Electronic Journal of Probability

Electron. J. Probab. **26** (2021), article no. 88, 1–50.
ISSN: 1083-6489 https://doi.org/10.1214/21-EJP647

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

Nicolas Broutin[†‡] Henning Sulzbach[§]

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/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 efficient to treat cases for which the mass measure on the real tree induced by natural encodings only provides weak estimates on the Hausdorff dimensions.

Keywords: continuum real tree; self-similarity; stochastic fixed point equation; fractal dimension.
MSC2020 subject classifications: 60C05; 60F17; 05C05.
Submitted to EJP on April 3, 2020, final version accepted on May 22, 2021.

1 Introduction

Since the pioneering work of Aldous [3, 5] who introduced the Brownian continuum random tree (Brownian CRT) 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

*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.

†Sorbonne Université, Campus Pierre et Marie Curie, 4 place Jussieu, 75252 Paris Cedex 05, France. E-mail: nicolas.broutin@upmc.fr

‡Institut Universitaire de France

§McGill University, 3480 University Street, H3A 0E9 Montreal, QC, Canada. Present address: School of Mathematics, University of Birmingham, Birmingham B15 2TT, Great Britain. E-mail: henning.sulzbach@gmail.com

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 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 [17, 36, 40, 46] and of Liouville quantum gravity [23] 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 CRT [6], of trees that are dual to recursive triangulations of the disk [19], but also of the genealogies of self-similar fragmentations [32]. 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 a random variable (r.v.) X taking values in some Polish space $\mathbb{S}$ of "objects" is an equation of the form

$$X \stackrel{d}{=} T((X_i)_{i\geq 1}, \Xi), \tag{1.1}$$

where $(X_i)_{i\geq 1}$ 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 random metric spaces is more involved. See display (2.4) below.) The fact that natural objects satisfy equations such as the one in (1.1) 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 it.

One of the striking features of random real trees that appear ubiquitous is their fractal nature. Among the most classical real trees one may cite the Lévy trees (including the Brownian CRT 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 [25, 26] and Picard [44]. 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 [32] and more recently by Stephenson [47]. In view of the recursive self-similarity of Equation (1.1), 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 CRT 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 be compared with the classical decomposition of the Brownian CRT into three pieces [2, 6, 18]. 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 (1.1):

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

Our aim in this paper is to provide answers to questions (i)–(v) in a general framework in which the limit objects are (most often) 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 [2] for the specific example where the fixed point equation is the one described by Aldous in [6] and that is satisfied by the Brownian CRT. In passing, our results answer a question in [2] 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 [19] and [14]), and but also other natural generalizations. Rembart and Winkel [45] have also very recently studied questions (i), (ii) and (iii) for a decomposition of the form (2.4) which is rather different from ours. See the remark at the end of Section 2.3 for details.

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. Various proofs of technical results are given in appendix.

2 Settings and preliminaries

2.1 Spaces, metrics and convergence

With the exception of Section 2.5, we assume throughout the paper that metric spaces are *compact*. General references on the topics that we are about to discuss include [15, 30, 31]. For measured spaces, we restrict our attention to *probability measures.*

The Gromov–Hausdorff–Prokhorov topology. For two compact metric spaces $(\mathcal{X}, d)$ and $(\mathcal{X}', d')$, the *Gromov–Hausdorff distance* $\mathrm{d_{gh}}((\mathcal{X}, d), (\mathcal{X}', d'))$ is defined as

$$\mathrm{d_{gh}}((\mathcal{X}, d), (\mathcal{X}', d')) = \inf_{Z,\phi,\phi'} \mathrm{d_h^Z}(\phi(\mathcal{X}), \phi'(\mathcal{X}')), \tag{2.1}$$

where the infimum is taken over all compact metric spaces (Z, d^Z), and isometries $\phi : \mathcal{X} \to Z$ and $\phi' : \mathcal{X}' \to Z$. 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 $A^\varepsilon = \{x \in Z : d^Z(x, A) \leq \varepsilon\}$. If $(\mathcal{X}, d)$ and $(\mathcal{X}', d')$ are isometric, then $\mathrm{d_{gh}}((\mathcal{X}, d), (\mathcal{X}', d')) = 0$. $\mathrm{d_{gh}}$ 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. [35, Theorem 2.1].

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

$$\mathrm{d_{ghp}}(\mathfrak{X}, \mathfrak{X}') = \inf_{Z,\phi,\phi'} \big\{ d^Z(\phi(\rho), \phi'(\rho')) + \mathrm{d_h^Z}(\phi(\mathcal{X}), \phi'(\mathcal{X}')) + \mathrm{d_p^Z}(\phi_*(\mu), \phi'_*(\mu')) \big\}.$$

Here, the infimum is to be understood as in (2.1), $\phi_*(\mu)$ is the push-forward of μ in Z, and $\mathrm{d}_\mathrm{p}^\mathrm{Z}$ denotes the Prokhorov metric on the set of probability measures on Z, that is,

$$\mathrm{d}_\mathrm{p}^\mathrm{Z}(\nu_1,\nu_2) = \inf\{\varepsilon > 0 : \nu_1(A) \le \nu_2(A^\varepsilon) + \varepsilon \text{ and } \nu_2(A) \le \nu_1(A^\varepsilon) + \varepsilon \\ \text{for all measurable sets } A \subseteq \mathcal{X}\}.$$

We call $\mathfrak{X}$ and $\mathfrak{X}'$ ghp-isometric if there exists a bijective isometry ϕ between X and X' that maps ρ to ρ' and such that $\phi_*(\mu) = \mu'$. 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]. For a compact rooted measured metric space $\mathfrak{X}$ (or an element of $\mathbb{K}^\mathrm{ghp}$), we set $\|\mathfrak{X}\| = \sup\{d(\rho,x) : x \in \mathcal{X}\}$.

The Gromov–Prokhorov topology. Analogously to the Gromov–Hausdorff–Prokhorov distance, for two compact rooted measured metric spaces $\mathfrak{X} = (\mathcal{X}, d, \mu, \rho)$ and $\mathfrak{X}' = (\mathcal{X}', d', \mu', \rho')$ [1] we define

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

We call $\mathfrak{X}$ and $\mathfrak{X}'$ gp-isometric if $\mathrm{d}_\mathrm{gp}(\mathfrak{X},\mathfrak{X}') = 0$ which happens to be the case if and only if there exists a bijective isometry $\phi : \mathrm{supp}(\mu) \to \mathrm{supp}(\mu')$ with $\rho' = \phi(\rho)$ and $\mu' = \phi_*(\mu)$. (Here $\mathrm{supp}(\mu)$ denotes the support of μ.) Endowed with d_gp, the set $\mathbb{K}^\mathrm{gp}$ of gp-isometry classes of compact rooted measured metric spaces becomes a Polish space [30, Proposition 5.6]. 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) $\mathrm{supp}(\mu) = \mathcal{X}$,

then, there exists a natural bijection ι between the spaces $\mathbb{K}_\mathrm{f}^\mathrm{ghp}$ and $\mathbb{K}^\mathrm{gp}$. The set $\mathbb{K}_\mathrm{f}^\mathrm{ghp}$ is measurable, and ι bimeasurable so we can and will consider any random variable with values in $\mathbb{K}_\mathrm{f}^\mathrm{ghp}$ also as random variable in $\mathbb{K}^\mathrm{gp}$ and vice versa. (Proving these statements makes use of Lemma 3.2 and Corollary 5.6 in [9] as well as the Lusin–Souslin theorem. See Lemma 25 in the appendix for details.)

Remark. We occasionally use results from [15, 22, 30, 31] which only treat the case of unrooted compact measured metric spaces. Incorporating a root vertex only generates marginal modifications that we do not discuss in detail.

2.2 Real trees, continuum trees and recursive decompositions

We are mostly 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 $\mathcal{T} \setminus \{x\}$. We call $x \in \mathcal{T}$ a *leaf* if $\deg(x) = 1$, and abbreviate $\mathscr{L}$

[1] When we do not introduce the components of $\mathfrak{X}$ with a given decoration explicitly, we always suppose that they would carry the same decoration as done here for $\mathfrak{X}'$.

for the set of leaves. We call $x \in \mathcal{T}$ a *branch point* if $\deg(x) \geq 3$. By compactness, the set of branch points $\mathscr{B}$ is at most countable.

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*. (Recall that we restrict our attention to the setting where the measure is a probability distribution.) For $x \in \mathcal{T}$, the distance $d(x, \rho)$ is called the height of x and $\|\mathfrak{T}\| := \sup\{d(x, \rho) : x \in \mathcal{T}\}$ the *height* of $\mathfrak{T}$. By $\mathbb{T}^{\text{ghp}} \subseteq \mathbb{K}^{\text{ghp}}$ we denote the closed subset of ghp-isometry classes of compact rooted measured real trees. Spaces carrying a measure with full support are particularly important, and we let $\mathbb{T}_{\text{f}}^{\text{ghp}} = \mathbb{K}_{\text{f}}^{\text{ghp}} \cap \mathbb{T}^{\text{ghp}}$ and call elements in $\mathbb{T}_{\text{f}}^{\text{ghp}}$ *continuum real trees* (or simply continuum trees). Note that both $\mathbb{K}_{\text{f}}^{\text{ghp}}$ and $\mathbb{T}_{\text{f}}^{\text{ghp}}$ are non-closed subsets of $\mathbb{K}^{\text{ghp}}$. In the literature on continuum real trees, see, e.g. [2, 5], one often finds the following two additional conditions:

$$\textbf{(C2)} \quad \mu \text{ has no atoms} \qquad \text{and} \qquad \textbf{(C3)} \quad \mu(\mathscr{L}) = 1\,.$$

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; see Proposition 7.

Real trees encoded by excursions. One natural way to define real trees is via an encoding by continuous excursions (see e.g., [28, 35]). Let $\mathbb{C}$ 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 $\mathbb{C}_{\text{ex}}$ denote the set of unit-length non-negative continuous excursions, that is, the set of functions $f \in \mathbb{C}$ such that $f(0) = f(1) = 0$ and $f(t) \geq 0$ for all $t \in (0, 1)$. For $f \in \mathbb{C}_{\text{ex}}$, define the pseudometric d_f by

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

Let $\mathcal{T}_f = [0, 1]/\!\sim$ where $x \sim y$ if and only if $d_f(x, y) = 0$. 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 $\mathscr{L}_f$ for the sets of leaves of $\mathcal{T}_f$.

As noted in [35, Remark following Theorem 2.2] (see also [24, Corollary 1.2]), for every compact real tree $(\mathcal{T}, d)$, there exists $f \in \mathbb{C}_{\text{ex}}$ such that $(\mathcal{T}, d)$ and $(\mathcal{T}_f, d_f)$ are isometric. 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 a continuous and strictly increasing function $\phi : [0, 1] \to [0, 1]$. In this case, we call the encoding excursions f and g *time-change-equivalent* or simply *equivalent*.

Let $\pi_f : [0, 1] \to \mathcal{T}_f$ be the canonical surjection. Then, we can turn the real tree $(\mathcal{T}_f, d_f)$ 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 is $\rho_f := \pi_f(0)$. Then, μ_f has full support, and thus $\mathfrak{T}_f$ satisfies **C1**. Finally, it is well-known that the map between $(\mathbb{C}_{\text{ex}}, \|\cdot\|)$ and $(\mathbb{T}_{\text{f}}^{\text{ghp}}, \mathrm{d}_{\text{ghp}})$ that associates $\mathfrak{T}_f$ to an $f \in \mathbb{C}_{\text{ex}}$ is (Lipschitz) continuous [1, Proposition 3.3]). Hence, for any random variable $\mathcal{Z}$ with values in $\mathbb{C}_{\text{ex}}$, the corresponding (ghp-equivalence class of the) compact rooted measured real tree $\mathfrak{T}_{\mathcal{Z}}$ is a random variable with values in $\mathbb{T}_{\text{f}}^{\text{ghp}}$.

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. A similar construction has already been used in the specific context of the Brownian CRT, see for instance [6] and [2].

Let Γ be a rooted plane tree[2] with K vertices, where $K \geq 2$. We call Γ the *structural tree* of the recursive decomposition, and accordingly, the decomposition of a space (or tree) will involve K subparts. We consider Γ as a labelled tree upon labelling the root by 1 and the remaining nodes in the *depth-first order*[2]. We write $i \prec j$ if j lies in the subtree rooted at i. (We always have $i \prec i$.) For $i \in [K] := \{1, \ldots, K\}$, we set

$$\Gamma_i = \{i \leq j \leq K : i \prec j\}, \quad \text{and} \quad E_i = \{1 \leq j < i : j \prec i\}. \tag{2.2}$$

For $i \geq 2$, we denote by $\varpi_i = \max E_i$ the label of the parent of node i. Next, fix $\alpha > 0$, and $\mathbf{r}, \mathbf{s} \in \Sigma_K := \{(x_1, x_2, \ldots, x_K) \in (0,1)^K : x_1 + \ldots + x_K = 1\}$. We consider the following construction (see Section 1.4 in [2] for a related construction): given compact rooted measured metric spaces $(\mathcal{X}_i, d_i, \mu_i, \rho_i)$, $i \in [K]$, construct a compact rooted measured metric space $(\mathcal{X}, d, \mu, \rho)$ as follows:

i) Independently sample points $\eta_i \in \mathcal{X}_i, i \in [K]$, according to the probability measures μ_i;

ii) Let $\mathcal{X}^\circ = \sqcup_{i=1}^K \mathcal{X}_i$ denote the disjoint union of the $\mathcal{X}_i$, $i = 1, \ldots, K$; let $\sim_\circ$ be the smallest equivalence relation[3] on X° for which $\rho_i \sim_\circ \eta_{\varpi_i}$ for all $i = 2, \ldots, K$. We define $\mathcal{X}$ as the quotient $\mathcal{X}^\circ / \sim_\circ$ and write φ° for the canonical surjection from $\mathcal{X}^\circ$ onto $\mathcal{X}$.

iii) Let d° be the maximal pseudometric on $\mathcal{X}^\circ$ that is not greater than $r_i^\alpha d_i$ on $\mathcal{X}_i$, and for which $d^\circ(x, y) = 0$ if $x \sim_\circ y$; define d as the metric induced on $\mathcal{X}$ by d° under φ°; (see Section 3.1.3 of [15], and especially Lemma 3.1.23, Corollary 3.1.24 and Theorem 2.1.27 there which guarantee existence and uniqueness of d°.)

iv) Let $\mu^\circ(\cdot) = \sum_{i=1}^K s_i \mu_i(\cdot \cap \mathcal{X}_i)$ be the unique probability measure on $\mathcal{X}^\circ$ that is compatible with $s_i \mu_i$ when restricted to $\mathcal{X}_i$; define μ as the push-forward of μ° under φ°.

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

Note that, because of the need to sample $\eta_1, \ldots, \eta_K$, the ghp-equivalence class of the resulting space $\mathfrak{X} = (\mathcal{X}, d, \mu, \rho)$ is random. It is crucial to observe that it is a random variable (Lemma 26 in Appendix) whose 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 and continuous when considering $\mathcal{M}_1(\mathbb{K}^{\text{ghp}})$ equipped with the Prokhorov distance. (For a technical proof of continuity using the concept of correspondences and a coupling theorem due to Strassen [48], we refer to Lemma 27 in the appendix.) For probability measures τ on Σ_K^2 and $\aleph$ on $\mathbb{K}^{\text{ghp}}$, we define the intensity measure

$$\Psi(\aleph, \tau)(A) := \mathbf{E}\left[\psi(\mathfrak{X}_1, \ldots, \mathfrak{X}_K, \mathcal{R}, \mathcal{S})(A)\right], \quad A \subseteq \mathbb{K}^{\text{ghp}} \quad \text{measurable}, \tag{2.3}$$

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. Here, and throughout the document, we use $\mathfrak{L}(\cdot)$ to denote the distribution of a random variable.

[2] A rooted plane tree is a subset t of $\cup_{n \geq 0} (\mathbb{N} \setminus \{0\})^n$, 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 the order induced on t by the lexicographic order.

[3] Formally, $\sim_\circ$ can be defined as follows: first, set $\rho_i \sim_1 \eta_{\varpi_i}$ and $\eta_{\varpi_i} \sim_1 \rho_i$ for all $i = 2, \ldots, K$. Then, for $x, y \in \mathcal{X}^\circ$, set $x \sim_\circ y$ if and only if $x = y$ or there exist $k \geq 0$ and $x_1, \ldots, x_k \in \mathcal{X}^\circ$ such that $x \sim_1 x_1 \sim_1 x_2 \ldots \sim_1 x_k \sim_1 y$.

Given Γ, $\alpha > 0$ and a probability distribution τ on Σ_K^2, we are interested in non-trivial laws $\aleph \in \mathcal{M}_1(\mathbb{K}^{\mathrm{ghp}})$ satisfying

$$\aleph = \Psi(\aleph, \tau). \tag{2.4}$$

We refer to (2.4) 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, one can define Ψ relying on gp-isometry classes, and we will occasionally use Ψ in this sense.

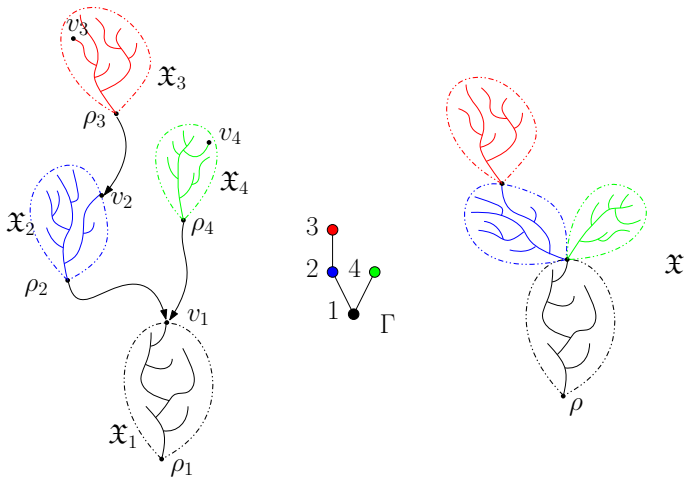

Figure 1: The construction of $\mathfrak{X}$ with law $\psi(\mathfrak{X}_1, \mathfrak{X}_2, \mathfrak{X}_3, \mathfrak{X}_4, \mathcal{R}, \mathcal{S})$ with the structural tree Γ shown in the middle. For simplicity, we have not rescaled any of the distances. An excursion point of view is depicted in Figure 2. (For the sake of representation, we have chosen the spaces to be tree-like.) Note that the roots ρ_2, ρ_4 are identified with the same point v_1 in the parent space $\mathcal{X}_1$.

Given Γ and τ, not every α is admissible for (2.4) to have non-trivial solutions. The parameter α is chosen as the unique value such that the height of an independent point sampled according to μ has finite mean, see the discussion of (2.6) below. Here, this condition can be expressed as follows. Recall Γ_i defined in (2.2), and define $\alpha > 0$ as the unique solution to

$$\mathbf{E}\left[\sum_{1\le i\le K} \mathcal{R}_i^\alpha \mathbf{1}_{\{J\in\Gamma_i\}}\right] = 1 \quad \text{with} \quad \mathbf{P}(J = j \,|\, \mathcal{R}, \mathcal{S}) = \mathcal{S}_j, \quad 1 \le j \le K. \tag{2.5}$$

Such an α always exists and lies in the interval $(0, 1)$ 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* (2.5).

The height of a random point. In a random compact rooted measured metric space $\mathfrak{X} = (\mathcal{X}, d, \mu, \rho)$, heights of points sampled according to μ play an important role. Recall E_i from (2.2). In our construction, with $\mathfrak{X}$ (or, rather its distribution) satisfying (2.4) and abbreviating $Y := d(\rho, \zeta)$ where ζ has distribution μ (given $\mathfrak{X}$), we have

$$Y \stackrel{d}{=} \sum_{i=1}^K \mathbf{1}_{\{J=i\}} \left[\mathcal{R}_i^\alpha Y^{(i)} + \sum_{j\in E_i} \mathcal{R}_j^\alpha Y^{(j)}\right] \stackrel{d}{=} \sum_{i=1}^K \beta_i^\alpha Y^{(i)}, \quad \beta_i = \mathcal{R}_i \mathbf{1}_{\{J\in\Gamma_i\}}, \tag{2.6}$$

and $Y^{(1)}, \dots, Y^{(K)}$ are distributed like Y, and $(J, \mathcal{R}, \mathcal{S}), Y^{(1)}, \dots, Y^{(K)}$ are independent with J as in (2.5). By [27, Theorem 2], the distribution of $d(\rho, \zeta)$ is uniquely determined by (2.6) in the space of probability distributions on $[0, \infty)$ up to a multiplicative constant. (To be more precise, in the notation of [27], we have[4] $\alpha = 1$ and $v'(1) < 0$ and therefore, as explained in the discussion following Theorem 2 in [27], the fixed points of (2.6) are parametrized by their means provided their existence.) [27, Theorem 2] further implies that, for any different choice of α, the random variable $d(\rho, \zeta)$ has either infinite mean, or, almost surely, $\mathcal{X} = \{\rho\}$.

Finally, we note that, from the last display, one can easily deduce that

$$\mathbf{P}\left(d(\rho, \zeta) > 0\right) \in \{0, 1\}. \tag{2.7}$$

Examples. **1)** The most celebrated example of a random measured real tree that satisfies a fixed point equation such as (2.4) is Aldous' Brownian CRT [3, 5]. Its recursive structure has been investigated in [6], where it has been proved that it satisfies a fixed point equation of the type (2.4), with $K = 3$, Γ is the tree on $\{1, 2, 3\}$ with 2 and 3 that are children of 1, $\mathcal{R} = \mathcal{S} = (\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3) \sim \text{Dirichlet}(1/2, 1/2, 1/2)$, and in this case, $\alpha = 1/2$.

2) An instance has also appeared in the context of recursive triangulations of the disk [14, 19]. There, $K = 2$ (and thus Γ is the tree on $\{1, 2\}$ where 2 is a child of 1), $\mathcal{S} = (\mathcal{S}_1, \mathcal{S}_2) \sim \text{Dirichlet}(2, 1)$, $\mathcal{R} = \mathcal{S}$ and α turns out to be given by $(\sqrt{17} - 3)/2$.

3) Another tree related to **2)** is given by $K = 2$, $\mathcal{S} = (\mathcal{S}_1, \mathcal{S}_2) \sim \text{Dirichlet}(2, 1)$, $\mathcal{R} \sim \text{Dirichlet}(1, 1)$ independent of $\mathcal{S}$, and for this case, one finds $\alpha = 1/3$. The tree has not been considered explicitly, but it appears in [14, 19] via one of its encoding processes.

