcal{X}(s) = 0$. For $n \geq 1$, let $\Lambda^n = \Lambda_{\vartheta_n}$ where $\vartheta_n = 1\ldots 1$ and $\vartheta_n \in \Theta_n$. Then for every $n \geq 0$, $\mathcal{T}^{\vartheta_n}$ is the subtree that contains the root ρ of $\mathfrak{T}$. Furthermore, Λ^n is the union of $\#\Lambda^n$ disjoint intervals. Clearly, $(\Lambda^n)_{n\geq 1}$ is decreasing and we set $\Lambda := \bigcap_{n\geq 1}\Lambda^n$. Since $X(\xi) > 0$ almost surely, it follows that $Z \subseteq \Lambda$. Next, note that $\#\Lambda^n = \#M(\vartheta_n) + 1$. Since $\mathbf{E}[\mathcal{S}_1] < 1$, Lemma 19 *i)* and a routine drift argument (see, e.g., Chapter 8 of [43]) shows that, almost surely, $\#M(\vartheta_n) = 1$ infinitely often, and thus $\#\Lambda^n = 2$ infinitely often. As $\{0,1\} \in Z$, for any $n \geq 1$ with this property, we have

$$\Lambda^n \subseteq [0, \inf\{t > 0 : t \notin \Lambda^n\}] \cup [\sup\{t < 1 : t \notin \Lambda^n\}, 1]\,.$$

Since $\mathcal{L}(\vartheta_n) = \mathrm{Leb}(\Lambda^n) \to 0$ with probability one, it follows that $Z \subseteq \Lambda = \{0,1\}$ almost surely, which implies that the root ρ of $\mathfrak{T}$ is a leaf.

ii) Since $\mu_{\mathcal{X}}$ has full support, this reduces to show that for a random point ζ sampled from $\mu_{\mathcal{X}}$ and any $\varepsilon > 0$, with probability one, there exists some $x \in \mathcal{T}_{\mathcal{X}}$ such that ζ lies in $\mathcal{T}(x)$, the subtree of $\mathcal{T}_{\mathcal{X}}$ rooted at x, that is the closure of the union of all connected components of $\mathcal{T}_{\mathcal{X}} \setminus \{x\}$ not containing the root, and $\mu_{\mathcal{X}}(\mathcal{T}(x)) < \varepsilon$. To prove this, we set $\zeta = \pi_{\mathcal{X}}(\xi)$ with ξ as in Lemma 19 *ii)*, and follow ζ in the refining decomposition of the tree according to Θ_n, as n increases. For $n \geq 1$, let $\tilde\vartheta_n \in \Theta_n$ be as in in Lemma 19 *ii)*. In particular, $\zeta \in \mathcal{T}_{\tilde\vartheta_n} := \mathcal{T}_{\mathcal{X}^{\tilde\vartheta_n}}$. Then, ζ lies in the subtree of $\mathcal{T}$ rooted at $\rho_{\tilde\vartheta_n}$, and thus it suffices to show that for any $\varepsilon > 0$,

$$N_\varepsilon = \inf\{n \geq 0 : \mu_{\mathcal{X}}(\mathcal{T}_{\rho_{\tilde\vartheta_n}}) < \varepsilon\} < \infty\,.$$

Observe that, if $\tilde{\mathcal{L}}_n = \mu_{\mathcal{X}}(\mathcal{T}_{\tilde\vartheta_n}) < \varepsilon$ and $\tilde M_n = 0$, then $N_\varepsilon \leq n$. As in the proof of *i)* above, since $\mathbf{E}[\mathcal{S}^n_{J_n}] < 1$, a classical drift argument shows that $(\tilde M_n)_{n\geq 0}$ is positive recurrent, and in particular, $\tilde M_n = 0$ infinitely often. Then, for some subsequence $(n_i)_{i\geq 1}$ with $n_i \geq i$, we have $\tilde M_{n_i} = 0$. But clearly, by Lemma 19, $\tilde{\mathcal{L}}_n \to 0$ almost surely, so that there is an i_0 for which $\tilde{\mathcal{L}}_{n_i} < \varepsilon$ for all $i \geq i_0$. One then has $N_\varepsilon \leq n_{i_0} < \infty$, which completes the proof.

Finally, we consider *iii)*. Since $d(\rho, \zeta) > 0$ almost surely, no mass can add up at exit points in the construction of the excursion X in (10). Hence,

$$\mathbf{E}\Bigg[\sup_{t\in[0,1]} \mu(\{\pi_X(t)\})\Bigg] \leq \mathbf{E}\left[\max(\mathcal{S}_1, \ldots, \mathcal{S}_K)\right] \cdot \mathbf{E}\Bigg[\sup_{t\in[0,1]} \mu(\{\pi_X(t)\})\Bigg].$$

It follows that the left-hand side is zero which concludes the proof. ☐

The claims in Theorem 1 *iii)* and *iv)* are simple consequences of Proposition 7 *i)–ii)* and its proof above:

Proof of Theorem 1 iii) and iv). We start with *iii)*: suppose that, with positive probability, there exists $s \in (0,1)$ such that $\mathcal{X}(s) = 0$. Then, again with positive probability, the root $\rho_{\mathcal{X}}$ has degree at least two, which contradicts Proposition 7 *i)*.

The argument we have used above for the proof of Proposition 7 *ii)* also implies *iv)*, that is $\mathcal{X}$ is nowhere monotonic: indeed, let ξ be uniform in $[0,1]$ and suppose that ζ is the corresponding point in $\mathcal{T}_{\mathcal{X}}$. Let $I \subseteq \mathbb{N}$ be the (a.s. infinite) set of indices n for which $\tilde{M}_n = 0$ in the proof of Proposition 7 *ii)* above. Then, for $n \in I$, the set $\Lambda_{\vartheta_n} \subseteq [0,1]$ consists of a single half-open interval, and $\xi \in \Lambda_{\vartheta_n}$. In particular, for every $n \in I$ large enough, Λ_{ϑ_n} is contained in an interval of length 2ε containing at ξ. But on the restriction of $\mathcal{X}$ to Λ_{ϑ_n} is non-monotonic since $\xi \in \Lambda_{\vartheta_n}$ and $\mathcal{X}(\inf \Lambda_{\vartheta_n}) = \mathcal{X}(\sup \Lambda_{\vartheta_n}) < \mathcal{X}(\xi)$ almost surely by Theorem 1 *iii)*. $\square$

Finally, we prove the characterization of the set of degrees of the points of $\mathcal{T}_X$. It relies crucially on the statements in Proposition 7 *i)* and *ii)* that we have just proved.

Proof of Proposition 7 iv). First note that $2 \in \mathscr{D}(\mathcal{T})$ almost surely: indeed, with probability one, μ has no atoms, the set of branch points of $\mathcal{T}$ is countable, and $\mathcal{T}$ is not reduced to a point. Furthermore, since $\mathcal{X}$ is no-where monotonic, there exist local minima, and hence branch points: it follows that the maximum degree of $\mathcal{T}$ is at least 3 (and, a priori, possibly infinite).

For the other degrees, let $(\zeta_i)_{i\ge 1}$ be a family of i.i.d. random points sampled from the mass measure μ, and set $\zeta_0 = \rho$ the root. Since μ of $\mathfrak{T}$ has full support on $\mathcal{T}$, if there exists $x \in \mathcal{T}$ with degree $k+1$, then all the connected components of $\mathcal{T} \setminus \{x\}$ have positive mass. In particular, with positive probability, any two of the segments $[\![\zeta_i, \zeta_j]\!]$, $0 \le i < j \le k$, intersect at the same point x.

We consider the construction of $\mathfrak{T}_{\mathcal{X}}$ from the excursion $\mathcal{X}$ of Section 3.4. Let $(\xi_i)_{i\ge1}$ be a family of i.i.d. uniform points in $[0,1]$ and $\xi_0 = 0$; then let $(\zeta_i)_{i\ge1}$ be the corresponding points in $\mathcal{T}_{\mathcal{X}}$. Suppose first that $k \ge 3$, and let $\mathcal{I}_k$ be the collection of points in $\mathcal{T}_{\mathcal{X}}$ that are contained in some intersection $[\![\zeta_i, \zeta_j]\!] \cap [\![\zeta_{i'}, \zeta_{j'}]\!]$, for i, i, j, j' distinct elements of $\{0, \dots, k\}$. Let G_k be the event that $\mathcal{I}_k$ is reduced to a single point of degree $k+1$ in $\mathcal{T}_{\mathcal{X}}$; we are interested in knowing whether G_k happens with positive probability or not. We rely on the recursive decomposition. Recall the collection of sets $(\Lambda_\sigma)_{\sigma\in\Theta}$ defined in Section 3.4. The idea of the proof is the following: initially, we set $G_k^0 = G_k$; then, as long as we cannot decide whether G_k^n occurs or not solely from the random variables $(\Lambda_\sigma)_{\sigma\in\Theta_n}$ within the first n levels of the decomposition, we expose more levels and replace the event G_k^n by another event G_k^{n+1} in such a way that G_k^{n+1} and G_k^n are almost surely equal; we then show that we can indeed determine the outcome of G_k after finitely many iterations.