Remark. Recently, Rembart and Winkel [45] have also studied recursive constructions of continuum random trees and the corresponding geometries using fixed points arguments. The decompositions they consider amount to seeing the trees as a forest of rescaled copies of random tree grafted on some segment (a so-called (random) *string of beads*) rather than on a copy of the tree itself. They typically involve an infinite number of pieces. Technically, Rembart and Winkel use a variant of the contraction method relying on Wasserstein distances on (a slightly modified version of) $\mathbb{K}_{\mathrm{f}}^{\mathrm{ghp}} \cap \mathbb{T}^{\mathrm{gh}}$ to verify uniqueness and attractiveness of fixed points. While both approaches rely on recursive decompositions and contraction arguments, the results seem largely disjoint. In particular many of the examples our results cover do not seem easily amenable to a spinal decomposition (e.g. lamination of the disk), and conversely, many examples covered by the results in [45] are not covered by our combinatorial decompositions.

2.4 Recursive decompositions: an excursion point of view

It is convenient to express the construction in Section 2.3 in terms of excursions. It is important to note that the excursion point of view developed here forces the spaces to be *real trees*, while in the previous section spaces were only required to be compact. However, we will see later on that this restriction is an important technical ingredient: we will prove that the support of the mass measure of any fixed point of the equations of interest is actually almost surely a real tree, see Theorem 1 *ii)*.

Let Γ be as in Section 2.3, $\mathbf{r}, \mathbf{s} \in \Sigma_K$ and assume for now that $\alpha > 0$ is arbitrary. We now describe a decomposition of the unit interval based on the tree Γ. 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 $\Gamma \setminus \partial\Gamma$. Set $L := 2K - |\partial\Gamma|$ and let $w_1, \dots, w_L \in \{1, \dots, K\}$ be the sequence of nodes visited by the depth first search process upon counting only the

[4] The α in the paper [27] should not be confused with the one defined in (2.5).

first and the *last* visit of a node. All nodes appear twice in this sequence except for leaves which appear once. For $i \in [K]$, let V_i be the set of times $1 \le j \le L$ with $w_j = i$. More formally, for $i \in [K]$, we can set $v_i = i + \#\{1 \le j < i : j \notin E_i \cup \partial\Gamma\}$ and obtain $V_i = \{v_i, v_i + 2|\Gamma_i| - \partial\Gamma_i - 1\}$. Observe that V_i contains a unique element if $i \in \partial\Gamma$, and two otherwise. For $\mathbf{u} = (u_1, \dots, u_K) \in (0,1)^K$, there is a unique decomposition of the unit interval into L half-open[5] intervals $I_1, \dots, I_L$ (following the natural order on $[0,1]$), such that,

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

Then, define

$$\Lambda_i = \bigcup_{k \in V_i} I_k \qquad \text{and} \qquad \varphi_i : \overline{\Lambda_i} \to [0,1] \tag{2.8}$$

as the unique function which is bijective, monotonically increasing and piecewise linear with constant slope. (Here, we $\overline{A}$ denotes the closure of a set A.) We can now define the continuous operator $\Phi : \mathbb{C}_{\text{ex}}^K \times \Sigma_K^2 \times (0,1)^K \to \mathbb{C}_{\text{ex}}$, such that $g = \Phi(f_1, \dots, f_K, \mathbf{r}, \mathbf{s}, \mathbf{u})$ is the unique excursion satisfying

$$g(x) - g(y) = r_{w_\ell}^\alpha \left[f_{w_\ell}(\varphi_{w_\ell}(x)) - f_{w_\ell}(\varphi_{w_\ell}(y)) \right], \tag{2.9}$$

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 (which may consist of two intervals). For an illustration see Figure 2 (and compare it with the corresponding version involving trees on Figure 1). By construction, for $f_1, \dots, f_K \in \mathbb{C}_{\text{ex}}, \mathbf{r}, \mathbf{s} \in \Sigma_K$ and Ξ uniformly distributed on $(0,1)^K$, we have

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

Thus, the distribution of the random compact rooted measured real tree $(\mathcal{T}_\mathcal{Z}, d_\mathcal{Z}, \mu_\mathcal{Z}, \rho_\mathcal{Z})$ satisfies (2.4) if

$$\mathcal{Z} \stackrel{d}{=} \Phi\big(\mathcal{Z}^{(1)}, \dots, \mathcal{Z}^{(K)}, \mathcal{R}, \mathcal{S}, \Xi\big), \tag{2.10}$$

where $\mathcal{Z}^{(1)}, \dots, \mathcal{Z}^{(K)}$ are independent copies of $\mathcal{Z}$, independent of $(\mathcal{R}, \mathcal{S}, \Xi)$, Ξ and $(\mathcal{R}, \mathcal{S})$ being independent and $\Xi = (\xi_1, \dots, \xi_K)$ being uniformly distributed on $(0,1)^K$; of course, in this case, α shall be chosen as in (2.5). The fixed point equation in (2.10) can be expressed alternatively as

$$\mathcal{Z}(\cdot) \stackrel{d}{=} \sum_{i=1}^K \mathcal{R}_i^\alpha \left[\mathbf{1}_{\Lambda_i}(\cdot) \mathcal{Z}^{(i)}(\varphi_i(\cdot)) + \sum_{j \in \Gamma_i \setminus \{i\}} \mathbf{1}_{\Lambda_j}(\cdot) \mathcal{Z}^{(i)}(\xi_i) \right].$$

Let us also note that, when asking for tree solutions to (2.4), 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 canonical external randomness.

Examples. **1)** An identity of the kind in (2.10) holds for the Brownian excursion $\mathbf{e}$, and is of course intimately related to the corresponding decomposition of the Brownian CRT. By [6, Corollary 3], we have

$$\mathbf{e} \stackrel{d}{=} \Phi\big(\mathbf{e}^{(1)}, \mathbf{e}^{(2)}, \mathbf{e}^{(3)}, \Delta, \Delta, \Xi\big), \tag{2.11}$$

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

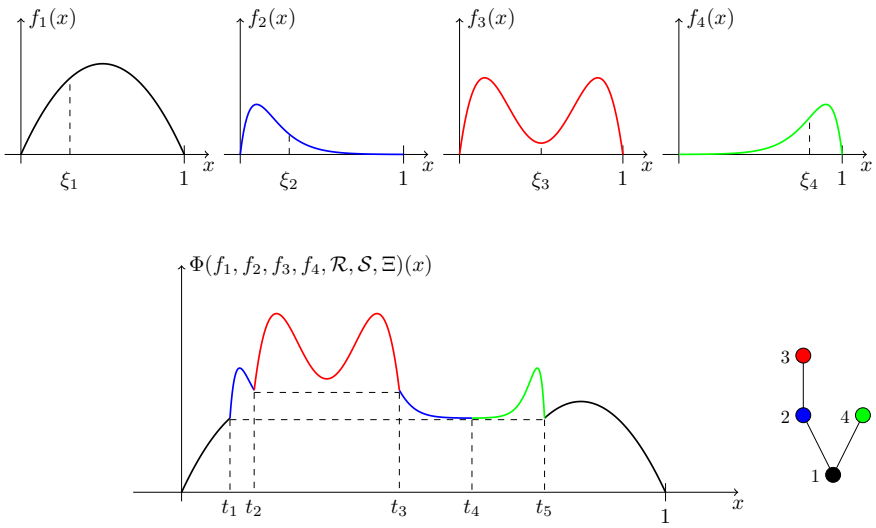

Figure 2: An example of the functional construction of Section 2.4: Here $K = 4$, $L = 6$, the structural tree Γ is the tree shown in the center right. The functions f_1, f_2, f_3, f_4 are composed into $\Phi(f_1, f_2, f_3, f_4, \mathcal{R}, \mathcal{S}, \Xi)$. In order to keep the focus on the structure 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. Observe that, in the tree encoded by f_1, the point corresponding to ξ_1 is not a leaf. The corresponding point of view using trees is depicted in Figure 1.

with conditions as in (2.10), where $\Delta \sim \text{Dirichlet}(1/2, 1/2, 1/2)$, $\alpha = 1/2$, and Γ is the tree of size three with nodes 2 and 3 attached to the root.

2) The functional version of the fixed point equation appearing in Example **2)** of Section 2.3 concerns a certain process $\mathscr{Z}$ that satisfies

$$\mathscr{Z} \stackrel{d}{=} \Phi\big(\mathscr{Z}^{(1)}, \mathscr{Z}^{(2)}, \Delta, \Delta, \Xi\big) \tag{2.12}$$

with conditions as in (2.10), where $\Delta \sim \text{Dirichlet}(2, 1)$ and $\alpha = (\sqrt{17} - 3)/2$.

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

$$\mathscr{H} \stackrel{d}{=} \Phi\big(\mathscr{H}^{(1)}, \mathscr{H}^{(2)}, \Delta, (W, 1 - W), \Xi\big) \tag{2.13}$$

with conditions as in (2.10), where $\Delta \sim \text{Dirichlet}(2, 1)$, W follows the uniform distribution on $[0, 1]$, Δ, W are independent and $\alpha = 1/3$.

2.5 Fractal properties of metric spaces

Let (S, d) be a metric space. (In this section, we do not assume the space to be compact.) For $\delta > 0$ and a non-empty relatively compact set B let $N_B(\delta)$ be the smallest number m such that there exist m open balls of radius δ covering B. We define the *lower Minkowski dimension* $\underline{\dim}_{\mathrm{m}}(B)$ and the *upper Minkowski dimension* $\overline{\dim}_{\mathrm{m}}(B)$ by

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

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 a set $A \subseteq S$ is defined using the family of (outer) Hausdorff measures $(H^s)_{s\geq 0}$ given by

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

(The map $H^s : \mathbb{K}^{\mathrm{gh}} \to [0,\infty]$ is measurable, see Lemma 29 in the appendix for a proof.) Here, for a set $U \subseteq S$, we let $|U| = \sup\{|x-y| : x, y \in U\}$. The Hausdorff dimension of $A \subseteq S$ 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. [29, Proposition 4.9] for a formulation in $\mathbb{R}^d$ which, together with its proof, applies analogously in any separable metric space). Let $B_r(x) := \{y \in S : d(x,y) < r\}$ denote the ball of radius r around x. Then, for a measurable set $A \subseteq S$, a finite measure ν on S with $\nu(A) > 0$ and $c > 0$, we have

$$\limsup_{r\to 0} \nu(B_r(x))/r^s \leq c \text{ for all } x \in A \quad \Rightarrow \quad \dim_{\mathrm{h}}(A) \geq s. \tag{2.14}$$

Lower and upper Minkowski dimension as well as Hausdorff dimension are invariant under bijective isometries. Furthermore, for a non-empty relatively compact 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 ω-Hölder continuous with $0 < \omega \leq 1$ (or Hölder with exponent ω) if there exists a constant $C > 0$ such that

$$|f(x) - f(y)| \leq C|x-y|^\omega, \quad 0 \leq x, y \leq 1.$$

For $f \in \mathbb{C}_{\mathrm{ex}}$, we let ω_f be the supremum of all Hölder exponents over excursions which are time-change equivalent to f.

Next, for $p > 0$, the p-variation $[f]_p(t), t \in [0,1]$, of a function $f \in \mathbb{C}$ is defined by

$$[f]_p(t) = \sup\left\{\sum_{i=0}^{n-1}|f(t_{i+1}) - f(t_i)|^p : n \geq 1 \text{ and } 0 \leq t_0 \leq \ldots \leq t_n \leq t\right\}.$$

Note that $[f]_p$ is monotonically increasing and, if $[f]_p(1) < \infty$, then $[f]_p \in \mathbb{C}$. It is easy to see that $\inf\{p > 0 : [f]_p(1) < \infty\} \leq \omega_f^{-1}$. Furthermore, if f is nowhere constant, the converse inequality follows from Theorem 3.1 in [16]. Combining this with Theorem 3.1 in Picard [44] shows that, for any $f \in \mathbb{C}_{\mathrm{ex}}$ which is nowhere constant, we have

$$\overline{\dim}_{\mathrm{m}}(\mathcal{T}_f) = \frac{1}{\omega_f}. \tag{2.15}$$

The last identity is used in Corollary 6.

3 Main results

3.1 Characterizations of solutions to (2.4) and (2.10)

Our first theorem clarifies the set of solutions to the fixed point equations (2.4) and (2.10). Recall that these involve the structural tree Γ, a probability distribution τ on Σ_K^2 and the value $\alpha \in (0,1)$ satisfying (2.5).

Theorem 1. Let $c > 0$. Then,

i) there exists a unique continuous excursion $\mathcal{Z}$ (in distribution) that satisfies (2.10) and $\mathbf{E}[\mathcal{Z}(\xi)] = c$ where ξ follows the uniform distribution on $[0,1]$ and $\mathcal{Z}, \xi$ are independent;
ii) for any $\mathbb{K}^{\mathrm{ghp}}$-valued r.v. $\mathfrak{X}$ satisfying (2.4) with $\mathbf{E}[d(\rho,\zeta)] = c$, where ζ is chosen according to μ (given $\mathfrak{X}$), the random tree $\mathfrak{T}_{\mathcal{Z}}$ encoded by $\mathcal{Z}$, and $(\mathrm{supp}(\mu), d, \mu, \rho)$ have the same distribution.

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

iii) $\mathcal{Z}(s) > 0$ for all $s \in (0,1)$;
iv) $\mathcal{Z}$ is nowhere monotonic.

Some comments are in order. First of all, as motivated by the formulation of point *ii)*, there can exist further solutions to (2.4) which are not almost surely continuum trees, see Proposition 3 below and the example discussed following Theorem 1.6 in [2]. Next, there does not exist any random compact rooted measured metric space with values in $\mathbb{K}^{\mathrm{ghp}}$ or $\mathbb{K}^{\mathrm{gp}}$ solving (2.4) for which one would have $\mathbf{E}[d(\rho,\zeta)] = \infty$. Similarly, any solution to (2.4) with values in $\mathbb{K}^{\mathrm{gp}}$ or $\mathbb{K}^{\mathrm{ghp}}_{\mathrm{f}}$ satisfying $\mathbf{E}[d(\rho,\zeta)] = 0$ must a.s. be the trivial space reduced to $\{\rho\}$.

Albenque and Goldschmidt [2] show that, in the case of the Brownian CRT, the fixed point in Example **1)** of Section 2.3 is attractive with respect to weak convergence on the space of probability measures on $\mathbb{K}^{\mathrm{gp}}$. They also raise the question whether this was true with respect to the Gromov–Hausdorff–Prokhorov distance. Our next result confirms that it is indeed the case, under certain moment conditions. (Note however that the results in [2] are not directly comparable to ours since the trees there are *unrooted*.)

In the following, we let $\phi_{\mathrm{ghp}} : \mathcal{M}_1(\mathbb{K}^{\mathrm{ghp}}) \to \mathcal{M}_1(\mathbb{K}^{\mathrm{ghp}})$ be the map that to $\aleph \in \mathcal{M}_1(\mathbb{K}^{\mathrm{ghp}})$ associates $\phi_{\mathrm{ghp}}(\aleph) = \Psi(\aleph, \tau)$, where τ is the given probability distribution on Σ_K^2 and Ψ is the map defined in (2.3). For $n \geq 1$, we write ϕ^n_{ghp} for the n-th iterate of ϕ_{ghp}. Analogously, we define ϕ_{gp} and $\phi^n_{\mathrm{gp}}, n \geq 1$ for probability measures on $\mathbb{K}^{\mathrm{gp}}$. Finally, for $\nu \in \mathcal{M}_1(\mathbb{K}^{\mathrm{ghp}})$ or $\mathcal{M}_1(\mathbb{K}^{\mathrm{gp}})$, we write $\mathbf{E}_\nu$ for the expectation with respect to spaces sampled from ν.

Theorem 2. Fix $c > 0$ and let $\mathcal{Z}$ be the unique continuous excursion (in distribution) satisfying (2.10) with $\mathbf{E}[\mathcal{Z}(\xi)] = c$ from Theorem 1. Then, we have the following two statements:

i) if $\nu \in \mathcal{M}_1(\mathbb{K}^{\mathrm{gp}})$ with $\mathbf{E}_\nu[d(\rho,\zeta)] = c$, then the sequence of distributions $(\phi^n_{\mathrm{gp}}(\nu))_{n\geq 1}$ converges weakly to the law of $\mathfrak{T}_{\mathcal{Z}}$.
ii) if $\nu \in \mathcal{M}_1(\mathbb{T}^{\mathrm{ghp}}_{\mathrm{f}})$ with $\mathbf{E}_\nu[d(\rho,\zeta)] = c$ and $\mathbf{E}_\nu[\|\mathfrak{T}\|^{\mathfrak{m}}] < \infty$, where

$$\mathfrak{m} = 1 + \lfloor 1/\alpha \rfloor = \min\left\{ m \in \mathbb{N} : \sum_{i=1}^{K} \mathbf{E}\left[\mathcal{R}_i^{m\alpha}\right] < 1 \right\}, \tag{3.1}$$

then the sequence of distributions $(\phi^n_{\mathrm{ghp}}(\nu))_{n\geq 1}$ converges weakly to the law of $\mathfrak{T}_{\mathcal{Z}}$.

Note that the assumptions in Theorem 2 *i)* are rather weak as no conditions on the probability distribution ν are imposed apart from $\mathbf{E}_\nu[d(\rho,\zeta)] = c$. For instance, if ν charges only two-point metric spaces $\{\rho, x\}$ such that $d(\rho, x) = c$, the theorem applies. 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 Γ and let η be a probability distribution on Σ_K. Then, there exists a distribution τ on Σ_K^2 such that $\tau(\Sigma_K \times \cdot\,) = \eta(\cdot)$, and furthermore, with α as in (2.5),

i) there exist infinitely many mutually singular fixed points of (2.4) on $\mathbb{K}^{\mathrm{ghp}}$ including some that are almost surely not real trees, such that, writing $\nu \in \mathcal{M}_1(\mathbb{K}^{\mathrm{ghp}})$ for such a distribution, and $\mathfrak{X} = (\mathcal{X}, d, \mu, \rho)$ distributed according to ν, we have $\mathbf{E}_\nu[d(\rho,\zeta)] = c$ and $\mathbf{E}[\|\mathfrak{X}\|^{1/\alpha}] = \infty$;

ii) there exists $\nu \in \mathcal{M}_1(\mathbb{K}^{\mathrm{ghp}})$ concentrated on $\mathbb{T}_{\mathrm{f}}^{\mathrm{ghp}}$ such that $\mathbf{E}_\nu[d(\rho,\zeta)] = c$, $\mathbf{E}_\nu[\|\mathfrak{T}\|^{1/\alpha}] = \infty$ and $(\phi_{\mathrm{ghp}}^n(\nu))_{n\geq 1}$ does not converge weakly to the law of $\mathfrak{T}_{\mathcal{Z}}$ with $\mathcal{Z}$ as in Theorem 2.

Let us comment on *ii)* above: According to Theorem 2 *i)*, under these conditions, one has convergence in the sense of Gromov–Prokhorov to $\mathfrak{T}_{\mathcal{Z}}$, where $\mathcal{Z}$ is the random excursion of Theorem 2; however, as one iterates ϕ_{ghp}, the support of the mass measure converges to $\mathfrak{T}_{\mathcal{Z}}$, but some of the branches of the tree are drained of their mass without becoming shorter (see the proof on page 23 for details).

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 (2.4). 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, the fixed point equation (2.4) should imply that we approximately have

$$N_{\mathcal{T}}(\delta) \approx \sum_{i=1}^{K} N_{\mathcal{R}_i^\alpha \mathcal{T}_i}(\delta) = \sum_{i=1}^{K} N_{\mathcal{T}_i}(\mathcal{R}_i^{-\alpha}\delta). \tag{3.2}$$

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 which 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 independent of $\mathcal{R}$ and uniformly chosen among $1, \ldots, K$.

Theorem 4 (Upper bound on $\overline{\dim}_{\mathrm{m}}$)**.** Assume that $\mathfrak{T} = (\mathcal{T}, d, \mu, \rho)$ satisfies (2.4) with values in $\mathbb{T}_{\mathrm{f}}^{\mathrm{ghp}}$. Moreover, assume that $\overline{\mathcal{R}}$ admits a density on $(0,1)$ (although an atom at 0 is allowed). Then, almost surely,

$$\overline{\dim}_{\mathrm{m}}(\mathcal{T}) \leq 1/\alpha.$$

Theorem 5 (Lower bound on $\dim_{\mathrm{h}}$)**.** Suppose that $\mathfrak{T} = (\mathcal{T}, d, \mu, \rho)$ satisfies (2.4) with values in $\mathbb{T}_{\mathrm{f}}^{\mathrm{ghp}}$, that $\mathbf{P}(\mathcal{T} \neq \{\rho\}) > 0$, and, for some $\delta > 0$, $\mathbf{E}[\overline{\mathcal{R}}^{-\delta}] < \infty$. Then

$$\dim_{\mathrm{h}}(\mathcal{T}) \geq 1/\alpha.$$

In the next result, recall that two excursions f and g are said to be equivalent if $f = g \circ \phi$ for a function $\phi : [0,1] \to [0,1]$ that is continuous and strictly increasing.

Corollary 6 (Optimal Hölder exponents)**.** Suppose that $\mathcal{Z}$ satisfies (2.10), that $\mathbf{P}(\|\mathcal{Z}\| > 0) > 0$ and that the conditions of Theorems 4 and 5 are satisfied. Then, almost surely:

(a) $\dim_{\mathrm{h}}(\mathcal{T}_{\mathcal{Z}}) = \underline{\dim}_{\mathrm{m}}(\mathcal{T}_{\mathcal{Z}}) = \overline{\dim}_{\mathrm{m}}(\mathcal{T}_{\mathcal{Z}}) = \alpha^{-1}$;

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

(c) for any $\gamma > \alpha$ and any distribution $\tilde{\nu}$ on $\mathbb{C}_{\mathrm{ex}}$ such that, for $\mathfrak{L}(\tilde{\mathcal{Z}}) = \tilde{\nu}$, $\mathcal{T}_{\tilde{\mathcal{Z}}}$ and $\mathcal{T}_{\mathcal{Z}}$ have the same distribution, $\tilde{\mathcal{Z}}$ has γ-Hölder continuous paths with probability zero.