Initially, set $\sigma_0 = \emptyset$. At some stage $n \ge 0$, we have ξ_i^n, $1 \le i \le k$, points that are uniform in Λ_{σ_n}, and $\xi_0^n = \inf \Lambda_{\sigma_n}$. Let J_i^n be the index j for which $\xi_i^n \in \Lambda_{\sigma_n j}$. Let $a \in \{1, 2, \dots, K\}$ be the node of Γ with maximal label such that $\{J_i^n : 1 \le i \le n\}$ is fully contained in the subtree of Γ rooted at a (note that we exclude J_0^n, for otherwise we would have $a = 1$). We distinguish a few cases depending on the number of $1 \le i \le k$ for which $J_i^n = a$:

i) if $\#\{1 \le i \le k : J_i^n = a\} = k$: set $\sigma_{n+1} = \sigma_n a$, $\xi_i^{n+1} = \xi_i^n$ for $1 \le i \le k$ and $\xi_0^{n+1} = \inf \Lambda_{\sigma_{n+1}}$. We let $\mathcal{I}_k^{n+1}$ and G_k^{n+1} be as before, but with the points ξ_i^{n+1}. Then, $G_k^{n+1} = G_k^n$ almost surely. Furthermore, conditional on that case, the points $(\xi_i^{n+1})_{1\le i\le k}$ are uniform on $\Lambda_{\sigma_{n+1}}$ and we recurse in $\Lambda_{\sigma_{n+1}}$;

ii) if $\#\{1 \le i \le k : J_i^n = a\} = k-1$: let ℓ be the unique index for which $J_\ell^n \ne a$. set $\sigma_{n+1} = \sigma_n a$, $\xi_i^{n+1} = \xi_i^n$ for $i \notin \{0, \ell\}$ and ξ_ℓ^{n+1} is the uniformly random point in $\Lambda_{\sigma_{n+1}}$ where the excursions corresponding to the children of $\sigma_n a$ are inserted, while $\xi_0^{n+1} = \inf \Lambda_{\sigma_{n+1}}$. Again, we let $\mathcal{I}_k^{n+1}$ and G_k^{n+1} be as before, but with the points ξ_i^{n+1}; we then have $G_k^{n+1} = G_k^n$ almost surely. Also, conditional on that event, the points $(\xi_i^{n+1})_{1\le i\le n}$ are uniform on $\Lambda_{\sigma_{n+1}}$ and we recurse in $\Lambda_{\sigma_{n+1}}$;

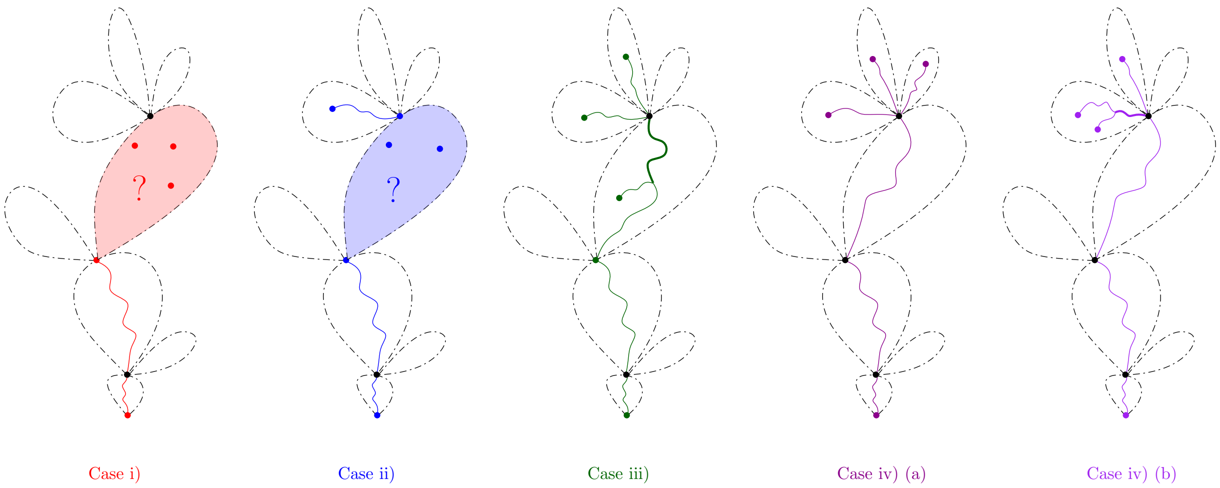

Figure 5: *An illustration of the different cases of the proof of Proposition 7 in the situation of $k = 4$ points. We can determine whether $\mathcal{I}_4$ (in fat) is reduced to a single point or not, and whether that point (if it exists) has degree 4, only from $J_1^n, \ldots, J_4^n$.*

iii) if $\#\{1 \leq i \leq k : J_i^n = a\} \in \{1, \ldots, k-2\}$: assume without loss of generality that $J_1^n = a$, and $J_2^n, J_3^n \neq a$. By Proposition 7 *ii)* the mass measure only charges leaves, therefore the segments $[\![\zeta_1^n, \zeta_2^n]\!]$ and $[\![\zeta_0^n, \zeta_3^n]\!]$ intersect on a portion of positive length and G_k^n does not occur. Set $N = n$ and stop.

iv) if $\#\{1 \leq i \leq k : J_i^n = a\} = 0$

(a) if the J_i^n are all distinct, then $\mathcal{I}_k^n$ is reduced to a single point x. The point x has degree at least $k+1$ in $\mathcal{T}_{\mathcal{X}}$, and this degree is an element of $\mathscr{D}(\Gamma)$, each possibility happening with positive probability, independently of n. We set $N = n$ and stop.

(b) otherwise, there exist $1 \leq i < j \leq k$ such that $J_i^n = J_j^n = b$. Suppose without loss of generality that $i = 1$ and $j = 2$. Since the point corresponding to $\inf \Lambda_{\sigma_n b}$ in $\mathcal{T}_{\mathcal{X}}$ has degree 1 almost surely by Proposition 7 *i)*, and since $k \geq 3$, the two segments $[\![\zeta_1^n, \zeta_0^n]\!]$ and $[\![\zeta_2^n, \zeta_3^n]\!]$ intersect on a portion of positive length, and thus G_k^n does not occur. We set $N = n$ and stop.

Note that whenever we go on, $G_k^{n+1} = G_k^n$ with probability one. Furthermore, since $k \geq 3$, case iii) always occurs with positive probability, and at each stage, we stop with the same positive probability, and thus $N < \infty$ almost surely. It follows that $G_k^N = G_k^0$. It follows readily that G_k occurs with positive probability if and only if $k+1 \in \mathscr{D}(\Gamma)$, and in particular $\sup \mathscr{D}(\mathcal{T}) < \infty$ almost surely.

Now, for the case $k = 2$, note that the procedure above does not always finish in finite time. If $\max \mathscr{D}(\Gamma) \geq 3$, then we do have $N < \infty$ a.s., since case iv) (a) does occur with positive probability (case iii) never happens); in this case, G_2^N occurs with positive probability if and only if $3 \in \mathscr{D}(\Gamma)$. On the other hand, if $\mathscr{D}(\Gamma) = \{1, 2\}$, then the procedure goes on forever. However, the proof for the cases $k > 2$ imply that the maximum degree is at most 3; since we have established that the maximum degree is at least 3, there must be points of degree 3.

Finally, to show that with probability one, for every k for which G_k happens with positive probability, there indeed exists a point $x \in \mathcal{T}_{\mathcal{X}}$ with degree k, it suffices to observe that for any $n \geq 0$, there is one such point with the same positive probability in every single one of the trees reduced to the sets Λ_σ, for $\sigma \in \Theta_n$, and that the events are independent. This concludes the proof. $\square$

5.4 Optimal Hölder exponents

We start with the proof of Corollary 6. At that point, it merely consists in putting together the information we have gathered in the previous sections.

Proof of Corollary 6. *(a)* Note that under the conditions of Theorem 4, $\mathbf{E}[\overline{\mathcal{R}}^{-\delta}] < \infty$ for any $\delta \in (0, \gamma)$. Thus the conclusion of Theorem 5 holds. Now, on the one hand, we have $\underline{\dim}_{\mathrm{M}}(\mathcal{T}) \geq \dim_{\mathrm{H}}(\mathcal{T}_X) \geq \alpha^{-1}$, because of Theorem 5 *ii)*; on the other hand, $\overline{\dim}_{\mathrm{M}}(\mathcal{T}) \leq (KC - \gamma)/\alpha = 1/\alpha$ when $C = \gamma$ and $\gamma(K-1) = 1$.

(b) By Thereom 1, X is no-where constant, and it follows that (15) holds, so that $\alpha = \alpha_f$. By definition of α_f, for any $\gamma < \alpha_f = \alpha$, almost surely, there exists a process $\tilde{X}$ equivalent to X with γ-Hölder continuous paths.

(c) Suppose that there exists a process X' with γ-Hölder continuous paths for $\gamma > \alpha$, and that the metric spaces $(\mathcal{T}_X, d_X)$ and $(\mathcal{T}_{X'}, d_{X'})$ are isometric. Then, we would have $\dim_{\mathrm{H}}(\mathcal{T}_{X'}) = \dim_{\mathrm{H}}(\mathcal{T}_X) \leq 1/\gamma < 1/\alpha$; this would contradict *(a)*, and thus there exists no such process. $\square$

To conclude the section, we return to the time change in Proposition 17 and show how it can be used to construct a re-parametrization of X satisfying (a variant of) Equation (10) (on an almost sure level) with matching scalings in space and in time. Since we do not draw further implications from these consideration, we remain brief and omit the proofs. First, it is not hard to see that, almost surely, τ^ϑ is nowhere constant, thus we can define its inverse $\sigma^\vartheta = (\tau^\vartheta)^{-1}$ in a straightforward way. For $i \in [K]$, define $\tilde{\xi}_i^\vartheta = \tau^\vartheta(\xi_i^\vartheta)$. It is important to note that, $\tilde{\Xi}^\vartheta$ depends only on $\mathcal{R}^\vartheta$ and Ξ^ϑ but *not* on $\mathcal{S}^\vartheta$. By construction, we have

$$\sigma^\vartheta = \Phi_1(\sigma^{\vartheta 1}, \ldots, \sigma^{\vartheta K}, \mathcal{S}^\vartheta, \mathcal{R}^\vartheta, \tilde{\Xi}^\vartheta). \tag{38}$$