Finally, we have the following results about the degrees in $\mathfrak{T}_{\mathcal{Z}}$. Recall that the degree of a point in a real tree is defined as the number of connected components in which the space decomposes upon removal of the point. Let $\mathscr{D}(\mathcal{T}_{\mathcal{Z}})$ be the (random) set of degrees of $\mathcal{T}_{\mathcal{Z}}$. Let also $\mathscr{D}(\Gamma) = \{1 + \#\{j : \varpi_j = i\} : 1 \leq i \leq K\}$; $\mathscr{D}(\Gamma)$ is the set of degrees in the tree obtained from Γ by connecting an additional node to its root. Observe that $1 \in \mathscr{D}(\Gamma)$, but that it is possible that 2 is not an element of $\mathscr{D}(\Gamma)$.

Proposition 7. Let $0 < c < \infty$ and $\mathcal{Z}$ be the unique solution (in distribution) of (2.10) in Theorem 1 with $\mathbf{E}[\mathcal{Z}(\xi)] = c$ and consider $\mathfrak{T}_{\mathcal{Z}} = (\mathcal{T}_{\mathcal{Z}}, d_{\mathcal{Z}}, \mu_{\mathcal{Z}}, \rho_{\mathcal{Z}})$. Then, almost surely,

i) the root $\rho_{\mathcal{Z}}$ is a leaf;
ii) a point sampled from $\mu_{\mathcal{Z}}$ is a leaf;
iii) $\mu_{\mathcal{Z}}$ has no atoms;
iv) the set $\mathscr{D}(\mathcal{T}_{\mathcal{Z}})$ of degrees of points in $\mathcal{T}_{\mathcal{Z}}$ is fully determined by $\mathscr{D}(\Gamma)$: we have

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

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 (on page 9). 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 $\varpi_3, \ldots, \varpi_K$ of the nodes $i = 3, \ldots, K$. Recall that continuum real tree refers to an element of $\mathbb{T}^{\mathrm{gp}}$ or $\mathbb{T}_{\mathrm{f}}^{\mathrm{ghp}}$.

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

(a) The Brownian CRT $(\mathcal{T}_{\mathbf{e}}, d_{\mathbf{e}}, \mu_{\mathbf{e}}, \pi_{\mathbf{e}}(0))$ is the unique continuum real tree satisfying (2.4) with $K = 3$, $\varpi_3 = 1$, $\alpha = 1/2$, $\mathcal{S} = \mathcal{R} = \Delta$ as in (2.11).
(b) The tree $(\mathcal{T}_{\mathscr{Z}}, d_{\mathscr{Z}}, \mu_{\mathscr{Z}}, \pi_{\mathscr{Z}}(0))$ is the unique continuum real tree satisfying (2.4) with $K = 2$, $\mathcal{S} = \mathcal{R} = \Delta$, as in (2.12) 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 continuum real tree satisfying (2.4) with $K = 2$, $\alpha = 1/3$, $\mathcal{S} = \Delta$, $\mathcal{R} = (W, 1 - W)$ with Δ, W as in (2.13).

Furthermore, these uniqueness results also hold in the spaces $\mathbb{K}_{\mathrm{f}}^{\mathrm{ghp}}$ and $\mathbb{K}^{\mathrm{gp}}$.

The assertion concerning the Brownian CRT (in the unrooted set-up and using a slightly different definition of continuum trees) in Corollary 8 is the main result in [2]. Similarly, the claim for the process $\mathscr{Z}$ in Corollary 9 has already been given in [14] in the space $\mathbb{C}$ 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 (2.11).
(b) The process $\mathscr{Z}$ is the unique continuous excursion satisfying (2.12).
(c) The process $\mathscr{H}$ is the unique continuous excursion satisfying (2.13).

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 [4] (see also [25] for the more general case of Lévy trees) and the Minkowski dimension of $\mathcal{T}_{\mathscr{Z}}$ in [14].

Corollary 10. Almost surely, for the processes $\mathbf{e}$, $\mathscr{Z}$ and $\mathscr{H}$ satisfying (2.11), (2.12) and (2.13),

(a) $\dim_{\mathrm{m}}(\mathcal{T}_{\mathbf{e}}) = \dim_{\mathrm{h}}(\mathcal{T}_{\mathbf{e}}) = 2$.
(b) $\dim_{\mathrm{m}}(\mathcal{T}_{\mathscr{Z}}) = \dim_{\mathrm{h}}(\mathcal{T}_{\mathscr{Z}}) = 2/(\sqrt{17} - 3)$,
(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 (2.10). For fixed point equations describing real-valued distributions, this idea is classical, see, e.g. [8,

Section 2.3]. For the Brownian CRT $\mathcal{T}_{\mathrm{e}}$, it was used in [18] as well as, in [2].

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,

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

be two independent sets of independent random variables where each $(\mathcal{R}^\vartheta, \mathcal{S}^\vartheta)$ has distribution τ 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

$$\mathcal{V}(\vartheta) = \prod_{e\in\pi_\vartheta} \mathcal{R}(e) \qquad \text{and} \qquad \mathcal{L}(\vartheta) = \prod_{e\in\pi_\vartheta} \mathcal{S}(e), \tag{3.3}$$

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 $c > 0$, one can construct a family of random excursions $\{\mathcal{Z}^\vartheta : \vartheta \in \Theta\}$, such that

$$\mathcal{Z}^\vartheta = \Phi\big(\mathcal{Z}^{\vartheta 1}, \dots, \mathcal{Z}^{\vartheta K}, \mathcal{R}^\vartheta, \mathcal{S}^\vartheta, \Xi^\vartheta\big),$$

where, for all $\vartheta \in \Theta$, the distribution of $\mathcal{Z}^\vartheta$ does not depend on ϑ, $\mathcal{Z}^\vartheta$ is measurable with respect to $\{(\mathcal{R}^{\vartheta\sigma}, \mathcal{S}^{\vartheta\sigma}, \Xi^{\vartheta\sigma}) : \sigma \in \Theta\}$, and $\mathbf{E}[\mathcal{Z}^\vartheta(\xi)] = c$.

It should be clear that, for any $\vartheta \in \Theta$ and $n \geq 0$, the unit interval is decomposed in half-open intervals such that, for $\sigma \in \Theta_n$, $\mathcal{Z}^{\vartheta\sigma}$ multiplied by $(\mathcal{V}(\vartheta\sigma)/\mathcal{V}(\vartheta))^\alpha$ governs the behaviour of the process $\mathcal{Z}^\vartheta$ on a set Λ_σ^ϑ 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 $\vartheta \in \Theta$ set $\Lambda_\emptyset^\vartheta = [0,1]$ and $j \in [K]$, let the set Λ_j^ϑ and the function φ_j^ϑ be defined as Λ_j and φ_j in (2.8) using the vector $(\mathcal{S}^\vartheta, \Xi^\vartheta)$. Then, given Λ_σ^ϑ and φ_σ^ϑ for $\sigma \in \Theta_n, n \geq 1$, for $j \in [K]$, let

$$\Lambda_{\sigma j}^\vartheta = \big(\varphi_\sigma^\vartheta\big)^{-1}(\Lambda_j^{\vartheta\sigma}), \tag{3.4}$$

and let $\varphi_{\sigma j}^\vartheta$ be the unique piece-wise linear bijective and increasing function with constant slope mapping $\overline{\Lambda_{\sigma j}^\vartheta}$ onto $[0,1]$. For $n \geq 0$, Λ_σ^ϑ is the disjoint union of the sets $\Lambda_{\sigma\omega}^\vartheta, \omega \in \Theta_n$. In particular, $\Lambda_\omega^\vartheta, \omega \in \Theta_n$ is a partition of the unit interval. Throughout the paper, we write $\Lambda_\vartheta = \Lambda_\vartheta^\emptyset$, $\vartheta \in \Theta$ and $\mathcal{Z} = \mathcal{Z}^\emptyset$ for the quantities at the root of Θ. (No confusion should arise in the notation as $\mathcal{Z}^\emptyset$ is indeed the process from Theorem 1.)

3.5 Organization of the proofs

The remainder of the paper is organized as follows: In Section 4, we prove Theorems 1 *i), ii)*, Theorem 2 and Proposition 3. Section 4.1 is devoted to showing that, up to a scaling constant, there exists at most one solution to (2.4) in $\mathbb{K}_{\mathrm{f}}^{\mathrm{ghp}}$ (or $\mathbb{K}^{\mathrm{gp}}$). In Section 4.2 this solution is constructed together with a unique continuous excursion satisfying (2.10). 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.

Section 5 contains the proofs to the remaining statements presented in Section 3. In Section 5.1, we start with the verification of Theorem 4. In Section 5.2, we give the proof of the lower bound on the Hausdorff dimension of Theorem 5. In Section 5.3, we discuss the proofs of Proposition 7 and the statements *iii)* and *iv)* of Theorem 1. Corollary 6 is discussed in Section 5.4.

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 [14, 19].

4 Proofs of the existence and uniqueness results

4.1 Uniqueness of the encoding function

The proof of Theorem 1 consists of two steps. First, in Proposition 12, we show that, in distribution, there exists at most one compact rooted measured metric space with full support satisfying (2.4). Second, as in [14], we use a variant of the functional contraction method developed in [41] to construct a solution to (2.10) whose supremum has finite moments of all orders. We indicate how to obtain the statements *i)* and *ii)* in Theorem 1 from the next two results right after proving Proposition 13.

For $n \in \mathbb{N} = \{0, 1, 2, \ldots\}$, let $\mathbb{M}_n = \{(m_{ij})_{0 \le i,j \le n} : m_{ij} = m_{ji} \ge 0, m_{ii} = 0\}$ be the set of symmetric $(n+1)$ by $(n+1)$ matrices with non-negative entries and zeros on the diagonal. We also write $\mathbb{M}_{\mathbb{N}}$ for the set of infinite dimensional matrices satisfying these properties. We always endow this set with the product topology and the corresponding Borel σ-field. For a fixed compact rooted measured metric space $\mathfrak{X} = (\mathcal{X}, d, \mu, \rho)$, let $(\zeta_i)_{i \ge 1}$ be independent and identically distributed (i.i.d.) test points with distribution μ on $\mathcal{X}$, and 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 distribution* $\nu^{\mathfrak{X}} \in \mathcal{M}_1(\mathbb{M}_{\mathbb{N}})$ for elements $\mathfrak{X}$ of $\mathbb{K}^{\mathrm{ghp}}$ (or $\mathbb{K}^{\mathrm{gp}}$)[6]. For a probability distribution $\kappa \in \mathcal{M}_1(\mathbb{K}^{\mathrm{ghp}})$ (or $\mathcal{M}_1(\mathbb{K}^{\mathrm{gp}})$), we also define the probability measure

$$\nu^{\kappa}(A) := \int \nu^{x}(A) d\kappa(x), \quad A \subseteq \mathbb{M}_{\mathbb{N}} \quad \text{measurable.}$$

The importance of $\nu^{\mathfrak{X}}$ and ν^{κ} 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 [31, Section 3 1/2] or Vershik [50, Theorem 4]. Statements *ii)* and *iii)* (again in the gp case and for unrooted structures) are covered by Corollary 3.1 in [30] (see also [37, 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) For fixed $\mathfrak{X}_1, \mathfrak{X}_2 \in \mathbb{K}^{\mathrm{ghp}}_{\mathrm{f}}$ (or $\mathbb{K}^{\mathrm{gp}}$) we have $\nu^{\mathfrak{X}_1} = \nu^{\mathfrak{X}_2}$ if and only if $\mathfrak{X}_1 = \mathfrak{X}_2$.

ii) For $\kappa_1, \kappa_2 \in \mathcal{M}_1(\mathbb{K}^{\mathrm{ghp}}_{\mathrm{f}})$ (or $\mathcal{M}_1(\mathbb{K}^{\mathrm{gp}})$) we have $\nu^{\kappa_1} = \nu^{\kappa_2}$ if and only if $\kappa_1 = \kappa_2$.

iii) For probability distributions κ and κ_n, $n \ge 1$, on $\mathbb{K}^{\mathrm{gp}}$, we have $\kappa_n \to \kappa$ weakly if and only if $\nu^{\kappa_n} \to \nu^{\kappa}$ weakly.

With Proposition 11 at hand, we can now compare the distance matrix distributions of two solutions of the fixed point equation (2.4). To this end, for a random variable $\mathfrak{X}$ with values in $\mathbb{K}^{\mathrm{ghp}}$ (or $\mathbb{K}^{\mathrm{gp}}$), set $\mathbf{E}[\nu^{\mathfrak{X}}] = \nu^{\mathfrak{L}(\mathfrak{X})}$, that is, $\mathbf{E}[\nu^{\mathfrak{X}}](A) = \mathbf{E}[\nu^{\mathfrak{X}}(A)]$ for a measurable set A. Further, for $n \ge 0$ and fixed $\mathfrak{X} \in \mathbb{K}^{\mathrm{ghp}}$ (or $\mathbb{K}^{\mathrm{gp}}$), we write $\nu_n^{\mathfrak{X}} \in \mathcal{M}_1(\mathbb{M}_n)$ for the distribution of the distance matrix induced by the first $n+1$ points $\zeta_0, \zeta_1, \ldots, \zeta_n$. The quantities ν_n^{κ} for a probability distribution $\kappa \in \mathcal{M}_1(\mathbb{K}^{\mathrm{ghp}})$ (or $\mathcal{M}_1(\mathbb{K}^{\mathrm{gp}})$) and $\mathbf{E}[\nu_n^{\mathfrak{X}}]$ for a random variable $\mathfrak{X}$ shall be defined correspondingly.

Proposition 12. Let $\mathfrak{X} = (\mathcal{X}, d, \mu, \rho)$ and $\mathfrak{U} = (\mathcal{U}, d', \mu', \rho')$ be $\mathbb{K}^{\mathrm{ghp}}_{\mathrm{f}}$ (or $\mathbb{K}^{\mathrm{gp}}$)-valued r.v.'s satisfying (2.4) (in distribution) and $\mathbf{E}[d(\rho, \zeta)] = \mathbf{E}[d'(\rho', \zeta')]$ where ζ (resp. ζ') is chosen on X (resp. $\mathcal{U}$) according to μ (resp. μ'). Then, $\mathbf{E}[\nu^{\mathfrak{X}}] = \mathbf{E}[\nu^{\mathfrak{U}}]$, hence $\mathfrak{L}(\mathfrak{X}) = \mathfrak{L}(\mathfrak{U})$.

[6] The map $\mathfrak{X} \to \nu^{\mathfrak{X}}$ is continuous for the Gromov–Prokhorov topology (via the equivalence with the so-called Gromov-weak topology; see [22, 30] for details).

Proof. For $n \geq 1$, let D_n (resp. D'_n) be a random variable on $\mathbb{M}_n$ with distribution $\mathbf{E}[\nu_n^{\mathfrak{X}}]$ (resp. $\mathbf{E}[\nu_n^{\mathfrak{U}}]$). Both D_1 and D'_1 satisfy fixed point equation (2.6). As elaborated in the discussion of (2.6) it follows from the results in [27] that D_1 and D'_1 are identically distributed since both random variables have mean c. Our aim is to show by induction on n that, for all $n \geq 2$, both D_n and D'_n satisfy the same stochastic fixed point equation known to admit at most one solution. Taking that for granted for now, it follows immediately that $\mathbf{E}[\nu_n^{\mathfrak{X}}] = \mathbf{E}[\nu_n^{\mathfrak{U}}]$ for all $n \geq 1$; Proposition 11 *ii)* then proves the assertion.

Let $A \subseteq \mathbb{M}_n$ be measurable. With $\tau = \mathfrak{L}((\mathcal{R}, \mathcal{S}))$, $\aleph = \mathfrak{L}(\mathfrak{X})$ and $\mathfrak{X}_1, \ldots, \mathfrak{X}_K$ i.i.d. with distribution $\aleph$, also independent of $(\mathcal{R}, \mathcal{S})$, we deduce from (2.4) that

$$\begin{aligned}\mathbf{P}(D_n \in A) &= \nu_n^{\aleph}(A) = \nu_n^{\Psi(\aleph,\tau)}(A) = \mathbf{E}[\nu_n^{\psi(\mathfrak{X}_1,\ldots,\mathfrak{X}_K,\mathcal{R},\mathcal{S})}(A)] \\ &= \int_{\Sigma_K^2} d\tau(\mathbf{r}, \mathbf{s}) \int_{(\mathbb{K}^{\mathrm{ghp}})^K} d\aleph^{\otimes K}(x_1, \ldots, x_K) \mathbf{E}[\nu_n^{\mathfrak{X}^*(x_1,\ldots,x_K,\mathbf{r},\mathbf{s})}](A), \end{aligned} \tag{4.1}$$

where $\mathfrak{X}^*(x_1, \ldots, x_k, \mathbf{r}, \mathbf{s})$ has distribution $\psi(x_1, \ldots, x_k, \mathbf{r}, \mathbf{s})$. (Here, we could apply Fubini's theorem thanks to the product measurability of the map ψ proved in the supplementary material.) We now keep $\mathbf{r} = (r_1, \ldots, r_K)$, $\mathbf{s} = (s_1, \ldots, s_K)$ and $x_1, \ldots, x_K$ fixed. Let $\bar{x}_1, \ldots, \bar{x}_K$ be arbitrary representatives of $x_1, \ldots, x_K$ and write $\bar{x}_i = (\bar{x}_i, \bar{d}_i, \bar{\mu}_i, \bar{\rho}_i)$. Let $\eta_i, i \in [K]$ be independent points on $\bar{x}_i$ with distribution $\bar{\mu}_i$. Let $x_* = (x_*, d_*, \mu_*, \rho_*)$ be the space constructed with the help of $\Gamma, \alpha, \mathbf{r}, \mathbf{s}, \bar{x}_1, \ldots, \bar{x}_K, \eta_1, \ldots, \eta_K$ following the steps ii) - iv) on page 6. To sample independent test points $(\zeta_\ell)_{1 \leq \ell \leq n}$, on x_* according to μ_*, we consider a family of independent random variables $\{\theta_{i,j} : i \in [K], j = 1, \ldots, n\}$ which is independent of the glue points η_i, $i \in [K]$, where each $\theta_{i,j}$ takes values in $\bar{x}_i$ and has distribution $\bar{\mu}_i$. Let also $\mathbf{J} = (J_1, \ldots, J_n)$ be a vector of i.i.d. random variables with values in $[K]$, independent of the remaining quantities, with $\mathbf{P}(J_\ell = j) = s_j$ for $j \in [K]$. Define $\zeta_\ell = \varphi^\circ(\theta_{J_\ell,\ell})$, where φ° is introduced in step *ii)* of the construction on page 6.

Then $(\zeta_i)_{1 \leq i \leq n}$ is a family of i.i.d. points with distribution μ_*. Set $\zeta_0 := \rho_*$. By construction, the matrix $W_n = (d(\zeta_i, \zeta_j))_{0 \leq i,j \leq n}$ has distribution $\mathbf{E}[\nu_n^{\mathfrak{X}^*(x_1,\ldots,x_k,\mathbf{r},\mathbf{s})}]$. (In particular, its distribution does not depend on the choice of the representatives $\bar{x}_1, \ldots, \bar{x}_K$.)

We now decompose W_n by looking in which of the subspaces (or, more precisely, their images under φ°) the random points $\zeta_1, \ldots, \zeta_n$ fall. In each space, the trace of the paths induce a rescaled distance matrix with a potentially different number of points which maybe the original points, the root of the subspace or glue points. To this end, first recall that the root $\rho_* = \zeta_0$ is the (image of the) root $\bar{\rho}_1$ of $\bar{x}_1$. Set $j_0 := 1$ and, for $\mathbf{j} = (j_1, \ldots, j_n) \in [K]^n$ and $i \in [K]$ define $L_i^{\mathbf{j}} = \{0 \leq \ell \leq n : j_\ell = i\}$. Then, set $\ell_1 = \#L_1^{\mathbf{j}} - 1$ and, for $2 \leq i \leq K$ define $\ell_i = \#L_i^{\mathbf{j}}$. Further, for $i \in [K]$, let $\ell_i^* = \ell_i + 1$ if there exists some $1 \leq \ell \leq n$ such that $j_\ell \in \Gamma_i \setminus \{i\}$ and $\ell_i^* = \ell_i$ otherwise. (ℓ_i is the number of test points falling in (the image of) $\bar{x}_i$, and ℓ_i^* accounts for the glue point η_i which plays a role if there is some segment $[\![\zeta_p, \zeta_q]\!]$ containing its image under φ°.)

For $\mathbf{j} \in [K]^n$ let $\mathcal{E}^{\mathbf{j}}$ be the event that $J_\ell = j_\ell$ for all $\ell \in [n]$. For integers $p \geq 1, i \in [K]$, we write $Y_{x_i}^{(p)}$ for a generic random variable with distribution $\nu_p^{x_i}$. Now note that, on $\mathcal{E}^{\mathbf{j}}$, the distance matrix W_n has the same distribution as a linear combination of deterministic linear operators evaluated at independent copies of $Y_{x_i}^{(\ell_i^*)}, i \in [K]$. To this end, for $\ell = 1, 2, \ldots, n$, let $\mathrm{rk}(\ell) := \#\{p \in L_{j_\ell}^{\mathbf{j}} : 1 \leq p \leq \ell\}$ be the rank of ℓ in the set $L_{j_\ell}^{\mathbf{j}}$, and set $\mathrm{rk}(0) = 0$. Then, for $i \in [K]$, we define operators $G_i^{(\mathbf{j})} : \mathbb{M}_{\ell_i^*} \to \mathbb{M}_n$ in two steps. Let $A \in \mathbb{M}_{\ell_i^*}$ and $1 \leq p, q \leq n$. First, set

$$G_i^{(\mathbf{j})}(A)_{p,q} = \begin{cases} A_{\mathrm{rk}(p),\mathrm{rk}(q)} & \text{if } p, q \in L_i^{\mathbf{j}}, \\ A_{0,\mathrm{rk}(p)} & \text{if } p \in L_i^{\mathbf{j}}, j_q \notin \Gamma_i, \\ A_{0,\mathrm{rk}(q)} & \text{if } q \in L_i^{\mathbf{j}}, j_p \notin \Gamma_i. \end{cases} \tag{4.2}$$

This does not define all the entries of $G_i^{(\mathbf{j})}(A)$. For the remaining entries: if $\ell_i^* = \ell_i$, then $G_i^{(\mathbf{j})}(A)_{p,q} = 0$ for all p, q which are not covered in one of the cases in (4.2); or else we have $\ell_i^* = \ell_i + 1$, and then

$$G_i^{(\mathbf{j})}(A)_{p,q} = \begin{cases} A_{\mathrm{rk}(p),\ell_i^*} & \text{if } p \in L_i^{\mathbf{j}}, j_q \in \Gamma_i \setminus \{i\}, \\ A_{\mathrm{rk}(q),\ell_i^*} & \text{if } q \in L_i^{\mathbf{j}}, j_p \in \Gamma_i \setminus \{i\}, \\ A_{0,\ell_i^*} & \text{if } j_p \in \Gamma_i \setminus \{i\}, j_q \notin \Gamma_i \text{ or } j_q \in \Gamma_i \setminus \{i\}, j_p \notin \Gamma_i, \end{cases} \tag{4.3}$$

and $G_i^{(\mathbf{j})}(A)_{p,q} = 0$ for all p, q which are covered neither in (4.2) nor in (4.3). Note that, if $\ell_i^* = 0$, we have $G_i^{(\mathbf{j})} \equiv 0$.