Setting $\tilde{X}^\vartheta = X^\vartheta \circ \sigma^\vartheta$, using (10), (32) and (38), on checks that

$$\tilde{X}^\vartheta = \Phi(\tilde{X}^{\vartheta 1}, \ldots, \tilde{X}^{\vartheta K}), \mathcal{R}^\vartheta, \mathcal{R}^\vartheta, \tilde{\Xi}^\vartheta).$$

Equivalently,

$$\tilde{X}^\vartheta(\cdot) = \sum_{i=1}^{K} \mathbf{1}_{\tilde{\Lambda}_i^\vartheta}(\,\cdot\,) \left[(\mathcal{R}_i^\vartheta)^\alpha \tilde{X}^{\vartheta i}(\tilde{\varphi}_i^{\vartheta i}(\,\cdot\,)) + \sum_{j \in E_i} (\mathcal{R}_j^\vartheta)^\alpha \tilde{X}^{\vartheta j}(\tilde{\xi}_j^\vartheta) \right], \tag{39}$$

where the function $\tilde{\varphi}_i^{\vartheta i}(\,\cdot\,) = \tau^{\vartheta i}(\varphi_i^\vartheta(\sigma^\vartheta(\,\cdot\,)))$ is piecewise linear on the set $\tilde{\Lambda}_i^\vartheta$ with constant slope (which is defined analogously to Λ_i but with $\mathcal{R}^\vartheta$ and $\tilde{\Xi}^\vartheta$). Furthermore, as desired, we have $\mathrm{Leb}(\tilde{\Lambda}_i^\vartheta) = \mathcal{R}_i^\vartheta$. Note that, on the one hand, $(\mathcal{R}^\vartheta, \tilde{\Xi}^\vartheta), \tilde{X}^{\vartheta 1}, \ldots, \tilde{X}^{\vartheta K}$ are independent and each $\tilde{X}^\vartheta$ has the same distribution. However, on the other hand, in general, $\tilde{\Xi}^\vartheta$ and $\mathcal{R}^\vartheta$ are not independent. Hence, the equation in (39) cannot be expressed as a fixed-point equation of the kind (10). In applications, when the conditions of Theorem 5 are satisfied, $\tilde{X}$ is a promising candidate for the process in Corollary 6 having optimal Hölder exponents. However, in the case of the process $\mathscr{H}$ in (13), we did not overcome the technical difficulties to show the desired properties of its re-parametrizations due to the lack of independence in the identity (39).

6 Applications

In the first three sections, we discuss the corollaries stated in Section 3.3 concerning the processes $\mathbf{e}$, $\mathscr{Z}$, and $\mathscr{H}$. Then, in the fourth section, we study a new application.

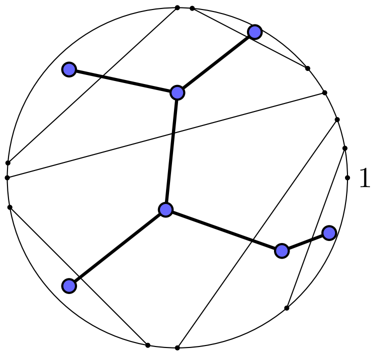

Figure 6: A lamination and the corresponding rooted dual tree. Distances in the tree correspond to the number of chords separating the fragments in the lamination.

6.1 The Brownian continuum random tree

The Brownian continuum random tree $\mathcal{T}_{\mathbf{e}}$, encoded by a Brownian excursion, is a fundamental tree arising as scaling limit for various classes of random trees. We quote the classical case of uniform random labelled trees [5, Theorem 2], but also binary unordered unlabeled trees (Otter trees) [42], random trees with a prescribed degree sequence [16], general unordered unlabeled trees (Pólya trees) [46], unlabeled unrooted trees [53], and random graphs from subcritical classes [47] to name a few examples. (See also [39].)

Theorem 20 (Aldous [5], see also [39])**.** *Let T_n be the family tree of a critical branching process with offspring mean one and finite offspring variance σ^2 conditioned on having n vertices. Let d_n denote the graph distance on T_n and μ_n the uniform probability measure on the leaves. Then, as $n \to \infty$, in distribution with respect to the Gromov–Hausdorff–Prokhorov distance,*

$$(T_n, \tfrac{\sigma}{2} \cdot n^{-1/2} d_n, \mu_n, \rho_n) \to (\mathcal{T}_{\mathbf{e}}, d_{\mathbf{e}}, \mu_{\mathbf{e}}, \rho_{\mathbf{e}}).$$

Corollary 8 *(a)* and Corollary 9 *(a)* follow immediately from (11) and Theorem 1. Similarly, Corollary 10 *(a)* follows from Theorems 4 and 5 noting that $\overline{\mathcal{R}}$ has the Beta$(1/2, 1)$ distribution with density $\frac{1}{2} t^{-1/2}$.

6.2 Random self-similar recursive triangulations of the disk

The processes $\mathscr{Z}$ and $\mathscr{H}$ arise in the problem of random recursive decompositions of the disk by non-crossing chords [17, 23]. They encode the trees that are the planar dual of the limit triangulation in the same sense that the Brownian continuum random tree is the dual of the limit uniform triangulation of the disk studied by Aldous [7, 8]. We now proceed to the precise definitions.

The unit disk $\mathscr{D} := \{z \in \mathbb{C} : |z| \leq 1\}$ is decomposed at discrete time steps as follows: At time $n = 1$, a chord is inserted connecting two uniformly chosen points on the boundary $\partial\mathscr{D}$. Then, given the configuration at time n, at time $n+1$: pick two independent points on the circle $\partial\mathscr{D}$ uniformly at random; add the chord connecting them if it does not intersect any previously inserted chord, otherwise reject the points and continue. This procedure yields an increasing sequence of non-intersecting chords $(L_n)_{n\geq 1}$, also called a lamination. For each $n \geq 1$, $\mathscr{D} \setminus L_n$ consists of a finite number of connected components, and by T_n, we denote the discrete tree which is planar dual to the decomposition (nodes correspond to connected components, and two nodes are adjacent if the corresponding connected components share a chord). The tree T_n is rooted at the node corresponding to a fragment containing a fixed-point on the circle, say 1. (See Figure 6). It has been proved in [17] that T_n suitably rescaled converges almost surely towards a limit tree encoded by a certain random process which satisfies a fixed-point equation of type (10). More precisely, for $\beta := (\sqrt{17} - 3)/2$, with respect to the Gromov–Hausdorff distance and as

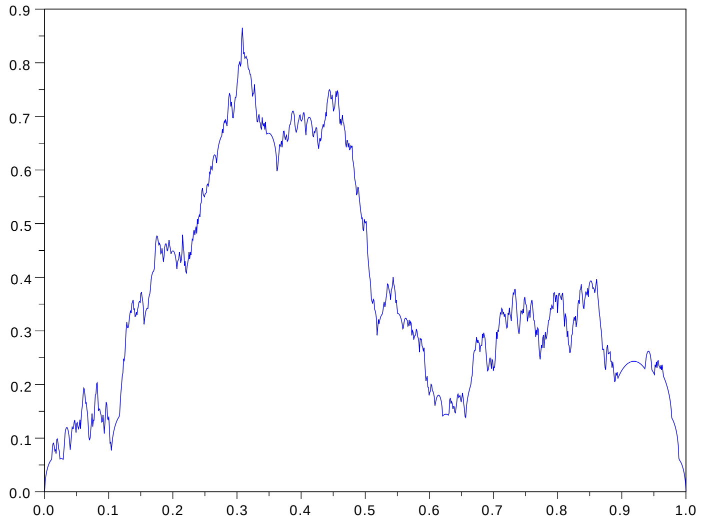

Figure 7: *The process $\mathscr{Z}$ encoding the dual tree of the self-similar recursive triangulation of the disk. It is the unique excursion that satisfies the fixed-point equation* (12) *up to a multiplicative constant.*

$n \to \infty$, we have

$$(T_n, n^{-\beta/2} d) \to (\mathcal{T}_{\mathscr{Z}}, d_{\mathscr{Z}}), \tag{40}$$

for the unique random excursion $\mathscr{Z}$ satisfying (12) with $\mathbf{E}[\mathscr{Z}(\xi)] = \kappa > 0$, where κ denotes a scaling constant whose value is irrelevant in the present context. (It is given in Theorem 3 in [17].)

Corollary 8 *(b)*, Corollary 9 *(b)* and Corollary 10 *(b)* follow from (12), Theorems 1, 4 and 5 since $\overline{\mathcal{R}}$ has the uniform distribution on the unit interval.

For any $\gamma < \beta$, by Corollary 6, there exists a process equivalent to $\mathscr{Z}$ with γ-Hölder continuous paths. Moreover, for $\gamma > \beta$, no equivalent process can be γ-Hölder continuous with positive probability. Indeed, by Theorem 1.1 in [23], the process $\mathscr{Z}$ itself has γ-Hölder continuous paths for any $\gamma < \beta$ and is therefore optimal with respect to regularity. $\mathscr{Z}$ is a good encoding of the real tree $\mathcal{T}_{\mathscr{Z}}$ since its fractal dimension corresponds precisely to what should be expected from the regularity of $\mathscr{Z}$. (The fact that the rescaling of T_n is $n^{-\beta/2}$ in (40) rather than $n^{-\beta}$ is reminiscent of the number of chords in L_n, which is only of order $\sqrt{n}$, so T_n has only order $\sqrt{n}$ nodes, see [17, 23].)

6.3 Homogeneous recursive triangulation of the disk

We have a completely different situation if we consider a partition of the disk $\mathscr{D} = \{z \in \mathbb{C} : |z| \leq 1\}$ using random chords, but this time, the chords are inserted using a different strategy that is homogeneous: in each step, given the current configuration, one connected component is chosen *uniformly at random* and split by the insertion of a chord linking two uniformly random points on the boundary conditioned on splitting the chosen component. Now, there is no rejection, and at time n we have a collection of chords L_n^h consisting of n elements. As before, we can define a tree that is dual to the lamination, and we denote it by T_n^h (the discrete tree T_n^h has $n+1$ nodes). It has been proved in [17] that a suitably rescaled version of T_n^h converges: in distribution with respect to the Gromov–Hausdorff distance and as $n \to \infty$, we have

$$(T_n^h, n^{-1/3} d) \to (\mathcal{T}_{\mathscr{H}}, d_{\mathscr{H}}), \tag{41}$$

where $\mathscr{H}$ is the unique random excursion satisfying (13) and $\mathbf{E}[\mathscr{H}(\xi)] = 1/\Gamma(4/3)$. [No characterization of $\mathscr{H}$ was given in [17] but is stated in Corollary 9 *(c)*.] The rescaling $n^{-1/3}$ in (41) suggests that

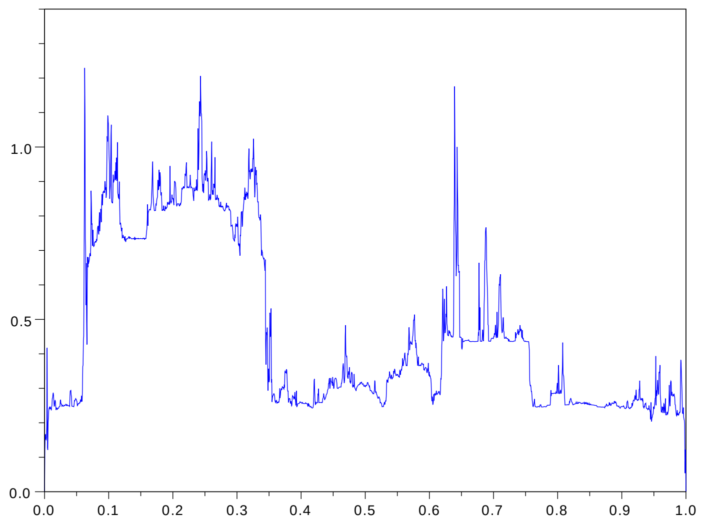

Figure 8: *The process $\mathscr{H}$ encoding the dual tree of the homogeneous recursive triangulation of the disk. It is the unique excursion that satisfies the fixed-point equation (13) up to a multiplicative constant.*

the limit tree $\mathcal{T}_{\mathscr{H}}$ should have fractal dimension 3. However, a first natural grasp that one has on the tree $\mathcal{T}_{\mathscr{H}}$ is the encoding excursion $\mathscr{H}$, but a quick look at Figure 8 suggests that some trouble is around the corner since $\mathscr{H}$ does not look Hölder with exponent $1/3 - \varepsilon$ for $\varepsilon > 0$ arbitrary.

It is precisely in this kind of situation that our general framework is most useful, since it permits to verify that $\mathcal{T}_{\mathscr{H}}$ indeed has fractal dimension 3, and more precisely that $\dim_{\mathrm{M}}(\mathcal{T}_{\mathscr{H}}) = \dim_{\mathrm{H}}(\mathcal{T}_{\mathscr{H}}) = 3$ with probability one. This is reminiscent of the fact that, for any $\gamma < 1/3$, there exists excursions equivalent to $\mathscr{H}$ that have γ-Hölder continuous paths. As expected, unlike the process $\mathscr{Z}$, $\mathscr{H}$ is suboptimal with respect to path regularity: the following proposition is given for the sake of completeness, and its proof can be found in the Appendix.

Proposition 21. *Let $\varrho = 1 - \frac{2}{3}\sqrt{2} = 0.057\ldots$ Then, almost surely,*

$$\sup\{\gamma > 0 : \mathscr{H} \text{ is } \gamma\text{-Hölder continuous}\} = \varrho.$$

Corollaries 8 (c), 9 (c) and 10 (c) follow as in the recursive case discussed in the previous section.

6.4 Recursive k-angulations

In this section we consider a generalization of the lamination process described in Sections 6.2 and 6.3, where, for some fixed $k \geq 2$, in each step, one adds the k-gon connecting k points sampled on the circle (for a precise definition, see below). Certain quantities in this model were studied by Curien and Peres [24]. Again, we are interested in non-intersecting structures and investigate both the recursive and the homogeneous model. Of course, for $k = 2$, we recover the processes studied in Sections 6.2 and 6.3. The techniques in [17] and [23] were sophisticated enough to yield detailed information on the height processes of the corresponding dual trees and their limits. For example, in [17], we gave explicit expressions for the leading constants and rates of convergence for the mean functions. Furthermore, the limit mean function had already been obtained in [23] (up to a multiplicative constant). Most of these precise results do not play a significant role in proving convergence of the dual trees or determining the fractal dimensions of the limiting objects. (The leading constants in Propositions 22 and 23 below could be given by lengthy implicit formulas and are of no particular relevance.)

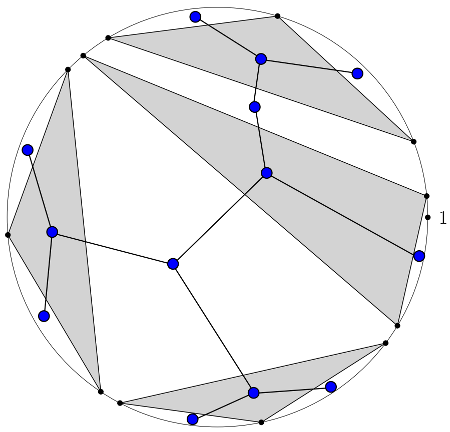

Figure 9: *An example of 3-angulation, together with its dual tree. The tree is rooted at the node contaning the point $1 \in \mathbb{C}$. The shaded portions correspond to the triangles inserted, while the white portions are essential fragments, i.e. the regions of the disk with a positive Lebesgue measure on the circle.*

The recursive k-angulation. In the recursive framework, in each step, we choose k points uniformly at random on the circle and insert the corresponding k-gon if none of its edges intersects any previously inserted. The dual tree $\mathcal{T}_n$ is defined analogously to the case $k = 2$ upon identifying fragments in the decomposition with nodes in $\mathcal{T}_n$. The *mass* of a fragment in the decomposition of the disk is the one-dimensional Lebesgue measure of its intersection with the circle. Fragments with positive mass will subsequently be called *essential* (of course, all fragments are essential for $k = 2$.) See Figure 9 for an illustration. Keeping the notation introduced in [17], by $\mathcal{C}_n(s)$, we denote the depth of the node associated to the fragment covering s in the tree $\mathcal{T}_n$. (We do not indicate k in the notation for the height functions.) Denote the first inserted k points in increasing order by $U_1, \ldots, U_k$ and define $\Delta_1 = 1 - U_k + U_1$, $\Delta_i = U_i - U_{i-1}, 2 \le i \le k$, as well as $\xi^* = U_1/\Delta_1$. Furthermore, let $I_n = (I_n^{(1)}, \ldots, I_n^{(k)})$, where $I_n^{(i)}$ is the number of attempted insertions up to time n in the fragment containing Δ_i. Given $(U_1, \ldots, U_k)$, for any $1 \le i \le k$, the random variable $I_n^{(i)}$ has the Binomial distribution with parameters $n-1, \Delta_i^k$. In particular, we have, almost surely,

$$\frac{I_n}{n} \to (\Delta_1^k, \ldots, \Delta_k^k). \tag{42}$$

Proposition 22. *Let $k \ge 2$ and N_n be the number of inserted k-gons at time n. Then, as $n \to \infty$, we have $n^{-1/k} N_n \to c_k$ in probability and with respect to all moments, where $c_k > 0$ is a constant.*

Proof. Let $\tau_1, \tau_2, \ldots$ be the times of homogeneous Poisson point process on the positive real axis. We consider the continuous-time analogue of $N_n, n \in \mathbb{N}$ denoted by $\mathcal{N}_t, t > 0$ where, for all times $\tau_i, i \ge 1$, a set of k independent points are drawn at random on the circle and the corresponding k-gon inserted if the decomposition remains non-crossing. In other words, $\mathcal{N}_t = N_i$ for $t \in [\tau_i, \tau_{i+1})$ where $\tau_0 := 0$. It is easy to see and explained in detail in [23] for $k = 2$, that this process can alternatively be obtained without the necessity of rejecting any k-gons as follows: starting with the disk at time $t = 0$, add a k-gon chosen uniformly at random after an exponentially distributed time with mean one. Then, independently on the k essential sub-fragments, run the same process with times slowed down by a factor x^k where x denotes the mass of the fragment. The masses of essential fragments at time $t > 0$ in this process constitute a conservative fragmentation process with index of self-similarity k and reproduction law Dirichlet$(1, \ldots, 1)$. Hence, by Theorem 1 in [13], we deduce $t^{-1/k}\mathcal{N}_t \to c_k$ with c_k as in the proposition in probability and in L_2. In particular, $\tau_n^{-1/k}\mathcal{N}_{\tau_n} \to c_k$ in probability as $n \to \infty$. In order to transfer moment convergence, note that, for any $\varepsilon > 0$, by monotonicity and since $N_n \le n$ almost surely,

$$\tau_n^{-2/k}\mathcal{N}_{\tau_n}^2 \le ((1-\varepsilon)n)^{-2/k}\mathcal{N}_{(1+\varepsilon)n}^2 + \tau_n^{2-2/k}\mathbf{1}_{\{|\tau_n - n| \notin (-\varepsilon n, \varepsilon n)\}}.$$