The following observation is the crucial ingredient of the proof: When conditioning on $\mathcal{E}^{\mathbf{j}}$, as the points $\theta_{k,\ell}$, $\ell = 1, \dots, n$, and η_k, are all independent and distributed on $\bar{x}_k$ according to $\bar{\mu}_k$, we may think of η_k as an additional test point on $\bar{x}_k$. Thus, on $\mathcal{E}^{\mathbf{j}}$, we obtain the following distributional equality

$$W_n \stackrel{d}{=} \sum_{i=1}^{K} r_i^\alpha \cdot G_i^{(\mathbf{j})}(Y_{x_i}^{(\ell_i^*)}), \tag{4.4}$$

where the $Y_{x_i}^{(\ell_i^*)}$, $i \in [K]$, are independent. (For the sake of convenience, we agree to set $Y_{x_i}^{(0)}$ to be the matrix containing a single entry which is 0.)

The cases where $\ell_i^* = n$ for some $i \in [K]$ need to be considered in detail: indeed, they are the cases that yield $(n+1) \times (n+1)$ distance matrices from the constituant subspaces, and are thus crucial to the fixed point argument. To this end, we define the following subsets of $[K]^n$: $C_i = \{(i, \dots, i)\}$ and

$$C_i^* = \bigcup_{k=1}^{K} \Big\{ (j_1, \dots, j_n) : j_\ell = i \text{ for all } \ell \neq k, j_k \in \Gamma_i \setminus \{i\} \Big\},$$

as well as $C = C_1 \cup \dots \cup C_K$ and $C^* = C_1^* \cup \dots \cup C_K^*$. Then, for $i \in [K]$ we have $\ell_i^* = n$ if and only if $\mathbf{j} \in C_i \cup C_i^*$. In the following, we distinguish these two cases. Recall the definition of E_i from (2.2).

a) If $\mathbf{j} \in C_i$, the operator $G_i^{(\mathbf{j})}$ is the identity and, as observed previously, $G_k^{(\mathbf{j})} \equiv 0$ if $k \notin E_i \cup \{i\}$ since $\ell_k^* = 0$. For $k \in E_i$, however, we have $\ell_k^* = 1$ and $G_k^{(\mathbf{j})} : \mathbb{M}_1 \to \mathbb{M}_n$ is defined in (4.2) and (4.3). Thus, in this case, (4.4) can be written as

$$W_n \stackrel{d}{=} r_i^\alpha \cdot Y_{x_i}^{(n)} + \sum_{k=1, k\neq i}^{K} r_k^\alpha \cdot G_k^{(\mathbf{j})}(Y_{x_k}^{(\ell_k^*)}), \tag{4.5}$$

where $Y_{x_i}^{(n)}, Y_{x_k}^{(1)}$, $k \in [K] \setminus \{i\}$ are independent. As just observed, $\ell_k^* \in \{0, 1\}$ for all $k \neq i$.

b) For $\mathbf{j} \in C_i^*$, letting $k^* \neq i$ denote the unique value with $L_{k^*}^{\mathbf{j}} \neq \emptyset$, we have $\ell_k^* = 1$ for all $k \in \{k^*\} \cup (E_{k^*} \setminus \{i\})$ and $\ell_k^* = 0$ for all $k \notin \{k^*\} \cup E_{k^*}$.

The operators $G_k^{(\mathbf{j})}$ for $k \neq i$ are defined in (4.2) and (4.3). The operator $G_i^{(\mathbf{j})}$ acts on a matrix $A \in \mathbb{M}_n$ by permuting the indices as follows: $0 \to 0, k \to \mathrm{rk}(k)$ for $k \neq m$ and $m \to \ell_i^* = n$. As the distribution of $Y_{x_i}^{(n)}$ is invariant under such permutations, the random variable W_n satisfies (4.5).

Upon collecting our findings for the different cases and performing the integration in (4.1), it follows that the random matrix D_n satisfies

$$D_n \stackrel{d}{=} U_n D_n + V_n, \tag{4.6}$$

where $U_n \in [0,1]$, $V_n \in \mathbb{M}_n$, the $(U_n, V_n), D_n$ are independent, and we have

$$U_n = \sum_{i=1}^{K} \mathcal{R}_i^{\alpha} \cdot \mathbf{1}_{\mathbf{J} \in C_i \cup C_i^*} \tag{4.7}$$

and

$$V_n = \sum_{i=1}^{K} \sum_{\mathbf{j} \in C_i \cup C_i^*} \mathbf{1}_{\mathbf{J}=\mathbf{j}} \sum_{k \neq i}^{K} \mathcal{R}_k^{\alpha} \cdot G_k^{(\mathbf{j})}(D_{\ell_k^*}^{(k)}) + \sum_{\mathbf{j} \notin C \cup C^*} \mathbf{1}_{\mathbf{J}=\mathbf{j}} \sum_{k=1}^{K} \mathcal{R}_k^{\alpha} \cdot G_k^{(\mathbf{j})}(D_{\ell_k^*}^{(k)}). \tag{4.8}$$

Here, recall that we have $\mathbf{J} = (J_1, \ldots, J_n)$, where, given $(\mathcal{R}, \mathcal{S})$, the random variables $J_1, \ldots, J_n$ are independent and each J_i is distributed as J in (2.5), and $\{D_\ell^{(k)} : 0 \leq \ell \leq n-1, k \in [K]\}$ is an independent family of random variables, which is independent of $(\mathcal{R}, \mathcal{S}, \mathbf{J})$, where each $D_\ell^{(k)}$ is distributed like D_ℓ.

A random variable satisfying a fixed point equation of type (4.6) is called a *perpetuity*. It follows from classical results on perpetuities, e.g. from [51, Theorem 1.5], that (4.6) has at most one solution (in distribution). Repeating the arguments shows that D'_n satisfies a distributional identity of the form $D'_n = U'_n D'_n + V'_n$ with $U'_n = U_n$ and the additive term V'_n, can be obtained from V_n by replacing each $D_\ell^{(k)}$ by a copy of D'_ℓ while maintaining the independence structure. Hence, by our induction hypothesis, (U_n, V_n) and (U'_n, V'_n) are identically distributed which shows that D_n and D'_n are identically distributed and concludes the proof of the induction. □

4.2 Construction of a solution

In this section, we construct the family of processes $\{\mathcal{Z}^\vartheta : \vartheta \in \Theta\}$ mentioned in Section 3.4 that plays a central role in a number of proofs later on. The following proposition is a generalization of Theorem 6 and Theorem 17 in [14]. We keep the presentation rather compact and refer to [14] for more details on technical points; this applies in particular to a number of tedious but straightforward inductions occurring throughout the proof. Recall that $\|\cdot\|$ denotes the uniform norm on $\mathbb{C}_{\text{ex}}$ and the definition of $\mathfrak{m}$ in (3.1). For $a > 0$, let

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

Fix $c > 0$ and let $Q_0^\vartheta = c, \vartheta \in \Theta$. Recall the map Φ defined in Section 2.4. Recursively, for $n \geq 1$ and $\vartheta \in \Theta$, define

$$Q_n^\vartheta = \Phi(Q_{n-1}^{\vartheta 1}, \ldots, Q_{n-1}^{\vartheta K}, \mathcal{R}^\vartheta, \mathcal{S}^\vartheta, \Xi^\vartheta). \tag{4.9}$$

Proposition 13. For any $\vartheta \in \Theta$, almost surely, the sequence Q_n^ϑ defined in (4.9) converges uniformly to a process $\mathcal{Z}^\vartheta$. For any $\vartheta \in \Theta$, we have, almost surely,

$$\mathcal{Z}^\vartheta = \Phi(\mathcal{Z}^{\vartheta 1}, \ldots, \mathcal{Z}^{\vartheta K}, \mathcal{R}^\vartheta, \mathcal{S}^\vartheta, \Xi^\vartheta). \tag{4.10}$$

Furthermore, $\mathfrak{L}(\mathcal{Z}^\vartheta)$ is the unique solution to (2.10) in the set $\mathcal{M}^c$ and $\mathbf{E}[\|\mathcal{Z}^\vartheta\|^m] < \infty$ for all $m \geq 1$.

Remark 1. The proof of the proposition also shows the following: Let $\mu \in \mathcal{M}^c$, and let $\{\mathcal{Y}^\vartheta : \vartheta \in \Theta\}$ be a family of independent random processes with $\mathfrak{L}(\mathcal{Y}^\vartheta) = \mu$ for all $\vartheta \in \Theta$ which is independent of $\{(\mathcal{R}^\vartheta, \mathcal{S}^\vartheta, \Xi^\vartheta) : \vartheta \in \Theta\}$. Then, the sequence Q_n^ϑ initiated with $\{\mathcal{Y}^\vartheta : \vartheta \in \Theta\}$, that is, $Q_0^\vartheta = \mathcal{Y}^\vartheta$, converges almost surely uniformly to $\mathcal{Z}^\vartheta$.

Proof of Proposition 13. First of all, note that, for $n \geq 1$, (4.9) says that, for all $t \in [0,1]$,

$$Q_n^\vartheta(t) = \sum_{i=1}^K \mathbf{1}_{\Lambda_i^\vartheta}(t) \left[(\mathcal{R}_i^\vartheta)^\alpha Q_{n-1}^{\vartheta i}(\varphi_i^\vartheta(t)) + \sum_{j \in E_i} (\mathcal{R}_j^\vartheta)^\alpha Q_{n-1}^{\vartheta j}(\xi_j^\vartheta) \right]. \tag{4.11}$$

Throughout the proof, ξ denotes a random variable with the uniform distribution on $[0,1]$ which is independent of all remaining quantities. Recalling the choice of α in (2.5) and the fact that $\mathbf{E}[Q_0^\vartheta(\xi)] = c$, it follows by induction that $\mathbf{E}[Q_n^\vartheta(\xi)] = c$ for all $n \geq 0$. From (4.9) (or (4.11)), it should be clear that the sequences $(Q_n^\vartheta)_{n\geq 0}, \vartheta \in \Theta$, are identically distributed. (A formal proof could again be given by induction.) Next, set $\Delta Q_n^\vartheta := Q_{n+1}^\vartheta - Q_n^\vartheta$ for $n \geq 0$. From (4.9), we have for all $t \in [0,1]$,

$$\Delta Q_n^\vartheta(t) = \sum_{i=1}^K \mathbf{1}_{\Lambda_i^\vartheta}(t) \left[(\mathcal{R}_i^\vartheta)^\alpha \Delta Q_{n-1}^{\vartheta i}(\varphi_i^\vartheta(t)) + \sum_{j \in E_i} (\mathcal{R}_j^\vartheta)^\alpha \Delta Q_{n-1}^{\vartheta j}(\xi_j^\vartheta) \right]. \tag{4.12}$$

In particular, one finds $\Delta Q_n^\vartheta = h(Q_{n-1}^{\vartheta 1}, \dots, Q_{n-1}^{\vartheta K}, \mathcal{R}^\vartheta, \mathcal{S}^\vartheta, \Xi^\vartheta)$ for a suitable deterministic continuous function h (see (4.12)). By induction on n, it follows that, for any fixed $n \geq 1$, the random variables ΔQ_n^ϑ, $\vartheta \in \Theta$, are identically distributed.

For $i \in [K]$, we now define $\beta_i^\vartheta = \mathcal{R}_i^\vartheta$ if $\xi \in \bigcup_{k \in \Gamma_i} \Lambda_k^\vartheta$ and $\beta_i^\vartheta = 0$ otherwise. The random variable β_i^ϑ is distributed like β_i defined in (2.6). From (4.12), using the independence of $(\mathcal{R}^\vartheta, \mathcal{S}^\vartheta)$ and Ξ^ϑ and the fact that, conditional on $\xi \in \Lambda_i^\vartheta$, the relative position of ξ in this interval is uniform and independent of $(\mathcal{R}^\vartheta, \mathcal{S}^\vartheta)$, it follows from the last display that $\mathbf{E}[\Delta Q_n^\vartheta(\xi)^2]$ is equal to

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

Here the Λ_i^ϑ have distjoint interior, and all the squared terms are collected in the first sum (recall the definition of β_i from (2.6)). Conditional on $\xi \in \Lambda_i^\vartheta$ and on $\mathcal{R}_{j_1}^\vartheta$ 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. Hence:

$$\mathbf{E}[\Delta Q_n^\vartheta(\xi)^2] = \sum_{i=1}^K \mathbf{E}[\beta_i^{2\alpha}] \mathbf{E}[\Delta Q_{n-1}^\vartheta(\xi)^2] = \mathbf{E}[\Delta Q_1^\vartheta(\xi)^2] \left(\sum_{i=1}^K \mathbf{E}[\beta_i^{2\alpha}] \right)^{n-1}. \tag{4.13}$$

Recalling that $q_2 := \sum_{i=1}^K \mathbf{E}[\beta_i^{2\alpha}] \in (0,1)$, this implies that $\mathbf{E}[\Delta Q_n^\vartheta(\xi)^2] \leq C_2 q_2^n$ for $C_2 = 1/q_2$.

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 $q_m \in (0,1)$. The previous argument verifies this claim for $m = 2$ (and $m = 1$ by the Cauchy–Schwarz inequality). Let $m \geq 3$ and assume it is true for all $1 \leq \ell \leq m-1$ and let $C_* = \max(C_1, \dots, C_{m-1})$, $q_* = \max(q_1, \dots, q_{m-1}) < 1$. Then, again from (4.12), we deduce that

$$\begin{aligned} \mathbf{E}[|\Delta Q_n^\vartheta(\xi)|^m] &\leq \sum_{i=1}^K \mathbf{E}[\beta_i^{m\alpha}] \mathbf{E}[|\Delta Q_{n-1}^\vartheta(\xi)|^m] \\ &\quad + \sum_{i=1}^K \sum_{j_1, \dots, j_m} \mathbf{E}\left[\mathbf{1}_{\Lambda_i^\vartheta}(\xi) \prod_{k=1}^m |\Delta Q_{n-1}^{\vartheta j_k}(\xi_{j_k}^\vartheta)| \right], \end{aligned} \tag{4.14}$$

where the inner sum of the second term ranges over all tuples $(j_1, \dots, j_m) \in (E_i \cup \{i\})^m$ for which $\#\{j_1, \dots, j_m\} \geq 2$. (Note that we have dropped some factors $(\mathcal{R}_j^\vartheta)^\alpha$ in the product in the right-hand side.) 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. Since no j_k appears more than $m-1$ times in the product, we can use the induction hypothesis to obtain the loose, but sufficient, bound

$$\begin{aligned}\mathbf{E}[|\Delta Q_n^\vartheta(\xi)|^m] \leq \sum_{i=1}^K \mathbf{E}[\beta_i^{m\alpha}]\mathbf{E}[|\Delta Q_{n-1}^\vartheta(\xi)|^m] \\ + \sum_{i=1}^K \sum_{j_1,\dots,j_m} (C_* q_*^{n-1})^{\#\{j_1,j_2,\dots,j_m\}},\end{aligned} \tag{4.15}$$

with $j_1, \dots, j_m$ as in the sum in (4.14). From here, since $m \geq 3$ we have $\sum_{i=1}^K \mathbf{E}[\beta_i^{m\alpha}] < 1$ and a simple induction on n shows that $\mathbf{E}[|\Delta Q_n^\vartheta(\xi)|^m]$ decays exponentially in n, as desired.

The exponential decay for all moments at a uniform point ξ can be bootstrapped to yield exponential decay for sufficiently high (hence all) moments of the supremum. For $m \geq 1$ it follows from (4.12) that $\mathbf{E}[\|\Delta Q_n^\vartheta\|^m]$ equals

$$\begin{aligned}&\mathbf{E}\left[\max_{1\leq i\leq K}\left\{\left\|(\mathcal{R}_i^\vartheta)^\alpha \Delta Q_{n-1}^{\vartheta i}(\varphi_i^\vartheta(t)) + \sum_{j\in E_i}(\mathcal{R}_j^\vartheta)^\alpha \Delta Q_{n-1}^{\vartheta j}(\xi_j^\vartheta)\right\|^m\right\}\right] \\ &\leq \mathbf{E}\left[\max_{1\leq i\leq K}\left\{(\mathcal{R}_i^\vartheta)^{m\alpha}\|\Delta Q_{n-1}^{\vartheta i}\|^m + \sum_{j_1,\dots,j_m}\|\Delta Q_{n-1}^{\vartheta i}\|^{\ell_i}\prod_k |\Delta Q_{n-1}^{\vartheta j_k}(\xi_{j_k}^\vartheta)|\right\}\right]\end{aligned}$$

with $j_1, \dots, j_m$ as in the sum in (4.14), $\ell_i = \#\{1 \leq k \leq m : j_k = i\}$ and the product over k only ranges over those values $1 \leq k \leq m$ with $j_k \neq i$. (Observe that $\ell_i < m$ for all i.) Bounding the maximum by the sum, abbreviating $C_{**} = \max(C_1, \dots, C_m)$ and $q_{**} = \max(q_1, \dots, q_m)$ and using the stochastic independence of $\Delta Q_{n-1}^{\vartheta 1}, \dots \Delta Q_{n-1}^{\vartheta K}, \xi_1^\vartheta, \dots, \xi_K^\vartheta$, gives

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

where we used $p_i^* = \#\{j_k : 1 \leq k \leq m, j_k \neq i\}$. Recall $\mathfrak{m}$ from (3.1); for $m \geq \mathfrak{m}$ we have $\sum_{i=1}^K \mathbf{E}[\mathcal{R}_i^{m\alpha}] < 1$ and a simple induction on n shows that $\mathbf{E}[\|\Delta Q_n^\vartheta\|^m]$ decays exponentially in n. From there, standard arguments (see, e.g. the proof Theorem 6 in [14]) imply that, almost surely, Q_n^ϑ converges uniformly and, writing $\mathcal{Z}^\vartheta$ for its limit, the identity in (4.10) holds. Furthermore, the random variable $\|\mathcal{Z}^\vartheta\|$ has finite polynomial moments of all orders.

It remains to show that the constructed process is the unique solution to (2.10) in distribution in $\mathcal{M}^c$. As this part does not require significantly new ideas, we remain brief. Let $\mu \in \mathcal{M}^c$ and consider a set of independent random variables $\{\mathcal{Y}_0^\vartheta, \vartheta \in \Theta\}$ with $\mathfrak{L}(\mathcal{Y}_0^\vartheta) = \mu$ that is independent of $\{(\mathcal{R}^\vartheta, \mathcal{S}^\vartheta, \Xi^\vartheta) : \vartheta \in \Theta\}$. Then, analogously to (4.9), for $n \geq 1$ and $\vartheta \in \Theta$, recursively define

$$\mathcal{Y}_n^\vartheta = \Phi(\mathcal{Y}_{n-1}^{\vartheta 1}, \dots, \mathcal{Y}_{n-1}^{\vartheta K}, \mathcal{R}^\vartheta, \mathcal{S}^\vartheta, \Xi^\vartheta).$$

By repeating the steps of the proof of uniform convergence of $(Q_n^\vartheta)_{n\geq 0}$, we find that the sequence $(\mathcal{Y}_n^\vartheta)_{n\geq 0}$ converges almost surely uniformly to a solution $\mathcal{Y}^\vartheta$ of (2.10). Furthermore, by following the same inductive arguments as before, one can show that $\mathbf{E}[(Q_n^\vartheta(\xi)-\mathcal{Y}_n^\vartheta(\xi))^2]\to 0$ exponentially fast. (To be precise, we have the same bound as in (4.13) for this term with $\mathbf{E}[\Delta Q_1^\vartheta(\xi)^2]$ replaced by $\mathbf{E}[(\mathcal{Y}_0^\vartheta(\xi)-c)^2]$.) The fact that $\mathbf{E}[|Q_n^\vartheta(\xi)-\mathcal{Y}_n^\vartheta(\xi)|^m]\to 0$ exponentially fast for all $1\leq m\leq \mathfrak{m}$ follows as in (4.14) and (4.15). Based on this, the same steps as before yield that $\mathbf{E}[\|Q_n^\vartheta-\mathcal{Y}_n^\vartheta\|^m]\to 0$ for all $1\leq m\leq \mathfrak{m}$. Hence, $\mathcal{Y}^\vartheta=\mathcal{Z}^\vartheta$ almost surely. This concludes the proof. □

Proof of Theorem 1 i) and ii). The statement in *ii)* is an immediate consequence of Proposition 12 and Proposition 11 *ii)* since $\mathbf{E}[\nu^{\mathfrak{X}}]$ remains invariant upon replacing $\mathcal{X}$ by the support of the measure. The uniqueness claim for the process in Theorem 1 *i)* can be deduced as follows: Let $\mathcal{Z}$ be the process constructed in Proposition 13, and assume that $\mathcal{Y}$ is a continuous excursion satisfying (2.10) with $\mathbf{E}[\mathcal{Y}(\xi)]=\mathbf{E}[\mathcal{Z}(\xi)]$. Then, by Proposition 12, $\mathbf{E}[\nu_n^{\mathfrak{T}_\mathcal{Z}}]=\mathbf{E}[\nu_n^{\mathfrak{T}_\mathcal{Y}}]$ for all $n\geq 1$. Let $f_n:\mathcal{M}_1(\mathbb{M}_n)\to\mathcal{M}_1([0,\infty))$ be the map that, to $\nu\in\mathcal{M}_1(\mathbb{M}_n)$, associates the law of $\sup\{A_{0,i}:0\leq i\leq n\}$, where $\mathfrak{L}(A)=\nu$. Then, as $n\to\infty$, we have the following weak convergences:

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

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

4.3 Attractiveness of the fixed points of (2.4)

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 $\mathbb{T}_{\mathrm{f}}^{\mathrm{ghp}}$. For technical reasons, we work in the space $\mathbb{D}$ of càglàd functions $f:[0,1]\to\mathbb{R}$ satisfying $f(t)=\lim_{s\uparrow t}f(s)$ for all $t\in(0,1]$ and for which the right-hand limits $f(t+)=\lim_{s\downarrow t}f(s)$ exist for all $t\in[0,1)$. The set $\mathbb{D}$ is equipped with the Skorokhod J_1-topology ([12, Chapter 3]). Let $\mathbb{D}_{\mathrm{ex}}\subset\mathbb{D}$ be the set of non-negative functions $f\in\mathbb{D}$ with $f(0)=f(0+)=0$ and $f(t)-f(t+)\geq 0$ for all $t\in[0,1]$. Analogously to continuous excursions, every $f\in\mathbb{D}_{\mathrm{ex}}$ encodes a compact rooted measured real tree $(\mathcal{T}_f,d_f,\mu_f,\rho_f)$ satisfying **C1** via the construction outlined in Section 2.2 ([24, Lemma 2.1]). The proof of the following lemma is found in the appendix.