By the L_2 convergence for the continuous-time process and the concentration of τ_n having the Gamma(n) distribution, the right hand side is uniformly integrable. Hence, $\tau_n^{-1/k}\mathcal{N}_{\tau_n} \to c_k$ in L_2. Since $\tau_n/n \to 1$ almost surely and with convergence of all moments and $\tau_n, \mathcal{N}_{\tau_n}$ are independent, we obtain the convergence in probability and in L_2. Finally, let $\tilde{N}_n = n^{-1/k}N_n$. Then,

$$\tilde{N}_n \stackrel{d}{=} 1 + \sum_{i=1}^{k} \left(\frac{I_n^{(i)}}{n} \right)^{1/k} \tilde{N}^{(i)}_{I_n^{(i)}},$$

where $(\tilde{N}_n^{(1)})_{n\geq 1}, \ldots, (\tilde{N}_n^{(k)})_{n\geq 1}$ are independent copies of $(N_n)_{n\geq 1}$, independent of I_n. Using (42), it is easy to prove that $\tilde{N}_n$ is bounded in $L_m, m \geq 1$ by induction over m since we have already shown it for $m = 1, 2$. □

By construction, the random process $(C_n(s))_{s\in[0,1]}$ satisfies the following recurrence in distribution on the space of cadlag functions $\mathcal{D}$:

$$\begin{aligned} C_n(\,\cdot\,) \stackrel{d}{=} & \mathbf{1}_{[0,U_1)}(\,\cdot\,) C^{(1)}_{I_n^{(1)}}\left(\frac{\cdot}{\Delta_1}\right) + \mathbf{1}_{[U_k,1]}(\,\cdot\,) C^{(1)}_{I_n^{(1)}}\left(\frac{\cdot - U_k}{\Delta_1}\right) \\ & + \sum_{i=2}^{k-1} \mathbf{1}_{[U_{i-1},U_i]}(\,\cdot\,) \left(C^{(i)}_{I_n^{(i)}}\left(\frac{\cdot - U_{i-1}}{\Delta_i}\right) + C^{(1)}_{I_n^{(1)}}(\xi^*) \right), \end{aligned} \tag{43}$$

where $(C_i^{(1)}(\cdot))_{i\geq 0}, \ldots, (C_i^{(k)}(\cdot))_{i\geq 0}$ are independent copies of $(C_i(\cdot))_{i\geq 0}$ independent of $(U_1, \ldots, U_k, I_n)$. The first step of our analysis is to investigate C_n at a uniformly chosen point ξ. In [17], an explicit expression for the mean was obtained by solving the underlying recursion. Here, we proceed as in [23] relying on results from fragmentation theory. Subsequently, let $\alpha_k \in (0, 1)$ be the unique solution to

$$\begin{aligned} g_k(x) &:= \mathbf{E}\left[\Delta_1^x\right] + (k-1)\mathbf{E}\left[\mathbf{1}_{[U_1,U_2]}(\xi)\Delta_2^x\right] \\ &= \frac{k!\Gamma(x+2)}{\Gamma(k+x+1)} + \frac{(k-1)k!\Gamma(x+2)}{\Gamma(k+x+2)} = 1 \end{aligned} \tag{44}$$

(Note that $g_k(x)$ is decreasing in x, $g_k(0) > 1$ and $g_k(1) < 1$. Hence, α_k exists.)

Proposition 23. *Let $k \geq 2$ and $\varepsilon_k = 2$ for $k \geq 3$, $\varepsilon_2 = 1$. As $n \to \infty$, in probability and with convergence of all moments, $n^{-\alpha_k/k}\mathcal{C}_n(\xi) \to X_k$ for some random variable X_k with mean $\kappa_k := \mathbf{E}[X_k] > 0$.*

Proof. We use the same continuous-time model as in the previous proof. As explained in [23], for $k = 2$, the masses of essential fragments corresponding to nodes on the path from 0 to ξ in the dual tree form a non-conservative fragmentation process with index of self-similarity 2 and reproduction law $\mathfrak{L}((\Delta_1, \mathbf{1}_{[U_1,U_2)}(\xi)\Delta_2))$. (We abbreviate that fragments of size 0 vanish instantaneously.) Let $\mathcal{C}_t(\xi)$ be the height of the node associated to ξ in the dual tree at time t and $E_t(\xi)$ be the number of essential fragments associated to nodes on the path from 0 to ξ. For $k \geq 3$, we have $\mathcal{C}_t(\xi) = 2(E_t(\xi) - 1)$. The sizes of essential fragments form a non-conservative fragmentation process with index of self-similarity k and reproduction law $\mathfrak{L}((\Delta_1, \mathbf{1}_{[U_1,U_2)}(\xi)\Delta_2, \ldots, \mathbf{1}_{[U_{k-1},U_k)}(\xi)\Delta_k))$. Hence, by [13, Theorem 1], as $t \to \infty$, $t^{-\alpha_k/k}\mathbf{E}\left[E_t(\xi)\right] \to \kappa_k/2$ for some $\kappa_k > 0$. Furthermore, by [13, Theorem 5], there exists a random variable X_k' such that, $t^{-\alpha_k/k}E_t(\xi) \to X_k'$ in L_2. With $X_k = 2X_k'$, the claim follows by standard depoissonization arguments as in the previous proof. □

Let $\mathcal{Y}_n(s) = \mathcal{C}_n(s)/\mathbf{E}\left[\mathcal{C}_n(\xi)\right]$. We expect that, as $n \to \infty$, we have $n^{-\alpha_k/k}\mathbf{E}\left[\mathcal{C}_n(s)\right] \to m_k(s)$ for some continuous excursion $m_k \in \mathcal{C}_{\text{ex}}$ with $\mathbf{E}\left[m_k(\xi)\right] = \kappa_k$. Thus, from (45), it follows that, if

$\mathcal{Y}_n(s) \to \mathcal{Z}(s)$ in $\mathcal{D}$ for some continuous process $\mathcal{Z}$, then the limit should have mean function m_k/κ_k and satisfy

$$\mathcal{Z}(\,\cdot\,) \stackrel{d}{=} \mathbf{1}_{[0,U_1)}(\,\cdot\,)\Delta_1^{\alpha_k}\mathcal{Z}^{(1)}\left(\frac{\cdot}{\Delta_1}\right) + \mathbf{1}_{[U_k,1]}(\,\cdot\,)\Delta_1^{\alpha_k}\mathcal{Z}^{(1)}\left(\frac{\cdot - U_k}{\Delta_1}\right) + \sum_{i=2}^{k}\mathbf{1}_{[U_{i-1},U_i]}(\,\cdot\,)\left(\Delta_i^{\alpha_k}\mathcal{Z}^{(i)}\left(\frac{\cdot - U_{i-1}}{\Delta_i}\right) + \Delta_1^{\alpha_k}\mathcal{Z}^{(1)}(\xi^*)\right), \tag{45}$$

where $\mathcal{Z}^{(1)}(\cdot), \ldots, \mathcal{Z}^{(k)}(\cdot)$ are independent copies of $\mathcal{Z}(\cdot)$ independent of $(U_1, \ldots, U_k)$. This fixed-point equation is of type (10) where

$$K = k, \quad L = k+1, \quad \vartheta_3 = \ldots = \vartheta_k = 1, \quad \mathcal{R} = \mathcal{S}, \quad \mathfrak{L}(\mathcal{S}) = \text{Dirichlet}(2, 1, \ldots, 1).$$

Let $\mathcal{Z}$ be the unique process (in distribution) solving (45) with $\mathbf{E}[\mathcal{Z}_k(\xi)] = \kappa_k$ whose existence is guaranteed by Theorem 1. The verification of $\mathcal{Y}_n \to \mathcal{Z}$ in distribution in the space $\mathcal{D}$ can be worked out by the same arguments as in [17, Section 3] relying on the contraction method both for real-valued random variables and for regular processes. Here, starting with an independent family $(U_1^{(i)}, \ldots, U_k^{(i)}), i \geq 1$ of copies of $(U_1, \ldots, U_k)$ one constructs the sequence $\mathcal{Y}_n$ and its limit $\mathcal{Z}$ satisfying (45) on the same probability space and shows convergence in probability. The steps are very similarly to the arguments in the proof of Proposition 13. First, one shows the convergence $\mathcal{Y}_n(\psi) \to \mathcal{Z}(\psi)$ in L_2, where ψ corresponds to the point ξ^* in the coupling. From the last proposition we know that this convergence also holds in L_m for all $m \geq 1$. Finally, one shows that $\mathbf{E}[\|\mathcal{Y}_n - \mathcal{Z}\|^m] \to 0$ for all $m \geq 1$. (The almost sure convergence in [17] relies on a convergence rate for the mean of $\mathcal{Y}_n(\xi)$. We do not pursue this line here but note that, sufficient rates in the continuous-time case can be extracted from [13], compare the discussion of Theorem 1 there.) Summarizing, we obtain the following result.

Theorem 24. *Let $k \geq 3$. In distribution with respect to the Gromov-Hausdorff topology, we have*

$$(\mathcal{T}_n, n^{-\alpha_k/k}d) \to (\mathcal{T}_\mathcal{Z}, d_\mathcal{Z}).$$

In distribution, the process $\mathcal{Z}$ is the unique continuous excursion to (45) up to a multiplicative constant. Almost surely, $\dim_{\mathrm{M}}(\mathcal{T}_\mathcal{Z}) = \dim_{\mathrm{H}}(\mathcal{T}_\mathcal{Z}) = \alpha_k^{-1}$. For any $\gamma < \alpha_k$, almost surely, there exists a process $\tilde{\mathcal{Z}}$ equivalent to $\mathcal{Z}$ which has γ-Hölder continuous paths.

Remark. 1) For $k = 2$, almost surely, the process $\mathcal{Z}$ has γ-Hölder continuous path for any $\gamma < \alpha_2 = \beta$ [23, Theorem 1.1]. We think that this remains true for all $k \geq 3$, that is, the function $\mathcal{Z}$ is a good encoding of the tree. However, a proof using the Kolmogorov–Chentsov theorem as in [23] seems difficult.
2) Let L_n be the set of chords inserted at time n. By Proposition 7 $\mathscr{D}(\mathcal{T}_\mathcal{Z}) = \{1, 2, k\}$, and it is not hard to see that $\overline{\bigcup_{n\geq 1} L_n}$ is indeed a k-angulation of the disk: every connected component in its complement is a convex k-gon with vertices on the circle.