Lemma 14. Let ν be a probability distribution on $\mathbb{T}_{\mathrm{f}}^{\mathrm{ghp}}$. Then, there exists a probability distribution η on $\mathbb{D}_{\mathrm{ex}}$ such that $\mathfrak{L}(\mathfrak{T}_\mathcal{Z})=\nu$ for a random variable $\mathcal{Z}$ with law η.

Proof of Theorem 2. i) For $m,n\geq 1$, write $\tilde{D}_n^{(m)}$ for a generic random variable with distribution $\nu_n^{\phi^m(\nu)}$. Recall from the proof of Proposition 12 that, if $\mathfrak{T}$ satisfies (2.10), then for $n\geq 1$, a r.v. D_n with distribution $\mathbf{E}[\nu_n^{\mathfrak{T}}]$ satisfies the fixed point equation (4.6), that is, $D_n\overset{d}{=}U_nD_n+V_n$, where U_n is a real-valued r.v. with $U_n\in(0,1)$ a.s. that is given in (4.7), V_n is a random matrix with non-negative entries given in (4.8), and $(U_n,V_n),D_n$ are independent. Since $\phi^{m+1}=\phi\circ\phi^m$, the arguments from the proof of Proposition 12 show that

$$\tilde{D}_n^{(m+1)}\overset{d}{=}U_n\tilde{D}_n^{(m)}+V_n^{(m)},$$

where $(U_n,V_n^{(m)})$ and $\tilde{D}_n^{(m)}$ are independent and $V_n^{(m)}$ is given as V_n in (4.8) upon replacing each copy of $D_\ell^{(k)}$ by a copy of $\tilde{D}_\ell^{(m)}$ and maintaining the independence between $\mathcal{R},\mathcal{S},\mathbf{J}$ and the copies of $\tilde{D}_1^{(m)},\dots,\tilde{D}_{n-1}^{(m)}$. The remainder of the proof consists 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. The proof relies on a contraction

argument. In the following, as matrices in $\mathbb{M}_n$ are symmetric with zeros on the diagonal, we use the natural identification of $\mathbb{M}_n$ 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 max-norm $\|\cdot\|_\infty$ is integrable. Recall the *Wasserstein* distance on $\mathcal{M}_1^1(\mathbb{R}^p)$ defined by, for $\mu_1, \mu_2 \in \mathcal{M}_1^1(\mathbb{R}^p)$,

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

For random variables X, Y in $\mathbb{R}^p$, we abbreviate $\ell_1(X, Y) := \ell_1(\mathfrak{L}(X), \mathfrak{L}(Y))$. 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 $m \to \infty$. First, observe that $\|G_i^{\mathbf{j}}(A)\|_\infty \leq \|A\|_\infty$ for all matrices A and linear operators $G_i^{\mathbf{j}}$ in and around (4.2)–(4.3). Since the random variables $\mathcal{R}_i, i \in [K]$, lie in $[0, 1]$, conditioning on $(\mathcal{R}, \mathcal{S}, \mathbf{J})$ shows that

$$\ell_1(\tilde{D}_n^{(m+1)}, D_n) \leq \mathbf{E}[U_n] \cdot \ell_1(\tilde{D}_n^{(m)}, D_n) + \sup_{1 \leq \ell \leq n-1} \ell_1(\tilde{D}_\ell^{(m)}, D_\ell).$$

From here, since $\mathbf{E}[U_n] < 1$ and $\ell_1(\tilde{D}_\ell^{(m)}, D_\ell) \to 0$ for all $1 \leq \ell \leq n-1$ by the induction hypothesis, it follows by induction on m that the sequence $(\ell_1(\tilde{D}_n^{(m)}, D_n))_{m \geq 1}$, is bounded. Taking the limit superior in the last display then shows that $\ell_1(\tilde{D}_n^{(m)}, D_n) \to 0$ as $m \to \infty$ since $\mathbf{E}[A_n] < 1$. This concludes the proof of the induction step, and it only remains to establish the base case $n = 1$. Note that, under our identification, $\tilde{D}_1^{(m)}$ and D_1 are real-valued and non-negative random variables. Distributional convergence $\tilde{D}_1^{(m)} \to D_1$ as $m \to \infty$ follows immediately from Theorem 2 b) in [27]. Furthermore, by construction ϕ^m preserves the expected distance between the root and an independent random point and therefore $\mathbf{E}[\tilde{D}_1^{(m)}] = \mathbf{E}[D_1]$ for all $m \geq 1$. But convergence in ℓ_1 for non-negative random variables is equivalent to distributional convergence together with convergence of the mean, see, e.g. [11, Lemma 8.3]. This concludes the proof of *i)*.

ii) If $\nu \in \mathcal{M}_1(\mathbb{T}_{\mathrm{f}}^{\mathrm{ghp}})$ the claim follows easily from the proof of Proposition 13. Indeed, by Lemma 14, there exists a probability distribution ν^* on $\mathbb{D}_{\mathrm{ex}}$ such that the (isometry class of the) tree encoded by a random variable with law ν^* has distribution ν. Then, for any $n \geq 0$, $\phi^n(\nu)$ is the distribution of the real tree encoded by $Q_n^\emptyset$ from the proof of Proposition 13 when Q_0^ϑ, $\vartheta \in \Theta_n$, are i.i.d. with distribution ν^*; see Remark 1. (All arguments in the proof of Proposition 13 apply analogously to càglàd functions.) As we have seen there, $Q_n^\emptyset$ converges almost surely to a solution of (2.4). (It is here where we need the moment assumption on $\|\mathfrak{T}\|$.) This proves the claim. □

Proof of Proposition 3. i) The example we provide generalizes the one by Albenque and Goldschmidt [2] in the special case of Example **1)**. Let $\tau = \mathfrak{L}((\mathcal{S}, \mathcal{S}))$ where $\mathfrak{L}(\mathcal{S}) = \eta$. (In other words, we choose $\mathfrak{L}(\mathcal{R}) = \eta$ and the coupling $\mathcal{R} = \mathcal{S}$.) Let ν_c be the unique law solving (2.4) 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 $\kappa \geq 1$. Let c_0^+ be the set of all non-negative sequences converging to zero. Let $\mathfrak{X} = (\mathcal{X}, d, \mu, \rho)$ be a fixed compact measured metric space. For a sequence of points $u = (u_n)_{n \geq 0}$ in $\mathcal{X}$ and $s \in c_0^+$, let $\chi_1(u, s)$ be the isometry class of the space obtained upon attaching κ disjoint segments of length s_i at the point u_i, each by one extremity, for all $i \geq 1$. As $s \in c_0^+$, the resulting space is compact. Hence, $\chi_1 : \mathcal{X}^{\mathbb{N}} \times c_0^+ \to \mathbb{K}^{\mathrm{ghp}}$ is well-defined. The map is continuous as proved in Lemma 28 in the appendix.

Let $\mathcal{P}$ be a Poisson point process with intensity measure $\mu \otimes s^{-1-1/\alpha} ds$ on $\mathcal{X} \times [0, \infty)$. $\mathcal{P}$ can be considered a $(\mathcal{X}^{\mathbb{N}} \times c_0^+)$-valued random variable. The isometry class $f(\mathcal{P})$ is a random variable whose distribution does not depend on the choice of the representative

of the isometry class of $\mathfrak{X}$. Hence, this operation defines a map $\psi_1 : \mathbb{K}^{\mathrm{ghp}} \to \mathcal{M}_1(\mathbb{K}^{\mathrm{ghp}})$, which is continuous (Lemma 28 in the appendix) 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 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, because of the choice of intensity measure with exponent $-1-1/\alpha$ in the length, the effects of a multiplication of the mass by C and of a multiplication lengths by C^α are equivalent, and thus Ψ and Ψ_1 commute. In other words, for $\upsilon \in \mathcal{M}_1(\mathbb{K}^{\mathrm{ghp}})$, 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 (2.4), the measure $\Psi_1(\nu)$ also solves (2.4). In particular, $\Psi_1(\nu_c)$ is such a fixed point and it charges only $\mathbb{T}^{\mathrm{ghp}} \setminus \mathbb{T}^{\mathrm{ghp}}_{\mathrm{f}}$.

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

$$\mathbf{P}(\|\mathfrak{T}\| \geq h) \geq \mathbf{P}(\mathrm{Po}(\textstyle\int_h^\infty s^{-1-1/\alpha} ds) \geq 1) = 1 - \exp(-\alpha h^{-1/\alpha}), \tag{4.16}$$

where $\mathrm{Po}(\lambda)$ denotes a Poisson random variable with parameter λ. Since $\mathfrak{m} \geq 1/\alpha$ (see above (3.1)) it follows readily that $\mathbf{E}[\|\mathfrak{T}\|^{\mathfrak{m}}] = \infty$.

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. Indeed, we can replace the segments by loops, or equivalently identify the extremities of all the segments together when $\kappa \geq 2$. This proves that there exist fixed points which are real trees with probability zero.

ii) The measure $\Psi_1(\nu_c)$ we have just constructed charges only spaces in $\mathbb{T}^{\mathrm{ghp}} \setminus \mathbb{T}^{\mathrm{ghp}}_{\mathrm{f}}$ since the hair has no mass; we now provide a modified example where (a) the metric part of the spaces is the unique (up to scaling) fixed point in $\mathbb{T}^{\mathrm{ghp}}_{\mathrm{f}}$ on which we graft hair, just as in *i)* before, but (b) we modify the mass measure so that it also charges the hair. We then prove that the diameter of spaces obtained by iteration does not converge in distribution to the diameter of $\mathfrak{T}_{\mathcal{Z}}$.

Let $\mathfrak{X}^\circ = (\mathcal{X}^\circ, d^\circ, \mu^\circ, \rho^\circ)$ be any *fixed* compact measured metric space where μ° has no atoms; to fix ideas, one may take $\mathcal{X}^\circ = [0,1]$ equipped with the Euclidean metric, the Lebesgue measure, and $\rho^\circ = 0$. From $\mathfrak{X}^\circ$, we construct the "hairy" compact measured metric space $(\mathcal{X}, d, \mu, \rho)$ using the decoration by a Poisson point process as above. Here, and subsequently, we focus on the case $\kappa = 1$. Now, we construct a new measure $\mu_{\mathcal{P}}$ on $\mathcal{X}$ as follows. We first set $\mu_{\mathcal{P}}(\mathcal{X}^\circ) = 0$. Then, for $(u,s) \in \mathcal{P}$, we associate a total $\mu_{\mathcal{P}}$ mass $\min\{s,1\}^{1/\alpha+1}$ to the segment of length s attached to u; we distribute this mass along this segment with density $\alpha \cdot (1+\alpha)^{-1} \cdot r^{1/\alpha} \mathbf{1}_{\{0 \leq r \leq s\}} dr$ if $s \leq 1$, and with density $e^{-r}(1-e^{-r})^{-1}\mathbf{1}_{\{0 \leq r \leq s\}} dr$ if $s > 1$. Then,

$$\begin{aligned}\mathbf{E}[\mu_{\mathcal{P}}(\mathcal{X})] = \int_0^\infty \min\{s,1\}^{1/\alpha+1} s^{-(1+1/\alpha)} ds &= \int_0^1 ds + \int_1^\infty s^{-(1+1/\alpha)} ds \\ &= 1 + \alpha < \infty\,,\end{aligned}$$

and it follows that $\mu_{\mathcal{P}}(\mathcal{X}) < \infty$ almost surely. We let μ^* be the unique probability measure on $\mathcal{X}$ that is proportional to $\mu + \mu_{\mathcal{P}}$. For a random point ζ^* sampled according to μ^*, we have

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

and

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

It follows that $\lambda^* := \mathbf{E}[d(\rho,\zeta^*)] < \infty$, so we may rescale the metric and define $d^*(\cdot,\cdot) = d(\cdot,\cdot)\times c/\lambda^*$. Finally, let $\mathfrak{X}^* = (\mathcal{X}, d^*, \mu^*, \rho)$ which satisfies $\mathbf{E}[d^*(\rho,\zeta^*)] = c$ and $\mathbf{E}[\|\mathfrak{X}^*\|^{1/\alpha}] = \infty$. As for the function ψ_1 in part *i)*, technical proofs which we omit show that the isometry class of $\mathfrak{X}^*$ is a random variable whose distribution only depends on the ghp-equivalence class of $\mathfrak{X}^\circ$. We denote this distribution by $\nu^* = \nu^*_c \in \mathcal{M}_1(\mathbb{K}^{\text{ghp}}_{\text{f}})$. Note that it is not difficult to construct explicitly a random variable $g \in \mathbb{C}_{\text{ex}}$ in terms of the points in $\mathcal{P}$ such that $\mathfrak{T}_g = \mathfrak{X}^*$ (with respect to isometry classes).

Fix $n \ge 1$ and let $\mathfrak{X}^*_n$ be a random variable with distribution $\phi^n_{\text{ghp}}(\nu^*)$. We now study the diameter of $\mathfrak{X}^*_n$. To this end, let $\tilde{\nu}$ be the distribution of the isometry class of $\tilde{\mathfrak{X}} = (\mathfrak{X}, d, \mu^*, \rho)$ and $\tilde{\mathfrak{X}}_n$ be a random variable with distribution $\phi^n_{\text{ghp}}(\tilde{\nu})$. As the rescaling for distances is deterministic, in distribution, we obtain $\mathfrak{X}^*_n$ from $\tilde{\mathfrak{X}}_n$ by multiplying distances in $\tilde{\mathfrak{X}}_n$ by c/λ^*. We can think of the random variable $\tilde{\mathfrak{X}}_n$ as constructed from a family of independent copies $\{\tilde{\mathfrak{X}}^\vartheta : \vartheta \in \Theta_n\}$ of $\tilde{\mathfrak{X}}$ relying on the family of independent rescaling factors $\{(\mathcal{R}^\vartheta, \mathcal{S}^\vartheta) : \vartheta \in \Theta_n\}$, where glue points are chosen with respect to the mass measures. (Formally, we can use the space $\mathfrak{T}_{Q^\emptyset_n}$ with $Q^\emptyset_n$ from the proof of Proposition 13 when $Q^\vartheta_0, \vartheta \in \Theta_n$, are i.i.d. copies of g mentioned above.) Recall that the distribution of $\mathcal{P}$ has been tailored precisely so that it is not affected by the rescaling: the distribution of lengths of the segments on the union of the rescaled copies of the $\tilde{\mathfrak{X}}^\vartheta$ is equal to the distribution of lengths of the segments on $\tilde{\mathfrak{X}}$ [7]. In particular, it follows that the diameter of $\tilde{\mathfrak{X}}_n$ is bounded from below in a stochastic sense by the largest segment appearing in the Poisson process. From these considerations and (4.16), it follows that, for $x > 0$,

$$\mathbf{P}(\text{diam}(\mathfrak{X}^*_n) \ge \lambda^* x/c) = \mathbf{P}(\text{diam}(\tilde{\mathfrak{X}}_n) \ge x) \ge 1 - \exp(-\alpha x^{-1/\alpha}).$$

As $\|\mathfrak{T}_{\mathcal{Z}}\|$ has finite moments of all orders and the last display is valid for all n, it follows that, for all x large enough, the left hand side of the last display does not converge to $\mathbf{P}(\text{diam}(\mathfrak{T}^*_{\mathcal{Z}}) \ge \lambda^* x/c)$. This concludes the proof. □

5 Proofs of the geometric properties

All arguments in this section rely on a generalization of the decomposition of the tree $\mathcal{T}_{\mathcal{Z}}$ into K subtrees corresponding to $\mathcal{T}_{\mathcal{Z}^1}, \dots, \mathcal{T}_{\mathcal{Z}^K}$ when the fixed point equation (2.10) is developed for several levels. Throughout Section 5, let $\mathfrak{T}^\vartheta = (\mathcal{T}^\vartheta, d^\vartheta, \mu^\vartheta, \rho^\vartheta) := \mathfrak{T}_{\mathcal{Z}^\vartheta}$ with $\mathcal{Z}^\vartheta$ as in Proposition 13 where, for the sake of convenience, we assume $\int_0^1 \mathbf{E}[\mathcal{Z}^\vartheta(t)]dt = 1$.

The structural tree Γ describes the way the trees $\mathcal{T}^1, \dots, \mathcal{T}^K$ (or rather, their images under the canonical surjection φ°), are arranged in $\mathcal{T} = \mathcal{T}_{\mathcal{Z}}$. Similarly, we introduce a structural tree for the decomposition at level $n \ge 1$ 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, that is, $\mathfrak{T}^\vartheta, \vartheta \in \Theta_n$, when carried out up to this level. We think of Γ^n as rooted at $1\dots1$. Observe that Γ^n is a random object, since the adjacency relations depend on the random points used to glue the trees. The tree Γ^n is measurable with respect to $\{(\mathcal{S}^\vartheta, \Xi^\vartheta) : 0 \le |\vartheta| < n\}$. Analogously to the construction on page 6, we shall consider $\mathcal{T} = \mathcal{T}_{\mathcal{Z}}$ as the disjoint union $\sqcup_{\vartheta\in\Theta_n}\mathcal{T}^\vartheta$ upon identifying each root of a tree $\mathcal{T}^\vartheta, \vartheta \ne 1\dots1$ with the glue point on the associated parent tree. We write $\varphi^\circ_n : \sqcup_{\vartheta\in\Theta_n}\mathcal{T}^\vartheta \to \mathcal{T}$ for the corresponding canonical surjection. See Figure 3 for an illustration.

[7] This is true, regardless of the fact that these segments may not be hair any longer in $\tilde{\mathfrak{X}}$, since the gluing occurs after the rescalings.

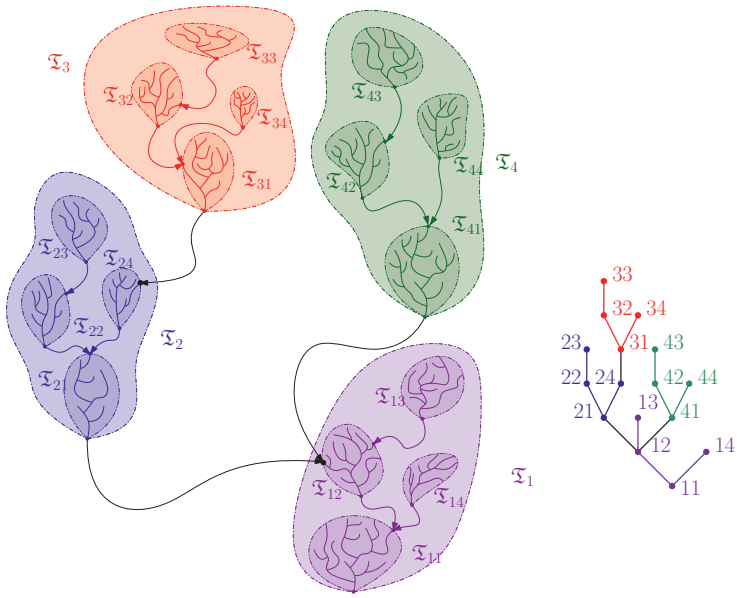

Figure 3: The construction of $\mathfrak{T}$ from the first two levels of Θ, and the corresponding structural tree Γ^2 on the right-hand side.

5.1 The upper Minkowski dimension: Proof of Theorem 4

In the following, for the sake of clarity, we write $a\mathcal{X}$ for the metric space $(\mathcal{X}, ad)$, where a constant $a > 0$ and a metric space $(\mathcal{X}, d)$. The approach here is to turn the heuristic presented in (3.2) into a rigorous argument.

Given $x > 0$, the idea consists in (1) expanding the recurrence for each subspace that has a scaling factor (for distances) greater than x^α, until we are left with spaces (potentially in different generations) with scaling factor no greater than x^α; (2) prove that, with high probability and for any $\epsilon > 0$, we can cover the whole space by using a single ball of radius $x^{\alpha-\epsilon}$ per subspace, which ends up being a collection of no more than $x^{-1-\epsilon}$ balls, and finally (3) show that the bounds ensure that this occurs for all x small enough with probability one. Let us move on to the details.

With $x > 0$ fixed, we first look for the suitable portion of the space that should intuitively be covered with balls of size roughtly x^α. For $\vartheta \in \Theta \setminus \{\varnothing\}$, we let ϑ^+ be the direct ancestor (parent) of ϑ in Θ. Let $\mathscr{L}_x := \{\vartheta \in \Theta : \mathcal{V}(\vartheta) \le x, \mathcal{V}(\vartheta^+) > x\}$. The set $\mathscr{L}_x$ separates the root $\varnothing$ from infinity, and no two elements of $\mathscr{L}_x$ lie on the same ancestral path. It follows that $\sum_{\vartheta \in \mathscr{L}_x} \mathcal{V}(\vartheta) = 1$ almost surely, and $\mathscr{L}_x$ provides a natural partition of the entire space $\mathcal{T}$. The cardinality of $\mathscr{L}_x$ has been studied finely by Janson and Neininger [33]; we only need rather weak results: it turns out that under the conditions of Theorem 4 (their Condition A), as one would expect, $\mathbf{E}[\#\mathscr{L}_x] = O(1/x)$ as $x \to 0$, and therefore, for any $\epsilon > 0$, by Markov's inequality,

$$\mathbf{P}(\mathscr{L}_x > x^{-1-\epsilon}) \le Cx^\epsilon,$$

for some constant C.

Now, for each $\vartheta \in \mathscr{L}_x$, the corresponding portion of the space is stochastically smaller (coupling of the scale factor) than $x\mathcal{X}$, and since the height has moments of all orders,

by choosing k as the smallest integer such that $k\epsilon > 2$, we have

$$\begin{aligned}\mathbf{P}(\exists\vartheta \in \mathscr{L}_x : \mathcal{V}(\vartheta)^\alpha \|\mathfrak{T}^\vartheta\| > x^{\alpha-\epsilon}) &\le \mathbf{P}(\exists\vartheta \in \mathscr{L}_x : \|\mathfrak{T}^\vartheta\| > x^{-\epsilon}) \\ &\le Cx^\epsilon + x^{-1-\epsilon}\mathbf{E}[\|\mathfrak{T}\|^k]x^{k\epsilon} \\ &\le Cx^\epsilon + C'x^{1-\epsilon},\end{aligned}$$

where the constant C' only depends on ϵ.

As a consequence, using balls of radius $x^{\alpha-\epsilon}$ centered at the roots ρ^ϑ for $\vartheta \in \mathscr{L}_x$, it follows that $N_{\mathcal{T}}(x^{\alpha-\epsilon})$ is such that $\mathbf{P}(N_{\mathcal{T}}(x^{\alpha-\epsilon}) > x^{-1-\epsilon}) \le C''x^\epsilon$ for $\epsilon \in (0, 1/2)$. Choosing $x = 2^{-n}$, a straightforward application of the Borel–Cantelli lemma implies that with probability one, $N_{\mathcal{T}}(2^{-n(\alpha-\epsilon)}) \le 2^{n(1+\epsilon)}$ for all but finitely many natural numbers n. This is then easily extended to any value of x by considering the unique $n \in \mathbb{N}$ for which $2^{-n} < x \le 2^{-n+1}$, and the fact that one then has $\mathcal{N}_{\mathcal{T}}(x^{\alpha-\epsilon}) \le \mathcal{N}(2^{-n})$. It follows that with probability one,

$$\limsup_{x\to 0} \frac{\log \mathcal{N}_{\mathcal{T}}(x^{\alpha-\epsilon})}{\log 1/x^{\alpha-\epsilon}} \le \frac{1+\epsilon}{\alpha-\epsilon}\,.$$

Since $\epsilon \in (0, 1/2)$ is arbitrary, this completes the proof of Theorem 4.

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

For a real tree $\mathcal{T}_f$ with $f \in \mathbb{C}_{\text{ex}}$, natural candidates for measures giving lower bounds on $\dim_{\text{h}}(\mathcal{T}_f)$ using the mass distribution principle (2.14) are the push-forward measures $\mu_f^* := \mu^* \circ \pi_f^{-1}$ of measures μ^* on $[0, 1]$ under the surjection $\pi_f : [0, 1] \to \mathcal{T}_f$. In our setting, in the case $\mathcal{R} = \mathcal{S}$ which covers both examples in (2.11) and (2.12), 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 (2.13), it turns out that one first has to find an appropriate time-change on the unit interval in order to re-correlate the masses of fragments in the tree with the extent of distances in the corresponding subtrees. Time-changes constructed in this context are typically random, and one is led to construct the pair "tree+time-change" simultaneously. It is this situation in which the almost sure construction in Section 4.2 turns out especially useful.

The construction of a suitable time-change is presented in the following proposition which requires the introduction of additional notation. Let Φ_1 be the analog of the map Φ defined in (2.9) for $\alpha = 1$ that combines functions of the space $\mathbb{C}_1 = \{f \in \mathbb{C} : f \ge 0, f(0) = 0, f(1) = 1\}$. In other words: $\Phi_1 : \mathbb{C}_1^K \times \Sigma_K^2 \times (0, 1)^K \to \mathbb{C}_1$, such that $g = \Phi(f_1, \dots, f_K, \mathbf{r}, \mathbf{s}, \mathbf{u})$ is the unique function of $\mathbb{C}_1$ for which

$$g(x) - g(y) = r_{w_\ell}\left[f_{w_\ell}(\varphi_{w_\ell}(x)) - f_{w_\ell}(\varphi_{w_\ell}(y))\right],$$

for all $1 \le \ell \le L$ and $x, y \in I_\ell$ (see around the definition of Φ for the notation). Further, for $i \in \Gamma^o$ let $\Lambda_i^{\vartheta-}$ and $\Lambda_i^{\vartheta+}$ be the two half-open intervals forming Λ_i^ϑ where $\inf \Lambda_i^{\vartheta-} \le \sup \Lambda_i^{\vartheta+}$. For the various quantities playing a role in the proposition, we refer the reader to Section 3.4.

Proposition 15. Almost surely, for any $\vartheta \in \Theta$, there exists a probability measure μ^ϑ on $[0, 1] = \Lambda_\emptyset^\vartheta$, such that, for every $\sigma \in \Theta$,

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

The (random) distribution function F^ϑ of μ^ϑ is measurable with respect to $\{\mathcal{R}^{\vartheta\sigma}, \mathcal{S}^{\vartheta\sigma}, \Xi^{\vartheta\sigma} : \sigma \in \Theta\}$, and if ξ is an independent random variable uniform in $[0, 1]$, then $\mathbf{E}[F^\vartheta(\xi)] = 1/2$.

All F^ϑ, $\vartheta \in \Theta$, have the same law and satisfy

$$F^\vartheta = \Phi_1(F^{\vartheta 1}, \dots, F^{\vartheta K}, \mathcal{R}^\vartheta, \mathcal{S}^\vartheta, \Xi^\vartheta)\,. \tag{5.1}$$

Furthermore, in distribution, F^ϑ is the unique continuous distribution function on $[0,1]$ satisfying (5.1).

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 $\mu_m^\vartheta(\Lambda_\sigma^\vartheta) = \mu_n^\vartheta(\Lambda_\sigma^\vartheta)$. 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 F_n^ϑ the distribution function of μ_n^ϑ. By construction, almost surely,

$$F_n^\vartheta = \Phi_1(F_{n-1}^{\vartheta 1}, \dots, F_{n-1}^{\vartheta K}, \mathcal{R}^\vartheta, \mathcal{S}^\vartheta, \Xi^\vartheta). \tag{5.2}$$

Note that $\{(F_n^\vartheta)_{n\geq 0} : \vartheta \in \Theta\}$ is a family of identically distributed random variables. The uniform convergence of F_n^ϑ is shown analogously to the convergence of Q_n^ϑ in Proposition 13: write F_n^ϑ as a telescoping sum, prove almost sure convergence of the corresponding series at an independent uniform point, and bootstrap to almost sure uniform convergence; we omit the details. Here, it is important to note that, since $F_0^\vartheta(t) = t$ for all $t \in [0,1]$, one can verify inductively that $\mathbf{E}[F_n^\vartheta(\xi)] = 1/2$ for all $n \geq 1$. The relevant constant that takes a value smaller than one, which allows to establish the convergence of F_n^ϑ at a uniformly chosen point by a contraction argument is $\sum_{i=1}^K \mathbf{E}[\beta_i^2]$ (this is similar to the constant in (4.13) with α replaced by 1, just as Φ_1 is similar to Φ with α replaced by 1). From the convergence of F_n^ϑ and (5.2), it follows that F^ϑ satisfies (5.1) and is a continuous distribution function. The induced measure μ^ϑ satisfies the desired properties. □

In the remainder of the section (and here only), let $\bar{\mu} = \mu^\emptyset \circ \pi_{\mathcal{Z}}^{-1}$ with $\mu^\emptyset$ as in Proposition 15.

Lemma 16. Assume that $\mathbf{E}[\overline{\mathcal{R}}^{-\delta}] < \infty$ for some $\delta > 0$ and let ζ be drawn on $\mathcal{T}$ according to μ and $\bar{\zeta}$ be drawn according to $\bar{\mu}$, where these random variables are independent given $\mathfrak{T}$. Then, there exists an $\varepsilon > 0$, such that,

$$\max\Big\{\mathbf{P}(d(\rho,\zeta) < r), \mathbf{P}(d(\rho,\bar{\zeta}) < r), \mathbf{P}(d(\zeta,\bar{\zeta}) < r)\Big\} = O(r^\varepsilon), \quad r \downarrow 0.$$

Remark. In most classical examples, we have $\mu = \bar{\mu}$ and invariance by rerooting at a random point distributed according to μ. This applies in particular to $\mathfrak{T}_{\mathscr{Z}}$ and $\mathfrak{T}_{\mathbf{e}}$. The tree $\mathfrak{T}_{\mathscr{H}}$ is invariant under rerooting but the time-change is non-trivial. The random variable $\mathbf{e}(\xi)$ has the Rayleigh distribution, thus $\mathbf{P}(\mathbf{e}(\xi) \leq t) = 1 - e^{-t^2/8} = t^2/8 + o(t^2)$ as $t \to 0$.

Proof. Since the arguments are similar: we give all the details for the bound on $d(\rho,\zeta)$ and only discuss the relevant modifications for $d(\rho,\bar{\zeta})$ and $d(\zeta,\bar{\zeta})$. Recall the construction discussed at the beginning of this section involving the random tree Γ^n for $n \geq 1$.

i) The lower bound on $d(\rho,\zeta)$. Fix $n \geq 1$ and let ζ be as in the lemma. We define Γ^* as the set of nodes of Γ^n containing all vertices $\vartheta \in \Theta_n$ such that $\varphi_n^\circ(\mathcal{T}^\vartheta)$ intersects the segment $[\![\rho,\zeta]\!]$. Note that Γ^* forms a path in Γ^n. By construction, $N = \#\Gamma^*$ is distributed as the number of individuals in the n-th generation of a discrete-time branching process with offspring distribution $\nu = \mathfrak{L}(1 + \#E_J)$. Note that $\nu(\{0\}) = 0$ and $\nu(\{1\}) < 1$. For $\vartheta \in \Gamma^*$, the contribution of the intersection of $[\![\rho,\zeta]\!]$ with $\varphi_n^\circ(\mathcal{T}^\vartheta)$ to $d(\rho,\zeta)$ is distributed

as a scaled copy of Y, the height of a random point (see around (2.6)). More precisely, we have

$$Y \stackrel{d}{=} \sum_{\vartheta \in \Gamma^*} \mathcal{V}(\vartheta)^\alpha Y^\vartheta, \tag{5.3}$$

with a family of independent random variables $\{Y^\vartheta : \vartheta \in \Theta_n\}$ distributed as Y which is independent of $\{(\mathcal{R}^\vartheta, \mathcal{S}^\vartheta) : |\vartheta| < n\}, \Gamma^*$. (Note that Γ^* and $\{(\mathcal{R}^\vartheta, \mathcal{S}^\vartheta) : |\vartheta| < n\}$ are typically not independent.) Therefore, with $m_n = \min\{\mathcal{V}(\vartheta) : |\vartheta| = n\}$ and $c > 0$, we obtain

$$\mathbf{P}(Y \le r) \le \mathbf{P}\left(\sum_{\vartheta \in \Gamma^*} e^{-c\alpha n} Y^\vartheta < r\right) + \mathbf{P}\left(m_n < e^{-cn}\right).$$

We now consider a crude (stochastic) bound on N which turns out to be sufficient: as $\nu(\{0\}) = 0$, we may alternatively think of the branching process as a path from the root to a leaf in level n such that all nodes on the path produce additional offspring in the next generation according to a copy of $\#E_J$, and all nodes created in this way reproduce according to ν. Upon keeping only a single node in the offspring of the latter set of particles, we can bound N from below by a random variable with a binomial distribution with parameters n and $\nu([2,\infty))$ hereafter denoted by $\mathrm{Bin}(n, \nu([2,\infty)))$. Thus, for $0 < \gamma < \nu([2,\infty))$, we have

$$\mathbf{P}(N < \lceil \gamma n \rceil) \ \le \ \mathbf{P}\left(\mathrm{Bin}(n, \nu([2,\infty))) < \lceil \gamma n \rceil\right) \ \le \ Ce^{-\varepsilon n},$$

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

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

With c large enough, the right-hand side is $O(e^{-\delta' n})$ for some $\delta' > 0$. Hence, with $n = \lceil C\log(1/r)\rceil$ and independent copies $Y_1, \dots, Y_n$ of Y, we obtain

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

Choosing $C < (\alpha c)^{-1}$ the exponent in the first term actually tends to $+\infty$ as $r \to 0$; hence this yields the assertion for $d(\rho, \zeta)$.

ii) The lower bound on $d(\rho, \bar{\zeta})$. We proceed similarly and write $\bar{\Gamma}$ for the set of nodes on the path in Γ^n containing those $\vartheta \in \Theta_n$ with $\varphi_n^\circ(\mathcal{T}^\vartheta)$ intersecting $[\![\rho, \bar{\zeta}]\!]$. Let $\vartheta^* = 11\dots1$ denote the root of Γ^n. Upon disregarding the contribution of $\mathcal{T}^{\vartheta^*}$, we have the following inequality which is the analogue of (5.3) ($\ge_{\mathbb{P}}$ denotes stochastic order):

$$d(\rho, \bar{\zeta}) \ge_{\mathbb{P}} \sum_{\vartheta \in \bar{\Gamma} \setminus \{\vartheta^*\}} \mathcal{V}(\vartheta)^\alpha Y^\vartheta. \tag{5.4}$$

Let $\bar{N} = \#\bar{\Gamma} - 1$. Then, $\bar{N}$ is distributed as the number of particles of type 2 in generation n in a two-type branching process, where each particle reproduces independently with the following dynamics: the process starts at time 0 with a type 1 particle which remains alive forever. We can think of this particle as the subtree in the decomposition

whose corresponding image contains $\bar{\zeta}$. Let $\bar{J}$ be a random variable satisfying $\mathbf{P}(\bar{J} = j \,|\, \mathcal{R}, \mathcal{S}) = \mathcal{R}_j$ for $j = 1, \dots, K$. In each generation, the immortal particle gives birth to an independent number of type 2 particles distributed as $\#E_{\bar{J}}$ corresponding to those subtrees intersecting $[\![\rho, \bar{\zeta}]\!]$ which do not contain $\bar{\zeta}$ and arise in the decomposition of the subtree associated with the immortal particle. Finally, since particles corresponding to (images of) subtrees intersecting $[\![\rho, \bar{\zeta}]\!]$ which were created on earlier levels are further decomposed, every type 2 particle generates type 2 offspring according to ν. (Type 2 particles do not generate offspring of type 1.)

Similarly to the bound on N derived above, upon only keeping track of type 2 children of the immortal particle, we can bound $\bar{N}$ from below by a random variable with a binomial distribution with parameters $n, \mathbf{P}\left(E_{\bar{J}} \geq 2\right)$. As $\mathbf{P}(E_{\bar{J}} \geq 2) > 0$, for $\gamma > 0$ sufficiently small, we find that $\mathbf{P}(\bar{N} < \lceil \gamma n \rceil)$ decays exponentially in n. The claim now follows as above.

iii) The lower bound on $d(\zeta, \bar{\zeta})$. Write Γ' for the path of nodes in Γ^n whose associated subtree (image) intersects $[\![\zeta, \bar{\zeta}]\!]$. Let ϑ_1, ϑ_2 be the two (non-necessarily distinct) end-points of Γ', and ϑ_3 their highest common ancestor. Similarly to (5.3) and (5.4), we have

$$d(\zeta, \bar{\zeta}) \geq_{\mathbb{P}} \sum_{\vartheta \in \Gamma' \setminus \{\vartheta_2, \vartheta_3\}} \mathcal{V}(\vartheta)^{\alpha} Y^{\vartheta}. \tag{5.5}$$

Describing $N' = \#\Gamma' - 2$ now leads to a branching process with as many types as kinds of portions of the path between ζ and $\bar{\zeta}$, which correspond to rescaled copies $d(\zeta, \bar{\zeta})$ (type 1), $d(\rho, \zeta)$ (type 2), $d(\rho, \bar{\zeta})$ (type 3), and $d(\zeta, \zeta')$ (type 4), where ζ' is an independent copy of ζ. The initial particle has type 1; once it reproduces, type 1 never reappears and there is an immortal particle of type 3; all other particles are of type 2 except the one corresponding to the portion of the path in ϑ_3 which as type 4; hence the representation in (5.5). Again, the number of children generated by the immortal particle up to time n grows linearly in n except on an event of exponentially small probability. We omit further details. □

The proof of Theorem 5 requires more details about the tree Γ^n: For a node $\vartheta \in \Gamma^n$, let $\mathcal{C}'(\vartheta)$ denote the (random) set of its children in Γ^n. By construction, if $\vartheta = \vartheta_1 \vartheta_2 \dots \vartheta_n$, then $\mathcal{C}'(\vartheta)$ only contains nodes of the form $\vartheta_1 \dots \vartheta_\ell \gamma 1 \dots 1$ with $0 \leq \ell < n$ and $2 \leq \gamma \leq K$ satisfying $\varpi_\gamma = \vartheta_{\ell+1}$ where $\vartheta_0 := 1$ (that is, γ is a child of $\vartheta_{\ell+1}$ in Γ). By $\mathcal{C}(\vartheta) \subseteq \mathcal{C}'(\vartheta)$ we denote the subset of children of ϑ where, for any $1 \leq \ell < n$, if $\vartheta_1 \dots \vartheta_\ell \gamma 1 \dots 1 \in \mathcal{C}'(\vartheta)$ for some γ, we keep only that child with minimal γ. (We also keep a child $\gamma 1 \dots 1$ with $\gamma \geq 2$ in $\mathcal{C}(\vartheta)$ if it exists in $\mathcal{C}'(\vartheta)$.) From Proposition 7 *iii)*, whose proof given in Section 5.3 does not make use of any results from the current section, we know that all the trees corresponding to nodes in $\mathcal{C}(\vartheta)$ are glued on the tree corresponding to ϑ at points that are distinct with probability one. By construction, $\mathcal{C}(\vartheta)$ is a maximal set with this property, and $\#\mathcal{C}(\vartheta) \leq n$. Informally, $\#\mathcal{C}(\vartheta)$ counts the number of distinct "exit points" of the subtree $\varphi_n^\circ(\mathcal{T}^\vartheta)$ in the decomposition of $\mathcal{T}$, that is, the points distinct from the root $\varphi_n^\circ(\rho^\vartheta)$ where geodesics may leave $\varphi_n^\circ(\mathcal{T}^\vartheta)$. See Figure 3 for an illustration of the construction and of the sets $\mathcal{C}(\vartheta)$. In particular, in this figure, we have $\mathcal{C}'(12) = \{13, 21, 41\}$, $\mathcal{C}(12) = \{13, 21\}$ while $\mathcal{C}'(24) = \mathcal{C}(24) = \{31\}$.

Proof of Theorem 5. Fix $n \in \mathbb{N}$ and $\gamma < 1/\alpha$. We need to show that $\dim_{\mathrm{h}}(\mathfrak{T}) \geq \gamma$ almost surely. Let $\mathcal{P} = \bigcup_{\vartheta \in \Theta_n} \varphi_n^\circ(\{\rho^\vartheta\})$. For $x \in \mathcal{T} \setminus \mathcal{P}$ let $\vartheta(x)$ be the unique node in Γ^n with $x \in \varphi_n^\circ(\mathcal{T}^{\vartheta(x)})$. Set $\mathfrak{T}(x) = \mathfrak{T}^{\vartheta(x)}$. Furthermore, let $H_x = d(x, \varphi_n^\circ(\rho^{\vartheta(x)}))$ be the height of x in (the image of) $\mathcal{T}(x)$ and $E_x = d(x, \mathcal{T} \setminus \varphi_n^\circ(\mathcal{T}(x)))$ the distance to exit (the image of)

$\mathcal{T}(x)$ from x. With $\mathcal{C}(x) := \mathcal{C}(\vartheta(x))$, for all $x \in \mathcal{T}$, we have

$$E_x = \begin{cases} \min_{\sigma \in \mathcal{C}(x)} d(x, \varphi_n^\circ(\mathcal{T}^\sigma)) & \text{if } \vartheta(x) = 1\dots1 \\ \min_{\sigma \in \mathcal{C}(x)} d(x, \varphi_n^\circ(\mathcal{T}^\sigma)) \wedge H_x & \text{if } \vartheta(x) \neq 1\dots1\,. \end{cases}$$

Recall that $B_r(x) = \{y \in \mathcal{T} : d(x,y) < r\}$ for $x \in \mathcal{T}$, $r > 0$. For any $x \in \mathcal{T} \setminus \mathcal{P}$ and $r > 0$, we have either $B_r(x) \subseteq \varphi_n^\circ(\mathcal{T}(x))$ or $E_x \leq r$.