From (45) it follows that $m_k(t) = \mathbf{E}[\mathcal{Z}(t)]$ is the unique continuous excursion on $[0,1]$ with $\mathbf{E}[m_k(\xi)] = \kappa_k$ such that, for all $t \in [0,1]$,

$$m_k(t) = \mathbf{E}\left[\mathbf{1}_{[0,U_1)}(t)\Delta_1^{\alpha_k}m_k\left(\frac{t}{\Delta_1}\right)\right] + \mathbf{E}\left[\mathbf{1}_{[U_k,1]}(t)\Delta_1^{\alpha_k}m_k\left(\frac{t-U_k}{\Delta_1}\right)\right] + \sum_{i=2}^{k}\mathbf{E}\left[\mathbf{1}_{[U_{i-1},U_i]}(t)\left(\Delta_i^{\alpha_k}m_k\left(\frac{t-U_{i-1}}{\Delta_i}\right) + \Delta_1^{\alpha_k}m_k(\xi^*)\right)\right].$$

For $k \geq 3$, we have no explicit expression for m_k. Using the last expression and some geometric arguments relying directly on the construction of the process, one can show that m_k is infinitely differentiable on $(0,1)$, symmetric at $t = 1/2$ and monotonically increasing and concave on $[0, 1/2]$. Since we do not use these observation, we omit the details.

The homogeneous k-angulation. In the homogeneous setting, in each step, one *essential* fragment is chosen uniformly at random and k uniformly chosen points selected to create a new k-gon. At time n this leads to a decomposition of the disk into $1+(k-1)n$ essential fragments and n non-essential fragments. The definitions of $\mathcal{T}_n^h, \mathcal{C}_n^h, U_1, \ldots, U_k, \Delta_1 = 1 - U_k + U_1, \Delta_i = U_i - U_{i-1}, 2 \le i \le K, \xi^* = U_1/\Delta_1$ as well as $I_n = (I_n^{(1)}, \ldots, I_n^{(k)})$ should be clear by now. By construction, the random variable I_n is independent of $(U_1, \ldots, U_k)$ and grows like the vector of occupation numbers in a Polya urn model. It is well-known that, almost surely,

$$\frac{I_n}{n} \to (\tilde{\Delta}_1, \ldots, \tilde{\Delta}_k),$$

where $\tilde{\Delta} = (\tilde{\Delta}_1, \ldots, \tilde{\Delta}_k)$ has the Dirichlet$(1/(k-1), \ldots, 1/(k-1))$ distribution. By construction, the random process $(C_n^h(s))_{s\in[0,1]}$ satisfies a recursion analogous to (43), the only difference being the distribution of $(U_1, \ldots, U_k, I_n)$.

Proposition 25. *Let $k \ge 2$. As $n \to \infty$, almost surely and with convergence of all moments, we have $n^{-1/(k+1)}\mathcal{C}_n^h(\xi) \to X_k^h$ for some random variable X_k^h with mean $\kappa_k^h := \mathbf{E}[X_k^h] > 0$.*

Proof. In the standard continuous-time embedding of the process, every essential fragment splits into k essential subfragments at unit rate, independently of its mass. Hence, the number of essential fragments $\mathcal{N}_t, t \ge 0$, forms a continuous-time branching process with offspring distribution δ_k, see, e.g. [11]. It is well-known that $e^{-t(k-1)}\mathcal{N}_t, t > 0$, is a uniformly integrable martingale with mean one converging almost surely to a limiting random variable denoted by $\mathcal{N}$ having the Gamma$((k-1)^{-1}, (k-1)^{-1})$ distribution. For the time of n-th split τ_n in the process, since $\mathcal{N}_{\tau_n} = 1 + (k-1)n$, by the optional stopping theorem, it follows that $(k-1)ne^{-\tau_n(k-1)} \to \mathcal{N}$ almost surely and in mean. Similarly, the number of essential fragments on the path from 0 to ξ in the dual tree denoted by $E_t^h(\xi)$ forms a branching process with offspring distribution $\mathfrak{L}(1 + \mathbf{1}_{[U_1,U_k)}(\xi))$. Again, the process $e^{-t(k-1)/(k+1)}E_t^h(\xi), t > 0$, is a uniformly-integrable martingale with unit mean and we denote its almost sure limit by $\mathcal{E}$. Writing

$$e^{-\tau_n(k-1)/(k+1)}E_{\tau_n}^h(\xi) = e^{-\tau_n(k-1)/(k+1)}((k-1)n)^{1/(k+1)} \cdot ((k-1)n)^{-1/(k+1)}E_{\tau_n}^h(\xi),$$

and noting that the random variables $\tau_n, E_{\tau_n}^h(\xi)$ are independent, it follows that $((k-1)n)^{1/(k+1)}E_{\tau_n}^h(\xi) \to \mathcal{E}'$ almost surely and in mean where $\mathcal{E} = \mathcal{E}'\mathcal{N}^{1/(k+1)}$. In particular,

$$\mathbf{E}[\mathcal{E}'] = (k-1)^{-1/(k-1)}\frac{\Gamma(\frac{1}{k-1})}{\Gamma(\frac{2}{k-1})}.$$

Convergence of higher moments can be deduced as in the recursive model. □

For $s \in [0,1]$, let $\mathcal{Y}_n^h(s) = \mathcal{C}_n^h(s)/\mathbf{E}[\mathcal{C}_n^h(\xi)]$. A limit $\mathcal{Z}^h(s)$ in $\mathcal{D}$ of $\mathcal{Y}_n^h(s)$ should satisfy $\mathbf{E}[\mathcal{Z}^h(\xi)] = \kappa_k^h$ and

$$\begin{aligned}\mathcal{Z}^h(\,\cdot\,) \stackrel{d}{=} \mathbf{1}_{[0,U_1)}(\,\cdot\,)\tilde{\Delta}_1^{1/(k+1)}\mathcal{Z}^{h,(1)}\left(\frac{\cdot}{\Delta_1}\right) + \mathbf{1}_{[U_k,1]}(\,\cdot\,)\tilde{\Delta}_1^{1/(k+1)}\mathcal{Z}^{h,(1)}\left(\frac{\cdot - U_k}{\Delta_1}\right)\\ + \sum_{i=2}^{k}\mathbf{1}_{[U_{i-1},U_i]}(\,\cdot\,)\left(\tilde{\Delta}_i^{1/(k+1)}\mathcal{Z}^{h,(i)}\left(\frac{\cdot - U_{i-1}}{\Delta_i}\right) + \tilde{\Delta}_1^{1/(k+1)}\mathcal{Z}^{h,(1)}(\xi^*)\right),\end{aligned} \tag{46}$$

where $\mathcal{Z}^{h,(1)}, \ldots, \mathcal{Z}^{h,(k)}$ are independent copies of $\mathcal{Z}^h$ independent of $(U_1, \ldots, U_k), \tilde{\Delta}$. This fixed-point equation is of type (10) where

$$L = k+1, \quad \vartheta_3 = \ldots = \vartheta_k = 1, \quad \mathcal{R} = \text{Dirichlet}(1,1,\ldots,1), \quad \mathfrak{L}(\mathcal{S}) = \text{Dirichlet}(2,1,\ldots,1),$$

and $\mathcal{R}$ and $\mathcal{S}$ are independent. As in the recursive model, one can prove the following theorem.

Theorem 26. *Let $k \geq 3$. In distribution with respect to the Gromov-Hausdorff topology, we have*

$$(\mathcal{T}_n^h, n^{-1/(k+1)} d_n^h) \to (\mathcal{T}_{\mathcal{Z}^h}, d_{\mathcal{Z}^h}).$$

In distribution, the process $\mathcal{Z}^h$ is the unique continuous excursion satisfying (46) up to a multiplicative constant. Almost surely, $\dim_{\mathrm{M}}(\mathcal{T}_{\mathcal{Z}^h}) = \dim_{\mathrm{H}}(\mathcal{T}_{\mathcal{Z}^h}) = k+1$. For any $\gamma < 1/(k+1)$, almost surely, there exists a process $\tilde{\mathcal{Z}}^h$ equivalent to $\mathcal{Z}^h$ which has γ-Hölder continuous paths.

The comments on the mean function m_k from the recursive model extend straightforwardly to m_k^h.

References

[1] R. Abraham, J.-F. Delmas, and P. Hoscheit. A note on the Gromov-Hausdorff-Prokhorov distance between (locally) compact metric measure spaces. *Electron. J. Probab.*, 18(14):21, 2013.

[2] M. Aizenman and A. Burchard. Hölder regularity and dimension bounds for random curves. *Duke Math. J.*, 99(3):419–453, 1999.

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

[4] D. Aldous. The continuum random tree. I. *Ann. Probab.*, 19(1):1–28, 1991.

[5] D. Aldous. The continuum random tree II: an overview. In M.T. Barlow and N.H. Bingham, editors, *Stochastic Analysis*, pages 23–70. Cambridge University Press, 1991.

[6] D. Aldous. The continuum random tree III. *Ann. Probab.*, 21:248–289, 1993.

[7] D. Aldous. Recursive self-similarity for random trees, random triangulations and Brownian excursion. *Ann. Probab.*, 22(2):527–545, 1994.

[8] D. Aldous. Triangulating the circle, at random. *Amer. Math. Monthly*, 101(3):223–233, 1994.

[9] D. J. Aldous and A. Bandyopadhyay. A survey of max-type recursive distributional equations. *Ann. Appl. Probab.*, 15(2):1047–1110, 2005.

[10] G. Alsmeyer and M. Meiners. Fixed points of the smoothing transform: two-sided solutions. *Probab. Theory Related Fields*, 155(1-2):165–199, 2013.

[11] K. B. Athreya and S. Karlin. Embedding of urn schemes into continuous time Markov branching processes and related limit theorems. *Ann. Math. Statist.*, 39:1801–1817, 1968.

[12] S. Athreya, W. Löhr, and A. Winter. The gap between Gromov-vague and Gromov-Hausdorff-vague topology. *Stochastic Process. Appl.*, 126(9):2527–2553, 2016.

[13] J. Bertoin and A. Gnedin. Asymptotic laws for nonconservative self-similar fragmentations. *Electron. J. Probab.*, 9:575–593, 2004.

[14] P. J. Bickel and D. A. Freedman. Some asymptotic theory for the bootstrap. *Ann. Statist.*, 9(6):1196–1217, 1981.

[15] P. Billingsley. *Convergence of probability measures.* New York-London-Sydney-Toronto: John Wiley and Sons, Inc. XII, 253 p. , 1968.

[16] N. Broutin and J.-F. Marckert. Asymptotics of trees with a prescribed degree sequence and applications. *Random Structures & Algorithms*, 44:290–316, 2014.

[17] N. Broutin and H. Sulzbach. The dual tree of a recursive triangulation of the disk. *Ann. Probab.*, 43(2):738–781, 2015.

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

[19] J. T. Chayes, L. Chayes, and R. Durrett. Connectivity properties of Mandelbrot's percolation process. *Probab. Theory Related Fields*, 77(3):307–324, 1988.

[20] V. V. Chistyakov and O. E. Galkin. On maps of bounded p-variation with $p > 1$. *Positivity*, 2(1):19–45, 1998.

[21] R. Cori and B. Vauquelin. Planar maps are well labeled trees. *Canadian Journal of Mathematics*, 33:1023–1042, 1981.

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

[23] N. Curien and J.-F. Le Gall. Random recursive triangulations of the disk via fragmentation theory. *Ann. Probab.*, 39(6):2224–2270, 2011.

[24] N. Curien and Y. Peres. Random laminations and multitype branching processes. *Electron. Commun. Probab.*, 16:435–446, 2011.

[25] A. Dembo and O. Zeitouni. *Large Deviation Techniques and Applications*. Springer, second edition, 1998.

[26] A. Depperschmitdt, A. Greven, and P. Pfaffelhuber. Marked metric measure spaces. *Electr. J. Probab.*, 16: 174–188, 2011.

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

[28] T. Duquesne. The coding of compact real trees by real valued functions. Preprint, arXiv:0604106.

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

[30] T. Duquesne and J.-F. Le Gall. The Hausdorff measure of stable trees. *ALEA Lat. Am. J. Probab. Math. Stat.*, 1:393–415, 2006.

[31] R. Durrett and T. M. Liggett. Fixed points of the smoothing transformation. *Z. Wahrsch. Verw. Gebiete*, 64 (3):275–3719, 1983.

[32] S. N. Evans. *Probability and real trees*, volume 1920 of *Lecture Notes in Mathematics*. Springer, Berlin, 2008. Lectures from the 35th Summer School on Probability Theory held in Saint-Flour, July 6–23, 2005.

[33] K. J. Falconer. *The Geometry of Fractal Sets*, volume 85 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1986.

[34] 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.

[35] M. Gromov. *Metric Structures for Riemannian and Non-Riemannian Spaces*. Birkhauser, 1999.

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

[37] S. Janson. Moments of gamma type and the brownian supremum process area. *Probab. Surveys*, 7:1–52, 2010.

[38] A. S. Kechris. *Classical descriptive set theory*, volume 156 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995.

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

[40] J.-F. Le Gall. Uniqueness and universality of the Brownian map. *Ann. Probab.*, 41:2880–2960, 2013.

[41] W. Löhr. Equivalence of Gromov-Prokhorov- and Gromov's $\underline{\square}_\lambda$-metric on the space of metric measure spaces. *Electr. J. Probab.*, 18:1–10, 2013.

[42] J.-F. Marckert and G. Miermont. The CRT is the scaling limit of unordered binary trees. *Random Structures & Algorithms*, 38:467–501, 2011.

[43] S.P. Meyn and R.L. Tweedie. *Markov chains and stochastic stability*. Springer, New York, 1993.

[44] G. Miermont. The Brownian map is the scaling limit of uniform random plane quadrangulations. *Acta Math.*, 210:319–401, 2013.

[45] R. Neininger and H. Sulzbach. On a functional contraction method. *Ann. Probab.*, 43(4):1777–1822, 2015.

[46] K. Panagiotou and B. Stufler. Scaling limits of random Pólya trees. Preprint, arXiv:1502.07180, 2015.

[47] K. Panagiotou, B. Stufler, and K. Weller. Scaling limits of random graphs from subcritical classes. *to appear in Ann. Probab.*, 2016.

[48] J. Picard. A tree approach to p-variation and to integration. *Ann. Probab.*, 36(6):2235–2279, 2008.

[49] E. E. Posey and J. E. Vaughan. Extrema and nowhere differentiable functions. *Rocky Mountain Journal of Mathematics*, 16(4):661–668, 1986.

[50] U. Rösler. A fixed point theorem for distributions. *Stochastic Process. Appl.*, 37:195–214, 1992.

[51] G. Schaeffer. *Combinaison d'arbres et cartes combinatoires aléatoires*. Phd thesis, Université de Bordeaux, 1998.

[52] R. Stephenson. General fragmentation trees. *Electron. J. Probab.*, 18(101), 2013.

[53] B. Stufler. The continuum random tree is the scaling limit of unlabelled unrooted trees. arXiv:1412.6333, 2014.

[54] A. M. Vershik. The universal Urysohn space, Gromov metric triples and random metrics on the natural numbers. *Russian Mathematical Surveys*, 53(5):921, 1998.

[55] W. Vervaat. On a stochastic difference equation and a representation of non-negative infinitely divisible random variables. *Adv. Appl. Probab.*, 11:750–783, 1979.

A Hölder continuity of $\mathscr{H}$

Proof of Proposition 21. The positive result for $\alpha < \varrho$ follows from bounds on the moments of $\mathscr{H}$ provided in Lemma 27 and Kolmogorov's criterion. Next, recall from [17] that $h(t) := \mathbf{E}[\mathscr{H}(t)] = \kappa'\sqrt{t(1-t)}$ for some $\kappa' > 0$. Let $\varrho < \alpha < \gamma$. With Q_n^ϑ defined in (22), where $Q_0^\vartheta = h$, the uniform almost sure limit $\mathcal{X}$ of $Q_n^\emptyset$ in Proposition 13 is measurable with respect to $\{(\mathcal{R}^\sigma, \mathcal{S}^\sigma, \Xi^\sigma) : \sigma \in \Theta\}$. Thus, since γ-Hölder continuity is a tail event of this σ-algebra, by Kolmogorov's zero-one law, it suffices to show that $\mathcal{X}$ fails to be γ-Hölder continuous with positive probability. For $n \geq 1$, let $C_n = \cup_{|\vartheta|=n}\partial\Lambda_\vartheta$. Observe that, for some $r > 0$, we have $|h(y)-h(x)| \geq r|y-x|^2$ for all $x, y \in [0,1/2]$ and $x, y \in [1/2, 1]$. Fix $\vartheta \in \Theta_\ell$. Let $x, y \in \Lambda_\vartheta$ with $(x,y) \cap C_\ell = \emptyset$ and denote by x_ℓ, y_ℓ their relative positions inside Λ_ϑ, that is $x_\ell = (x - \inf \Lambda_\vartheta)/\mathcal{L}(\vartheta)$, analogously for y_ℓ. If $x_\ell, y_\ell \in [0, 1/2]$ or $x_\ell, y_\ell \in [1/2, 1]$, then

$$\mathcal{V}(\vartheta)^{1/3} \geq \ell^{2-\alpha}\mathcal{L}(\vartheta)^\alpha$$

implies

$$\begin{aligned} |Q_\ell^\emptyset(y) - Q_\ell^\emptyset(x)| &= \mathcal{V}(\vartheta)^{1/3}|h(y_\ell) - h(x_\ell)| \\ &\geq \ell^{2-\alpha}\mathcal{L}(\vartheta)^\alpha|h(y_\ell) - h(x_\ell)| \geq r\ell^{2-\alpha}|y-x|^\alpha|y_\ell - x_\ell|^{2-\alpha}. \end{aligned}$$

As Λ_ϑ is the union of at most $\ell + 1$ intervals, we can find $x, y \in \Lambda(\vartheta), (x,y) \cap C_\ell = \emptyset$ satisfying the latter inequality with $|y_\ell - x_\ell| \geq 1/(4\ell)$ where $x_\ell, y_\ell \in [0, 1/2]$ or $x_\ell, y_\ell \in [1/2, 1]$. Hence, for these x, y we deduce

$$|Q_\ell^\emptyset(y) - Q_\ell^\emptyset(x)| \geq \frac{|y-x|^\alpha}{16/r}.$$

As $n \to \infty$, almost surely, the maximal distance between consecutive points in C_n converges to zero. Hence, $\mathcal{X}$ is not γ-Hölder continuous if there exists $n \in \mathbb{N}$ and an infinite path $\vartheta = \varepsilon_1\varepsilon_2\ldots$ such that, for all $k \in \mathbb{N}$, with $\vartheta_n = \varepsilon_1 \ldots \varepsilon_n$,

$$\mathcal{V}(\vartheta_{kn})^{1/3} \geq (kn)^{2-\alpha}\mathcal{L}(\vartheta_{kn})^\alpha.$$

Below, we will show that this event has positive probability for some $n \in \mathbb{N}$ (in fact, for all n large enough). For $\vartheta, \sigma \in \Theta$, let

$$A_\sigma^\vartheta = \left\{ \frac{\mathcal{V}(\vartheta\sigma)^{1/3}}{\mathcal{V}(\vartheta)^{1/3}} \geq |\sigma|^{2-\alpha}\frac{\mathcal{L}(\vartheta\sigma)^\alpha}{\mathcal{L}(\vartheta)^\alpha} \right\}.$$