We aim at using the mass distribution principle with the measure $\bar{\mu} = \mu^\emptyset \circ \pi_{\mathcal{Z}}^{-1}$ constructed thanks to Proposition 15. As in the previous proof, we let $\bar{\zeta}$ be a random variable drawn on $\mathcal{T}$ according to $\bar{\mu}$ (given $\mathfrak{T}$). Formally, we can let $\bar{\zeta} = \pi_{\mathcal{Z}}((F^\emptyset)^{-1}(\xi))$ for a random variable ξ with the uniform distribution on $[0,1]$ which is independent of all remaining quantities. As $\bar{\mu}$ has no atoms ($F^\emptyset$ is continuous and Proposition 7 *iii)*), we obtain

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

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

$$\mathbf{P}(E_{\bar{\zeta}} \leq r) \leq \mathbf{P}\left(H_{\bar{\zeta}} \leq r\right) + \mathbf{P}\left(\bigcup_{\ell=0}^{n-1} \{d(\bar{\zeta}, \varphi_n^\circ(\mathcal{T}^{\sigma_\ell})) \leq r, \sigma_\ell \in \mathcal{C}(\bar{\zeta})\}\right). \tag{5.6}$$

Let $\bar{\zeta}' = \pi_{\mathcal{Z}}((F^\emptyset)^{-1}(\xi'))$ and $\zeta'' = \pi_{\mathcal{Z}}(\xi'')$, where ξ', ξ'' are independent random variables with uniform distribution on $[0,1]$, independent of the remaining quantities. Then, for $\ell \in \{0, 1, \dots, n-1\}$, we have

$$\begin{aligned} \mathbf{P}\left(d(\bar{\zeta}, \varphi_n^\circ(\mathcal{T}^{\sigma_\ell})) \leq r, \sigma_\ell \in \mathcal{C}(\bar{\zeta})\right) &= \mathbf{P}(\mathcal{V}(\vartheta(\bar{\zeta}))^\alpha \cdot d(\bar{\zeta}', \zeta'') \leq r, \sigma_\ell \in \mathcal{C}(\zeta)) \\ &\leq \mathbf{P}(\mathcal{V}(\vartheta(\bar{\zeta}))^\alpha \cdot d(\bar{\zeta}', \zeta'') \leq r). \end{aligned}$$

Similarly, $H_{\bar{\zeta}}$ is distributed like $\mathcal{V}(\vartheta(\bar{\zeta}))^\alpha \cdot d(\rho, \bar{\zeta}')$, for $\bar{\zeta}'$ an independent copy of $\bar{\zeta}$. Let $\eta \in (0,1)$ be a parameter to be chosen later. Applying the union bound on the right-hand side of (5.6) yields

$$\begin{aligned} \mathbf{P}(E_{\bar{\zeta}} \leq r) \leq n\big\{&\mathbf{P}(\mathcal{V}(\vartheta(\bar{\zeta})) \leq r^{\eta/\alpha}) + \mathbf{P}(d(\rho, \bar{\zeta}) \leq r^{1-\eta}) \\ &+ \mathbf{P}(d(\bar{\zeta}, \zeta'') \leq r^{1-\eta})\big\}\,. \end{aligned}$$

As $\bar{\mu}(\varphi_n^\circ(\mathcal{T}(\bar{\zeta}))) = \mathcal{V}(\vartheta(\bar{\zeta}))$, combining the bounds and using Lemma 16, we see that there exists universal constants $\varepsilon_1 > 0$ and $C > 0$ such that

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

We now choose the parameters. Note that $-\log \mathcal{V}(\vartheta(\bar{\zeta}))$ is distributed like the sum of n independent copies of the random variable $-\log \sum_{i=1}^K \mathbf{1}_{\{\bar{J}=i\}} \mathcal{R}_i$; furthermore, as $\mathcal{R}^* := \sum_{i=1}^K \mathbf{1}_{\{\bar{J}=i\}} \mathcal{R}_i$ is stochastically larger than $\overline{\mathcal{R}}$, the tail bound on $\overline{\mathcal{R}}$ implies that $-\log \mathcal{R}^*$ has exponential moments, and hence that the expected value $q := \mathbf{E}[-\log \mathcal{R}_{\bar{J}}] < \infty$ governs the asymptotics. First, let $\eta \in (\gamma\alpha, 1)$. Then choose $\delta \in (\gamma/q, \eta/(q\alpha))$. Finally, let $n = n(r) = \lfloor -\delta \log r \rfloor$. 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 $\mathbf{P}(\mathcal{V}(\vartheta(\bar{\zeta})) > r^\gamma) \leq C_2 r^{\varepsilon_2}$ and $\mathbf{P}(\mathcal{V}(\vartheta(\bar{\zeta})) \leq r^{\eta/\alpha}) \leq C_2 r^{\varepsilon_2}$. Summarizing, there exists $C > 0$ (which may depend on all parameters but not on r), such that,

$$\mathbf{P}(\bar{\mu}(B_r(\bar{\zeta})) > r^\gamma) \leq C \log(1/r) \cdot (r^{\varepsilon_2} + r^{\varepsilon_1(1-\eta)}), \quad 0 < r < 1.$$

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

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

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

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

Thus, denoting $A = \{x \in \mathcal{T} : \limsup_{r\to 0} \bar{\mu}(B_r(x))/r^{\gamma} \leq 2\}$, we have $1 = \mathbf{P}\left(\bar{\zeta} \in A\right) = \mathbf{E}[\bar{\mu}(A)]$ implying $\bar{\mu}(A) = 1$ almost surely. From (2.14), it follows that, almost surely, $\dim_{\mathrm{h}}(\mathcal{T}) \geq \dim_{\mathrm{h}}(A) \geq \gamma$ which completes the proof. □

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

The proofs of the missing parts of Theorem 1, i.e., *iii)* and *iv)*, rely on the dynamics governing the number of points in the set $\mathcal{C}(\vartheta)$ as ϑ follows a path in Θ away from the root. The following lemma is straightforward from the construction, since the exit points counted by $\#\mathcal{C}(\vartheta)$ are chosen on $\mathcal{T}^{\vartheta}$ according to the mass measure μ^{ϑ}. Recall that, for $\vartheta \in \Theta$, we write $\mathcal{L}(\vartheta)$ for the Lebesgue measure of the set Λ_{ϑ} (see (3.3)). Recall also that $\Gamma^o = \Gamma \setminus \partial\Gamma$, where $\partial\Gamma$ is the set of leaves of Γ.

Lemma 17. i) Let $\varepsilon_1, \varepsilon_2, \ldots \in [K]$ and, for each $n \geq 1$, $\vartheta_n = \varepsilon_1 \ldots \varepsilon_n \in \Theta_n$. Then, $\mathcal{L}(\vartheta_{n+1}) = \mathcal{L}(\vartheta_n) \cdot \mathcal{S}^{\vartheta_n}_{\varepsilon_{n+1}}$. The sequence $(\#\mathcal{C}(\vartheta_n), \mathcal{L}(\vartheta_n)), n \geq 0$ is a Markov chain on $\mathbb{N} \times [0,1]$ starting at $(0,1)$, whose evolution can be described as follows: given $(\#\mathcal{C}(\vartheta_n), \mathcal{L}(\vartheta_n))$, we have

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

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 $(\tilde{\mathcal{C}}_n, \tilde{\mathcal{L}}_n) := (\#\mathcal{C}(\tilde{\vartheta}_n), \mathcal{L}(\tilde{\vartheta}_n))$. Then, the sequence $(\tilde{\mathcal{C}}_n, \tilde{\mathcal{L}}_n), n \geq 0$ is a homogeneous Markov chain on $\mathbb{N} \times [0,1]$ starting at $(0,1)$, whose evolution can be described as follows: given $(\tilde{\mathcal{C}}_n, \tilde{\mathcal{L}}_n)$, we have

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

where $(\mathcal{S}^n)_{n\geq 0}$ is a family of independent 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)*. Let $A = \{s \in [0,1] : \mathcal{Z}(s) = 0\}$ denote the zero-set of $\mathcal{Z}$. For $n \geq 1$, let $\Lambda^n = \Lambda_{\vartheta_n}$ where $\vartheta_n = 1\ldots 1$. Then for every $n \geq 0$, $\varphi_n^{\circ}(\mathcal{T}^{\vartheta_n})$ is the subtree that contains the root ρ of $\mathfrak{T}$. Furthermore, Λ^n is the union of $\#\mathcal{C}(\vartheta_n) + 1$ disjoint intervals. Clearly, $(\Lambda^n)_{n\geq 1}$ is decreasing and we set $\Lambda := \bigcap_{n\geq 1} \Lambda^n$. Since $\mathbf{E}[\mathcal{S}_1] < 1$, Lemma 17 *i)* and a routine drift argument (see, e.g., Chapter 8 of [39]) shows that, almost surely, $\#\mathcal{C}(\vartheta_n) = 1$ infinitely often, and thus Λ^n consists of only two intervals for infinitely many n. As $\{0,1\} \in A$, 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 $\Lambda = \{0,1\}$ almost surely.

Now, for any $s \notin \Lambda$, there is some n large enough for which $s \notin \Lambda^n$ and Λ^n consists of two intervals. Then, since the path between ρ and the projection of s in the tree must cross $\varphi_n^{\circ}(\mathcal{T}^{\vartheta_n})$, $\mathcal{Z}(s)$ is at least a rescaled copy of $\mathcal{Z}(\xi)$, for a uniform random variable ξ which is independent of $\mathcal{Z}$. However, in the context of fixed point equation (2.6), we

have already seen that $\mathcal{Z}(\xi) > 0$ almost surely (see (2.7)). It follows that $A \subseteq \Lambda$, and since $\Lambda = \{0, 1\}$ the root ρ of $\mathfrak{T}$ is a leaf.

ii) Since $\mu_{\mathcal{Z}}$ has full support, this reduces to showing that for a random point ζ sampled from $\mu_{\mathcal{Z}}$ and any $\varepsilon > 0$, with probability one, there exists some $x \in \mathcal{T} = \mathcal{T}_{\mathcal{Z}}$ such that ζ lies in the subtree of $\mathcal{T}$ rooted at x, namely $\mathcal{T}^{\uparrow}(x) := \{u \in \mathcal{T} : x \in [\![\rho, u]\!]\}$, and $\mu_{\mathcal{Z}}(\mathcal{T}^{\uparrow}(x)) < \varepsilon$. To prove this, we choose $\zeta = \pi_{\mathcal{Z}}(\xi)$ with ξ as in Lemma 17 *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 Lemma 17 *ii)*. In particular, $\zeta \in \varphi_n^{\circ}(\mathcal{T}_{\mathcal{Z}^{\tilde{\vartheta}_n}})$. Thus, ζ lies in the subtree of $\mathcal{T}$ rooted at $\varphi_n^{\circ}(\rho^{\tilde{\vartheta}_n})$, and thus it suffices to show that for any $\varepsilon > 0$,

$$A_\varepsilon = \inf\{n \geq 0 : \mu_{\mathcal{Z}}(\varphi_n^{\circ}(\mathcal{T}_{\tilde{\vartheta}_n})) < \varepsilon\} < \infty\,.$$

Observe that, if $\tilde{\mathcal{L}}_n = \mu_{\mathcal{Z}}(\varphi_n^{\circ}(\mathcal{T}_{\tilde{\vartheta}_n})) < \varepsilon$ and $\tilde{\mathcal{C}}_n = 0$, then $A_\varepsilon \leq n$. As in the proof of *i)* above, since $\mathbf{E}[\mathcal{S}_{J_n}^n] < 1$, a classical drift argument shows that $(\tilde{\mathcal{C}}_n)_{n\geq 0}$ is positive recurrent, and in particular, $\tilde{\mathcal{C}}_n = 0$ infinitely often. Then, for some subsequence $(n_i)_{i\geq 1}$ with $n_i \to \infty$, we have $\tilde{\mathcal{C}}_{n_i} = 0$. But, by Lemma 17, $\tilde{\mathcal{L}}_n \to 0$ 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 $A_\varepsilon \leq n_{i_0} < \infty$, 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 $\mathcal{Z}$ in (2.10). Hence,

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

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

Proof of Theorem 1 iii) and iv). The point *iii)* was established in the proof of Proposition 7 *i)*. The argument we have used above for the proof of Proposition 7 *ii)* also implies *iv)*, that is, almost surely, $\mathcal{Z}$ is nowhere monotonic: indeed, if this were not the case, then, with positive probability, a randomly chosen point would be contained in an interval where $\mathcal{Z}$ is monotonic. In particular, for some $\varepsilon > 0$ and a uniformly chosen point ξ, $\mathcal{Z}(\xi)$ would fall in a interval of length at least ε on which $\mathcal{Z}$ is monotonic with positive probability. Recall the quantites $\tilde{\vartheta}_n$ and $\tilde{\mathcal{C}}_n$ from the proof of Proposition 7 *ii)*. Let $I \subseteq \mathbb{N}$ be the (a.s. infinite) set of indices n for which $\tilde{\mathcal{C}}_n = 0$. For $n \in I$, the set $\Lambda_{\tilde{\vartheta}_n} \subseteq [0, 1]$ consists of a single half-open interval, and $\xi \in \Lambda_{\tilde{\vartheta}_n}$. For every $n \in I$ large enough, we have $\mathcal{L}(\tilde{\vartheta}_n) < \varepsilon/2$. But $\mathcal{Z}$ is non-monotonic on $\Lambda_{\tilde{\vartheta}_n}$ since $\xi \in \Lambda_{\tilde{\vartheta}_n}$ and $\mathcal{Z}(\inf \Lambda_{\tilde{\vartheta}_n}) = \mathcal{Z}(\sup \Lambda_{\tilde{\vartheta}_n}) < \mathcal{Z}(\xi)$ almost surely by Theorem 1 *iii)*. This concludes the proof. □

Proof of Proposition 7 iv). Let $\mathfrak{T} = \mathfrak{T}_{\mathcal{Z}}$. First note that $2 \in \mathscr{D}(\mathcal{T})$ almost surely since μ has no atoms, the set of branch points of $\mathcal{T}$ is at most countable, and $\mathcal{T}$ is not reduced to a point. Further, since $\mathcal{Z}$ is nowhere 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).

Next, for every $k \in \mathscr{D}(\Gamma), k \geq 3$, almost surely, there exists a point $x \in \mathcal{T}$ with degree k. This follows immediately from the fixed point equation satisfied by $\mathfrak{T}$ and the fact that the root has degree 1 and μ is concentrated on the leaves (see parts *i)* and *ii)* of the proposition). For example, if $a \in \Gamma$ has degree k, then (the image of) the root of $\mathcal{T}_{a+1}$ is a point in $\mathcal{T}$ with degree k. The main part of the proof consists in showing that, if $k + 1 \notin \mathscr{D}(\Gamma), k \geq 3$, then there do not exist points in $\mathcal{T}$ with degree $k + 1$. To this end, let $(U_i)_{i\geq 1}$ be a family of independent r.v. with the uniform distribution on $[0, 1]$ and $\zeta_i = \pi_{\mathcal{Z}}(U_i)$ be the corresponding test points on $\mathcal{T}$. Further, let $U_0 = 0$ and $\zeta_0 = \rho$.

Since μ has full support on $\mathcal{T}$, if there exists $x \in \mathcal{T}$ with degree $k+1$, then all 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 a point with degree $k+1$. Let $\mathcal{I}$ be the collection of points in $\mathcal{T}$ that are contained in some intersection $[\![\zeta_i, \zeta_j]\!] \cap [\![\zeta_{i'}, \zeta_{j'}]\!]$, for distinct elements $i, i, j, j' \in \{0, \dots, k\}$. Let $T \ge 1$ be the minimal integer such that, for some pairwise distinct nodes $\sigma_i, i = 0, \dots, k$ in Θ_T, we have $U_i \in \Lambda_{\sigma_i}, i = 0, \dots, k$. Clearly, $T < \infty$ almost surely as $\max_{\sigma \in \Theta_n} \mathcal{L}(\sigma) \to 0$ almost surely. Let us denote the most recent common ancestor of the nodes $\sigma_i, i = 1, \dots, k$ by σ^*. (That is, σ^* is the node of maximal depth such that all $\sigma_i, i = 1, \dots, K$ are contained in its subtree.) Write $\sigma^* = \bar{\sigma}^* \varepsilon$ with $\bar{\sigma}^* \in \Theta_{T-1}$ and $\varepsilon \in [K]$. $\mathcal{I}$ is a singleton if and only if the nodes $\sigma_i, i = 1, \dots, k$ lie in subtrees of pairwise distinct nodes of the form $\bar{\sigma}^* \gamma$, where $\varpi_\gamma = \varepsilon$. In particular, x is equal to the roots of trees $\mathcal{T}_\sigma, \sigma \in \{\bar{\sigma}^* \gamma : \varpi_\gamma = \varepsilon\}$ under the canonical surjection $\sqcup_{\sigma \in \Theta_T} \mathcal{T}_\sigma \to \mathcal{T}$. As roots have degree 1 and the mass measure is concentrated on the leaves, the degree of x must be equal to the degree of ε in Γ. In particular, the degree of x lies in the set $\mathscr{D}(\Gamma)$. Summarizing, almost surely, the set $\mathcal{I}$ does not exist of a singleton with degree $k+1 \notin \mathscr{D}(\Gamma)$ which concludes the argument. $\sup \mathscr{D}(\mathcal{T}) < \infty$ almost surely can be deduced easily. For, if this was not the case, for any natural number $L \ge 3$, the above construction with $k = L-1$ would show that, with positive probability, there exists $L' \ge L$ such that $\mathcal{I}$ is reduced to a singleton of degree $L' \in \mathscr{D}(\Gamma)$. Choosing L strictly larger than the maximal degree in Γ contradicts this fact. It remains to prove that, if $\max \mathscr{D}(\Gamma) \ge 4$ and $3 \notin \mathscr{D}(\Gamma)$, then, almost surely, there are no nodes of degree 3 in $\mathcal{T}$. Set $\mathcal{I} = [\![\zeta_1, \zeta_2]\!] \cap [\![\zeta_1, \zeta_3]\!] \cap [\![\zeta_2, \zeta_3]\!]$. Almost surely, $\mathcal{I}$ is a singleton, say x. There exists a node with degree 3 in $\mathcal{T}$ with positive probability if and only if, with positive probability, x has degree 3. Let $\vartheta_1, \vartheta_2 \in [K]$ such that $U_i \in \Lambda_{\vartheta_i}, i = 1, 2$. Let σ^* be their most recent common ancestor. If $U_1, U_2 \in \Lambda_{\sigma^*}$, then let $U_i' = \varphi_{\sigma^*}(U_i)$, where φ_{σ^*} is defined in (2.8). If $U_1 \in \Lambda_{\sigma^*}$ and $U_2 \notin \Lambda_{\sigma^*}$, then set $U_1' = \varphi_{\sigma^*}(U_1)$ and $U_2' = \xi_{\sigma^*}$. Proceed analogously if the roles of U_1, U_2 are interchanged. Note that, given that one of these cases occurs, the random variables U_1', U_2' are independent, uniformly distributed on $[0, 1]$ and independent of $\{(\mathcal{R}^\vartheta, \mathcal{S}^\vartheta, \Xi^\vartheta) : |\vartheta| > 0\}$. Let $\zeta_i' = \pi_{\mathcal{Z}^{\sigma^*}}(U_i'), i = 1, 2, 3$, where $U_3' = 0$. Note that, x has degree 3 if and only if the node $[\![\zeta_1', \zeta_2']\!] \cap [\![\zeta_1', \zeta_3']\!] \cap [\![\zeta_2', \zeta_3']\!]$ in $\mathcal{T}_{\sigma^*}$ has degree 3. By the same arguments as above, if $U_1, U_2 \notin \Lambda_{\sigma^*}$, then the degree of x is equal to the degree of σ^* in Γ and therefore at least 4. It follows that

$$\mathbf{P}\left(\deg(x) = 3\right) = \mathbf{P}\left(\deg(x) = 3\right)\left(1 - \mathbf{P}\left(U_1, U_2 \notin \Lambda_{\sigma^*}\right)\right)$$

As $\mathbf{P}\left(U_1, U_2 \notin \Lambda_{\sigma^*}\right) > 0$ it follows that $\mathbf{P}\left(\deg(x) = 3\right) = 0$ which concludes the proof. □

5.4 Optimal Hölder exponents

The proof of Corollary 6 merely consists in putting together the information we have gathered in 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. On the one hand, we have $\underline{\dim}_{\mathrm{m}}(\mathcal{T}_{\mathcal{Z}}) \ge \dim_{\mathrm{h}}(\mathcal{T}_{\mathcal{Z}}) \ge \alpha^{-1}$ because of Theorem 5; on the other hand, $\overline{\dim}_{\mathrm{m}}(\mathcal{T}_{\mathcal{Z}}) \le \alpha^{-1}$.

(b) By Theorem 1, $\mathcal{Z}$ is nowhere constant, hence (2.15) holds and $\alpha = \omega_{\mathcal{Z}}$. By definition of $\omega_{\mathcal{Z}}$, for any $\gamma < \omega_{\mathcal{Z}} = \alpha$, almost surely, there exists a process $\tilde{\mathcal{Z}}$ time-change equivalent to $\mathcal{Z}$ with γ-Hölder continuous paths.

(c) This follows immediately from part *(a)* since the existence of a process $\mathcal{Z}'$ whose sample paths are γ-Hölder continuous with positive probability for $\gamma > \alpha$ would contradict the statement there. □

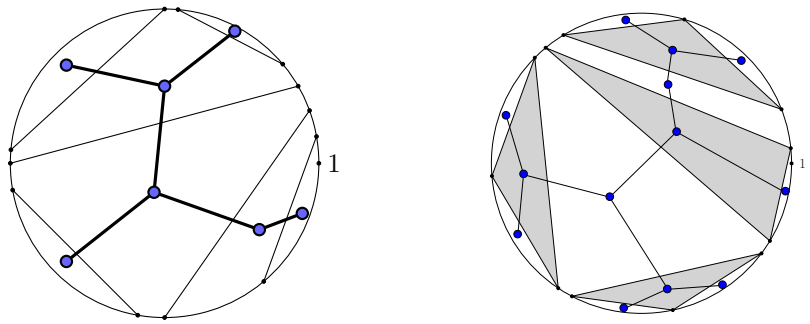

Figure 4: On the left, a lamination and the corresponding rooted dual tree. Distances in the tree correspond to the number of chords separating the fragments in the lamination. On the right, an example of 3-angulation, together with its dual tree. The tree is rooted at the node containing the point $(1,0) \in \partial\mathscr{D}$. 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.

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.

6.1 The Brownian continuum random tree

The Brownian CRT $\mathcal{T}_{\mathbf{e}}$ 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 [4, Theorem 2], but also binary unordered unlabeled trees (Otter trees) [38], random trees with a prescribed degree sequence [13], general unordered unlabeled trees (Pólya trees) [42], unlabeled unrooted trees [49], and random graphs from subcritical classes [43] to name a few examples. (See also [35].)

Theorem 18 (Aldous [4], see also [35])**.** 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 (2.11) 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 $\frac{1}{2}t^{-1/2}\mathbf{1}_{[0,1]}(t)$.

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 [14, 19]. They encode the trees that are the planar dual of the limit triangulation in the same sense that the Brownian CRT is the dual of the limit uniform triangulation of the disk studied by Aldous [6, 7]. We now proceed to the precise definitions.

The unit disk $\mathscr{D} := \{x \in \mathbb{R}^2 : \|x\| \leq 1\}$ 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

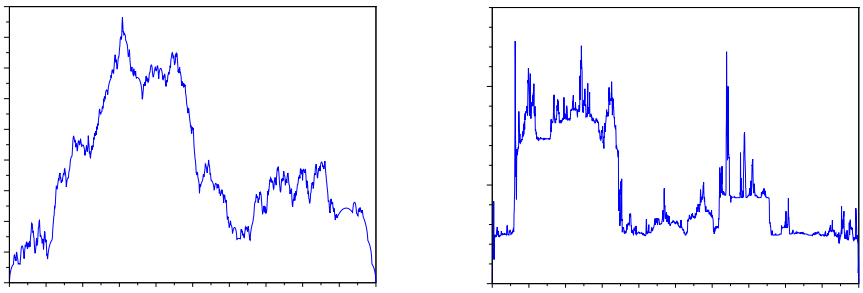

Figure 5: A comparison of the processes $\mathscr{Z}$ and $\mathscr{H}$. On the left, the process $\mathscr{Z}$ encoding the dual tree of the self-similar recursive triangulation of the disk. On the right, the process $\mathscr{H}$ encoding the dual tree of the homogeneous recursive triangulation of the disk. (The scales are not given since they are irrelevant.)