Let $N := 2^n$ and Θ^* be the complete N-ary tree with nodes on level k denoted by $\Theta_{k,n}$. Moreover, let S be the random subtree of Θ^* in which a node $\vartheta^* = \vartheta_1^* \ldots \vartheta_k^*$ with $\vartheta_1^*, \ldots, \vartheta_k^* \in \Theta_n$ on level k exists if, for all $0 \leq i \leq k-1$, the event $A_{\vartheta_{i+1}^*}^{\vartheta_1^*\ldots\vartheta_i^*}$ occurs. By construction, for fixed $n \geq 1$,

$$\left\{ \left(\mathbf{1}_{A_\sigma^\vartheta}\right)_{\sigma \in \Theta_n} : \vartheta \in \Theta_{k,n}, k \in \mathbb{N} \cup \{0\} \right\}$$

is a family of independent and identically distributed random vectors. Thus, S is a branching process with offspring mean

$$\sum_{\vartheta \in \Theta_n} \mathbf{P}(A_\vartheta^\emptyset) = \sum_{\vartheta \in \Theta_n} \mathbf{P}\left(\mathcal{V}(\vartheta)^{1/3} > n^{2-\alpha}\mathcal{L}(\vartheta)^\alpha\right).$$

By the elementary formula $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$ it is easy to see that

$$\sum_{\vartheta \in \Theta_n} \mathbf{P}(A_\vartheta^\emptyset) = 2^n \mathbf{P}\left(\prod_{i=1}^n W_i^{1/3} > n^{2-\alpha}\prod_{i=1}^n \overline{W}_i^\alpha\right),$$

where $W_1, \overline{W}_1, \ldots, W_n, \overline{W}_n$ are independent and identically uniformly distributed on the unit interval. We may assume $\alpha < 1/3$. Let $\delta > 0$ and $E_1, F_1, \ldots, E_n, F_n$ be independent standard exponentials. Then, by an application of Cramér's theorem for sums of independent and identically distributed random variables with finite momentum generating function in a neighborhood of zero, for all n sufficiently large,

$$\mathbf{P}\left(\prod_{i=1}^{n} W_i^{1/3} > n^{2-\alpha} \prod_{i=1}^{n} \overline{W}_i^{\alpha}\right) \geq \mathbf{P}\left(\sum_{i=1}^{n} \alpha F_i - \frac{1}{3} E_i > \delta n\right) = \exp(-I(\delta) n + o(n)),$$

where $I(x), x \in [\alpha - 1/3, \infty)$ denotes the large deviations rate function of the random variable $\alpha F_1 - E_1/3$ [see, e.g., 25] given by

$$I(x) = \sup_{-3<s<1/\alpha} sx - \log \frac{3}{3+s} \frac{1}{1-s\alpha} \leq \frac{x}{\alpha} - \inf_{-3<s<1/\alpha} \log \frac{3}{3+s} \frac{1}{1-s\alpha} = \frac{x}{\alpha} - \log \frac{12\alpha}{(3\alpha+1)^2}.$$

Thus, for all n large enough,

$$\sum_{\vartheta \in \Theta_n} \mathbf{P}(A_{\vartheta}^{\emptyset}) \geq (2c(1+o(1)))^n, \quad c = \frac{24\alpha}{(3\alpha+1)^2} e^{-\delta/\alpha}.$$

Since $\alpha > \varrho$, upon choosing $\delta > 0$ sufficiently small, we obtain $c > 1/2$. Thus, for all n sufficiently large, with positive probability, there exists an infinite path $\vartheta_1^* \vartheta_2^* \ldots$ in Θ^* with $\vartheta_i^* \in \Theta_n$ such the events $A_{\vartheta_{i+1}^*}^{\vartheta_1^* \ldots \vartheta_i^*}$ occur for all $i \in \mathbb{N} \cup \{0\}$. Along this path written as $\vartheta = \varepsilon_1 \varepsilon_2 \ldots$, we deduce

$$\mathcal{V}(\vartheta_{kn}) \geq n^{(2-\alpha)k} \mathcal{L}(\vartheta_{kn})^{\alpha} \geq (kn)^{2-\alpha} \mathcal{L}(\vartheta_{kn})^{\alpha}.$$

for all $k \in \mathbb{N}$. This concludes the proof. $\square$

The final lemma generalizes Proposition 4.1 in [23] to non-integer values of p.

Lemma 27. *For all $\varepsilon > 0$ and $p \in [0, \infty)$, there exists $K > 0$ such that, for all $x \in [0, 1]$,*

$$\mathbf{E}[\mathscr{H}(x)^p] \leq K(x(1-x))^{2p/(p+3)-\varepsilon}$$

Proof. We provide the minor modifications necessary to extend Proposition 4.1 in [23] to the non-integer case without presenting tedious calculations. First, by Jensen's inequality, since $\mathbf{E}[\mathscr{H}(x)] = \kappa' \sqrt{x(1-x)}$ with $\kappa' = 1/\Gamma(4/3)$, we have

$$\mathbf{E}[\mathscr{H}(x)^p] \leq (\kappa')^p (x(1-x))^{p/2}, \quad 0 \leq p \leq 1, \tag{47}$$

$$\mathbf{E}[\mathscr{H}(x)^p] \geq (\kappa')^p (x(1-x))^{p/2}, \quad p \geq 1. \tag{48}$$

Thus, for $0 \leq p \leq 1$, the assertion follows immediately from (47). For $p \in (1, \infty)$, we do not have a integral recursion for $m_p(t) = \mathbf{E}[\mathscr{H}(t)^p]$ such as (17) in [23] unless p is integer. However, applying the inequality $(a+b)^p \leq a^p + b^p + C_1(a^{p-1}b + ab^{p-1})$ for $a, b \geq 0$ and some $C_1 = C_1(p)$, to the stochastic fixed-point equation (13), we have, in a stochastic sense,

$$\begin{aligned}
\mathscr{H}(t)^p \leq & \mathbf{1}_{[0,U_1)}(t) W^{p/3} (\mathscr{H}^{(1)})^p \left(\frac{t}{\Delta_1}\right) + \mathbf{1}_{[U_2,1]}(t) W^{p/3} (\mathscr{H}^{(1)})^p \left(\frac{t-\Delta_2}{\Delta_1}\right) \\
& + \mathbf{1}_{[U_1,U_2)}(t) \left((1-W)^{p/3} (\mathscr{H}^{(2)})^p \left(\frac{t-U_1}{\Delta_2}\right) + W^{p/3} (\mathscr{H}^{(1)})^p (\xi)\right) \\
& + \mathbf{1}_{[U_1,U_2)}(t) \left(C_1 (1-W)^{(p-1)/3} W^{1/3} (\mathscr{H}^{(2)})^{p-1} \left(\frac{t-U_1}{\Delta_2}\right) \mathscr{H}^{(1)}(\xi)\right) \qquad (49) \\
& + \mathbf{1}_{[U_1,U_2)}(t) \left(C_1 (1-W)^{1/3} W^{(p-1)/3} (\mathscr{H}^{(2)}) \left(\frac{t-U_1}{\Delta_2}\right) (\mathscr{H}^{(1)})^{p-1} (\xi)\right), \qquad (50)
\end{aligned}$$

with conditions as in (13) on the right hand side. Subsequently, we consider $0 \le t \le 1/2$ which suffices by symmetry. With $q = 3/(3+p)$, taking the expectation on both sides of the last display leads to

$$m_p(t) \le 2q(1-t)^2 \int_0^t m_p(x)(1-x)^{-3}dx + 2qt^2 \int_t^1 m_p(x)x^{-3}dx + s_p(t),$$

where $s_p(t)$ is the sum of the expectation of (49) and (50). As in the proof of Proposition 4.1 in [23], relying on first, Lemma 4.2 there for $\mathscr{H}$ instead of M, and second, a stochastic inequality inverse to the display above based on $(a+b)^p \ge a^p + b^p$ for $a, b \ge 0$, one can show that the first summand has negligible contribution as $t \to 0$. In other words, for any $\delta > 0$, there exists t_0 such that, for $t \le t_0$, we have

$$m_p(t) \le 2(1+\delta)qt^2 \int_t^1 m_p(x)x^{-3}dx + 2s_p(t). \tag{51}$$

Furthermore, for some $C_2 = C_2(p, t_0, \delta) > 0$,

$$m_p(t) \le 2(1+\delta)qt^2 \int_t^{t_0} m_p(x)x^{-3}dx + 2s_p(t) + C_2t^2.$$

Now, if we were to drop $s_p(t)$, then, by applying Gronwall's lemma to the function $m_p(t)t^{-2}$, we could deduce

$$m_p(t) \le C_3 t^{2p/(p+3)-\delta q}$$

for all $t \in [0,1]$ and some $C_3 = C_3(p, t_0, \delta)$. This would give the assertion as δ was chosen arbitrarily. For a rigorous verification, we start with the case $1 < p \le 2$. Then, a direct computation shows that, for some $C_4 = C_4(p)$, we have $s_p(t) \le C_4 t^{(p+2)/2}$. Thus, by (48), $s_p(t)$ is asymptotically negligible compared to $m_p(t)$. Using (51), for any $\delta' > 0$, upon decreasing t_0 if necessary, we have

$$m_p(t) \le 2(1+\delta)(1+\delta')qt^2 \int_t^1 m_p(x)x^{-3}dx.$$

As indicated, the claim now follows from Gronwall's lemma with a suitable choice of δ and δ'. For $p > 2$, we proceed by induction. We may assume that $\varepsilon > 0$ is chosen small enough such that $3/2 + 2(p-1)/(p+2) - \varepsilon > 2$. By the induction hypothesis, there exists $C_5 = C_5(p)$, such that $s_p(t) \le C_5 t^{3/2+2(p-1)/(p+2)-\varepsilon}$. Thus, $s_p(t)t^{-2}$ is bounded on $[0,1]$ and the result follows as indicated from inequality (51). $\square$