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 $(1, 0)$. (See Figure 4). It has been proved in [14] 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 (2.10). More precisely, for $\beta := (\sqrt{17} - 3)/2$, with respect to the Gromov–Hausdorff distance and as $n \to \infty$, we have

$$(T_n, n^{-\beta/2} d_n) \to (\mathcal{T}_{\mathscr{Z}}, d_{\mathscr{Z}}), \tag{6.1}$$

for the unique random excursion $\mathscr{Z}$ satisfying (2.12) 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 [14].)

Corollary 8 *(b)*, Corollary 9 *(b)* and Corollary 10 *(b)* follow from (2.12), Theorems 1, 4 and 5 since $\overline{\mathcal{R}}$ has the uniform distribution on $[0, 1]$.

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 [19], 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 (6.1) 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 [14, 19].)

6.3 Homogeneous recursive triangulation of the disk

We have a completely different situation if we consider a partition of the disk $\mathscr{D}$ 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 [14] 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

$$(T_n^h, n^{-1/3} d_n) \to (\mathcal{T}_{\mathscr{H}}, d_{\mathscr{H}}), \tag{6.2}$$

where $\mathscr{H}$ is the unique random excursion satisfying (2.13) and $\mathbf{E}[\mathscr{H}(\xi)] = 1/\Gamma(4/3)$. (No characterization of $\mathscr{H}$ was given in [14]; one is given in Corollary 9 *(c)*.) The rescaling $n^{-1/3}$ in (6.2) suggests that 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 5 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 $\dim_{\mathrm{m}}(T_{\mathscr{H}}) = \dim_{\mathrm{h}}(T_{\mathscr{H}}) = 3$ 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 supplementary material.

Proposition 19. 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 [20]. 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 [14] and [19] yield detailed information on the height processes of the corresponding dual trees and their limits. For example, in [14], 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 [19] (up to a multiplicative constant). Most of these 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 20 and 21 could be given by lengthy implicit formulas but they are of no particular relevance.)

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 one. The dual tree $\mathcal{T}_n$ is defined analogously to the case $k = 2$ upon identifying fragments in the decomposition with nodes in $\mathcal{T}_n$. It is endowed with the graph distance d_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* (all fragments are essential for $k = 2$.) See Figure 4 for an illustration. Keeping the notation introduced in [14], we denote by $\mathcal{C}_n(s)$ the depth of the node associated to the fragment covering the point s in the tree $\mathcal{T}_n$. Here, and subsequently, we identify the unit interval with $\partial\mathscr{D}$ through $s \mapsto (\cos 2\pi s, \sin 2\pi s)$. (We do not indicate k in the notation for the height

functions.) Denote the first inserted k points in increasing order by $0 \le U_1 \le \ldots \le U_k \le 1$ 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 of mass 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 $n-1$ and Δ_i^k. In particular, we have, almost surely,

$$\frac{I_n}{n} \to (\Delta_1^k, \ldots, \Delta_k^k). \tag{6.3}$$

Proposition 20. 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 a Poisson point process with unit intensity on $\mathbb{R}^+$. 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 [19] 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 [10], we deduce $t^{-1/k}\mathcal{N}_t \to c_k$ in probability and in L_2, with c_k as in the proposition In particular, $\tau_n^{-1/k}\mathcal{N}_{\tau_n} \to c_k$ in probability as $n \to \infty$. In order to obtain 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 (which has a 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,

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

where $(\tilde{N}_n^{(1)})_{n\ge1}, \ldots, (\tilde{N}_n^{(k)})_{n\ge1}$ are independent copies of $(N_n)_{n\ge1}$, independent of I_n. Using (6.3), it is easy to prove that $\tilde{N}_n$ is bounded in $L_m, m \ge 1$ by induction over m since we have already shown it for $m = 1, 2$. □

By construction, the random process $(\mathcal{C}_n(s))_{s\in[0,1]}$ satisfies the following recurrence in distribution on the space of càdlàg functions endowed with the Skorokhod J_1-topology:

$$\begin{aligned}\mathcal{C}_n(\,\cdot\,) \stackrel{d}{=} {} & \mathbf{1}_{[0,U_1)}(\,\cdot\,)\mathcal{C}_{I_n^{(1)}}^{(1)}\left(\frac{\cdot}{\Delta_1}\right) + \mathbf{1}_{[U_k,1]}(\,\cdot\,)\mathcal{C}_{I_n^{(1)}}^{(1)}\left(\frac{\cdot - U_k}{\Delta_1}\right) \\ & + \sum_{i=2}^{k-1} \mathbf{1}_{[U_{i-1},U_i]}(\,\cdot\,)\left(\mathcal{C}_{I_n^{(i)}}^{(i)}\left(\frac{\cdot - U_{i-1}}{\Delta_i}\right) + \mathcal{C}_{I_n^{(1)}}^{(1)}(\xi^*)\right).\end{aligned} \tag{6.4}$$

Here $(\mathcal{C}_i^{(1)}(\cdot))_{i\geq 0}, \ldots, (\mathcal{C}_i^{(k)}(\cdot))_{i\geq 0}$ are independent copies of $(\mathcal{C}_i(\cdot))_{i\geq 0}$ independent of $(U_1, \ldots, U_k, I_n)$. The first step of our analysis is to investigate $\mathcal{C}_n(\xi)$ for uniformly chosen point ξ. In [14], an explicit expression for the mean was obtained by solving the underlying recursion. Here, we proceed as in [19] relying on results from fragmentation theory. Subsequently, let $\alpha_k \in (0,1)$ be the unique solution to

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

(Note that $g_k(x)$ decreases in x, $g_k(0) > 1$ and $g_k(1) < 1$. Thus, α_k exists.)

Proposition 21. Let $k \geq 2$. As $n \to \infty$, in probability and with convergence of all moments, we have $n^{-\alpha_k/k}\mathcal{C}_n(\xi) \to X_k$ 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. 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 $\mathcal{C}_t(\xi) = 2(E_t(\xi) - 1)$. As explained in [19] in the case $k = 2$, 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 [10, 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 [10, 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 $m_k \in \mathbb{C}_{\text{ex}}$ with $\mathbf{E}\left[m_k(\xi)\right] = \kappa_k$. Thus, from (6.5), it follows that, if $\mathcal{Y}_n(s) \to \mathscr{Z}(s)$ for some continuous process $\mathscr{Z}$, then the limit should have mean function m_k/κ_k and satisfy

$$\begin{aligned} \mathscr{Z}(\,\cdot\,) \stackrel{d}{=} {} & \mathbf{1}_{[0,U_1)}(\,\cdot\,)\Delta_1^{\alpha_k}\mathscr{Z}^{(1)}\left(\frac{\cdot}{\Delta_1}\right) + \mathbf{1}_{[U_k,1]}(\,\cdot\,)\Delta_1^{\alpha_k}\mathscr{Z}^{(1)}\left(\frac{\cdot - U_k}{\Delta_1}\right) \\ & + \sum_{i=2}^{k}\mathbf{1}_{[U_{i-1},U_i]}(\,\cdot\,)\left(\Delta_i^{\alpha_k}\mathscr{Z}^{(i)}\left(\frac{\cdot - U_{i-1}}{\Delta_i}\right) + \Delta_1^{\alpha_k}\mathscr{Z}^{(1)}\left(\xi^*\right)\right), \end{aligned} \tag{6.5}$$

where $\mathscr{Z}^{(1)}(\cdot), \ldots, \mathscr{Z}^{(k)}(\cdot)$ are independent copies of $\mathscr{Z}(\cdot)$ independent of $(U_1, \ldots, U_k)$. This fixed point equation is of type (2.10) where $K = k, L = k+1, \varpi_3 = \ldots = \varpi_k = 1, \mathcal{R} = \mathcal{S} \sim \text{Dirichlet}(2, 1, \ldots, 1)$. Let $\mathscr{Z}$ be the unique process (in distribution) solving (6.5) with $\mathbf{E}\left[\mathscr{Z}(\xi)\right] = \kappa_k$ whose existence is guaranteed by Theorem 1. (We use the same notation for the limit process as in Section 6.2 without indicating the choice of k.) The verification of $\mathcal{Y}_n \to \mathscr{Z}$ in distribution in the space of càdlàg functions can be worked out by the same arguments as in [14, 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 $\mathscr{Z}$ satisfying (6.5) 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 $\mathcal{Y}_n(\psi) \to \mathscr{Z}(\psi)$ 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 $\mathbf{E}\left[\|\mathcal{Y}_n - \mathscr{Z}\|^m\right] \to 0$ for all $m \geq 1$. (The almost sure convergence in [14] 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 [10], compare the discussion of Theorem 1 there.) Summarizing, we obtain the following result.

Theorem 22. Let $k \geq 3$. In distribution with respect to the Gromov–Hausdorff topology, we have

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

In distribution, the process $\mathscr{Z}$ is the unique continuous excursion solving (6.5) up to a multiplicative constant. Almost surely, $\dim_{\mathrm{m}}(\mathcal{T}_{\mathscr{Z}}) = \dim_{\mathrm{h}}(\mathcal{T}_{\mathscr{Z}}) = \alpha_k^{-1}$. For any $\gamma < \alpha_k$, almost surely, there exists a process $\tilde{\mathscr{Z}}$ equivalent to $\mathscr{Z}$ which has γ-Hölder continuous paths.

Remark. **1)** For $k = 2$, almost surely, the process $\mathscr{Z}$ has γ-Hölder continuous path for any $\gamma < \alpha_2 = \beta$ [19, Theorem 1.1]. We think that this remains true for all $k \geq 3$, that is, the function $\mathscr{Z}$ is a good encoding of the tree. However, we do not pursue this here.
2) Let L_n be the set of chords inserted at time n. By Proposition 7, $\mathscr{D}(\mathcal{T}_{\mathscr{Z}}) = \{1, 2, k\}$, 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.

For $k \geq 3$, we have no explicit expression for $m_k = \mathbf{E}[\mathscr{Z}(t)]$. It follows from (6.5) that m_k is the unique continuous excursion on $[0, 1]$ with $\mathbf{E}[m_k(\xi)] = \kappa_k$ such that $m_k(t) = \mathbf{E}[\Phi(m_k, \ldots, m_k, \Delta, \Delta, \Xi)(t)], t \in [0, 1]$ where $K = k, L = k+1, \varpi_3 = \ldots = \varpi_k = 1$ and $\Delta \sim \text{Dirichlet}(2, 1, \ldots, 1)$.

Using this observation 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 and the proofs.

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 $\mathcal{T}_n^h, \mathcal{C}_n^h, U_1, \ldots, U_k, \Delta_1 = 1 - U_k + U_1, \Delta_i = U_i - U_{i-1}, 2 \leq i \leq K, \xi^* = U_1/\Delta_1$ 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, $I_n/n \to (\tilde{\Delta}_1, \ldots, \tilde{\Delta}_k)$, where $\tilde{\Delta} = (\tilde{\Delta}_1, \ldots, \tilde{\Delta}_k) \sim \text{Dirichlet}(1/(k-1), \ldots, 1/(k-1))$. By construction, the random process $(\mathcal{C}_n^h(s))_{s\in[0,1]}$ satisfies a recursion analogous to (6.4), the only difference being the distribution of $(U_1, \ldots, U_k, I_n)$.

Proposition 23. Let $k \geq 2$. 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 $\mathcal{N}_t, t \geq 0$, forms a continuous-time branching process with offspring distribution δ_k. 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

$$\begin{aligned} &e^{-\tau_n(k-1)/(k+1)} E^h_{\tau_n}(\xi) \\ &= e^{-\tau_n(k-1)/(k+1)}((k-1)n)^{1/(k+1)} \cdot ((k-1)n)^{-1/(k+1)} E^h_{\tau_n}(\xi), \end{aligned}$$

and noting that the random variables $\tau_n, E^h_{\tau_n}(\xi)$ are independent, it follows that $((k-1)n)^{1/(k+1)} E^h_{\tau_n}(\xi) \to \mathcal{E}'$ a.s. and in mean where $\mathcal{E} = \mathcal{E}'\mathcal{N}^{1/(k+1)}$. The claimed convergence follows from the identity $\mathcal{C}^h_n(\xi) = 2(E^h_{\tau_n}(\xi) - 1)$. Convergence of moments can be deduced as in the recursive model. □

For $s \in [0,1]$, let $\mathcal{Y}^h_n(s) = \mathcal{C}^h_n(s)/\mathbf{E}[\mathcal{C}^h_n(\xi)]$. A limit $\mathscr{H}(s)$ of $\mathcal{Y}^h_n(s)$ should satisfy $\mathbf{E}[\mathscr{H}(\xi)] = \kappa^h_k$ and

$$\begin{aligned} \mathscr{H} &\stackrel{d}{=} \Delta_1^{1/(k+1)} \left(\mathbf{1}_{[0,U_1)}(\,\cdot\,) \tilde{\mathscr{Z}}^{h,(1)} \left(\frac{\cdot}{\Delta_1} \right) + \mathbf{1}_{[U_k,1]}(\,\cdot\,) \tilde{\mathscr{Z}}^{h,(1)} \left(\frac{\cdot - U_k}{\Delta_1} \right) \right) \\ &+ \sum_{i=2}^{k} \mathbf{1}_{[U_{i-1},U_i]}(\,\cdot\,) \left(\tilde{\Delta}_i^{1/(k+1)} \mathscr{Z}^{h,(i)} \left(\frac{\cdot - U_{i-1}}{\Delta_i} \right) + \tilde{\Delta}_1^{1/(k+1)} \mathscr{Z}^{h,(1)}(\xi^*) \right), \end{aligned} \tag{6.6}$$

where $\mathscr{Z}^{h,(1)}, \ldots, \mathscr{Z}^{h,(k)}$ are independent copies of $\mathscr{H}$ independent of $(U_1, \ldots, U_k), \tilde{\Delta}$. This fixed point equation is of type (2.10) where $K = k, L = k+1, \varpi_3 = \ldots = \varpi_k = 1, \mathcal{R} \sim \text{Dirichlet}(1,1,\ldots,1)$ and $\mathcal{S} \sim \text{Dirichlet}(2,1,\ldots,1)$, and $\mathcal{R}$ and $\mathcal{S}$ are independent. (We use the same notation for the limit process as in Section 6.3.) As in the recursive model, one can prove the following theorem.

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

$$(\mathcal{T}^h_n, n^{-1/(k+1)} d^h_n) \to (\mathcal{T}_{\mathscr{H}}, d_{\mathscr{H}}).$$

In distribution, the process $\mathscr{H}$ is the unique continuous excursion satisfying (6.6) up to a multiplicative constant. Almost surely, $\dim_{\mathrm{m}}(\mathcal{T}_{\mathscr{H}}) = \dim_{\mathrm{h}}(\mathcal{T}_{\mathscr{H}}) = k+1$. For any $\gamma < 1/(k+1)$, almost surely, there exists a process $\tilde{\mathscr{H}}$ equivalent to $\mathscr{H}$ which has γ-Hölder continuous paths.

A Continuity and measurability statements

Lemma 25. The set $\mathbb{K}^{\mathrm{ghp}}_{\mathrm{f}}$ is a measurable subset of $\mathbb{K}^{\mathrm{ghp}}$ and the bijection ι between $\mathbb{K}^{\mathrm{ghp}}_{\mathrm{f}}$ and $\mathbb{K}^{\mathrm{gp}}$ is bimeasurable.

Proof. For any $\delta > 0$, the quantity $\kappa_\delta(\mathfrak{X}) = \inf\{\mu(\{y \in \mathcal{X} : d(x,y) \leq \delta\}) : x \in \mathcal{X}\}$, only depends on the ghp-equivalence class of $\mathfrak{X}$. Moreover, $\mathbb{K}^{\mathrm{ghp}}_{\mathrm{f}} = \{\mathfrak{X} \in \mathbb{K}^{\mathrm{ghp}} : \kappa_\delta(\mathfrak{X}) > 0 \text{ for all } \delta > 0\}$. A straightforward application of the Portemanteau lemma shows that κ_δ is upper semi-continuous with respect to $\mathrm{d}_{\mathrm{ghp}}$. (The details are given in the proof of Lemma 3.2 in [9].) It follows easily that $\mathbb{K}^{\mathrm{ghp}}_{\mathrm{f}}$ is a measurable set. The map ι is continuous since convergence with respect to the Gromov–Hausdorff–Prokhorov topology implies convergence with respect to the Gromov–Prokhorov topology. Its inverse ι^{-1} is not continuous, and the space $\mathbb{K}^{\mathrm{ghp}}_{\mathrm{f}}$ (considered as subspace of $\mathbb{K}^{\mathrm{ghp}}$) is not complete. Nevertheless, it is important to note that, by [9, Corollary 5.6], $\mathbb{K}^{\mathrm{ghp}}_{\mathrm{f}}$ endowed with the relative topology generated by $\mathrm{d}_{\mathrm{ghp}}$ is Polish. It follows from the Lusin–Souslin theorem, see, e.g. [34, Theorem 15.1], that ι^{-1} is measurable. (We thank Stephan Gufler for pointing out the short argument showing the measurability of ι^{-1}.) □

Proof of Lemma 14. Let $(\mathcal{T}, d, \mu, \rho)$ be a compact rooted measured real tree. We borrow notation from [24]. Recall that $\mathscr{B}$ denotes the set of branch points of $\mathcal{T}$. Let $\sigma \in \mathscr{B}^* :=$

$\mathscr{B} \cup \{\rho\}$ and note that the degree of a point $\sigma \in \mathscr{B}^*$ is at most countably infinite. For $\sigma \in \mathscr{B}^*$, let

$$\mathcal{I}_\sigma = \begin{cases} \emptyset & \text{if } \deg(\sigma) = 1, \\ \{1, 2, \dots, \deg(\sigma) - 1\} & \text{if } 1 < \deg(\sigma) < \infty, \\ \mathbb{N} \setminus \{0\} & \text{if } \deg(\sigma) = \infty. \end{cases}$$

Further, let $\mathcal{C}_\sigma$ be the set of connected components in $\mathcal{T} \setminus \{\sigma\}$ which do not contain ρ. The cardinality of $\mathcal{C}_\sigma$ is given by the cardinality of $\mathcal{I}_\sigma$. Fix a bijection $p_\sigma : \mathcal{I}_\sigma \to \mathcal{C}_\sigma$ (e.g. by considering the order in which the elements of C_σ are visited by a fixed dense sequence in $\mathcal{T}$). Let $\mathcal{D} = \{(\sigma, k) : \sigma \in \mathscr{B}^*, k \in \mathcal{I}_\sigma\}$ and $\mathcal{U}$ be the set of elements $u \in [0,1]^{\mathcal{D}}$ such that $u_i \neq u_j$ for all $i \neq j$, $i, j \in \mathcal{D}$. For $\sigma, \sigma' \in \mathcal{T}$, set $\sigma \leq \sigma'$ if $\sigma \in [\![\rho, \sigma']\!]$ and $\sigma' \leq \sigma$ if $\sigma' \in [\![\rho, \sigma]\!]$. Otherwise, denoting by $\gamma = \sigma \wedge \sigma' \in \mathscr{B}^*$ the most recent common ancestor of σ and σ' (that is, the unique node satisfying $[\![\rho, \sigma \wedge \sigma']\!] = [\![\rho, \sigma]\!] \cap [\![\rho, \sigma']\!]$), and assuming that $\sigma \neq \gamma \neq \sigma'$, there exist $1 \leq i \neq i'$ such that $\sigma \in p_\gamma(i)$ and $\sigma' \in p_\gamma(i')$. Set $\sigma \leq \sigma'$ if $u_{(\gamma, p_\gamma(i))} < u_{(\gamma, p_\gamma(i'))}$ and $\sigma' \leq \sigma$ otherwise. By Proposition 2.4 in [24], $\leq$ is a total order on $\mathcal{T}$ satisfying the following two properties:

(Or1) For $\sigma, \sigma' \in \mathcal{T}$, $\sigma \in [\![\rho, \sigma']\!]$ implies $\sigma \leq \sigma'$.

(Or2) For $\sigma_1 \leq \sigma_2 \leq \sigma_3$ and γ defined by $[\![\rho, \gamma]\!] = [\![\rho, \sigma_1]\!] \cap ([\![\rho, \sigma_2]\!] \cup [\![\rho, \sigma_3]\!])$, we have $\gamma \in [\![\rho, \sigma_2]\!]$.

Note that, by Lemma 2.5 in [24], for $\gamma \in \mathscr{B}^*$ and two connected components $C, C' \in \mathcal{C}_\gamma$, we have either $\sigma \leq \sigma'$ for all $\sigma \in C, \sigma' \in C'$ or $\sigma' \leq \sigma$ for all $\sigma \in C, \sigma' \in C'$.

Next, let $\leq$ be a total order on $\mathcal{T}$ satisfying **(Or1)** and **(Or2)**. We construct a function $h \in \mathbb{D}_{\text{ex}}$ such that $\mathfrak{T}_h = \mathfrak{T}$ (with respect to isometry classes). First, set $m(\sigma) = \mu(\{\sigma' \in \mathcal{T} : \sigma' \leq \sigma\})$. By construction, $m : \mathcal{T} \to [0,1]$ is monotonically increasing with respect to the order $\leq$. As μ has full support and upon setting $\sigma < \sigma'$ if $\sigma \leq \sigma'$ and $\sigma \neq \sigma'$, m is strictly increasing; in particular, $m(y) > 0$ for all $y \neq \rho$. For $\sigma \neq \rho$, we have $m(\sigma-) := m(\sigma) - \mu(\{\sigma\}) = \lim_{\sigma' \to \sigma, \sigma' < \sigma} m(\sigma')$. For $\sigma \notin \mathscr{B}^*$ we also have $m(\sigma) = \lim_{\sigma' \to \sigma, \sigma' > \sigma} m(\sigma')$. For $\mu(\{\rho\}) \leq x < 1$, let $m^{-1}(x) = \sup\{\sigma \in \mathcal{T} : m(\sigma) < x\}$ be the generalized inverse of m. For $x < \mu(\{\rho\})$, set $m^{-1}(x) = \rho$. Finally, let $m^{-1}(1) = \lim_{x \to 1} m^{-1}(x)$. Then, $m^{-1}(x) = \lim_{y \to x, y < x} m^{-1}(y)$ and $m^{-1}(x+) = \lim_{y \to x, y > x} m^{-1}(y)$. Set $h(x) = d(\rho, m^{-1}(x)), x \in [0,1]$. It immediately follows that $h \in \mathbb{D}_{\text{ex}}$. Further, $\mathfrak{T}$ and $\mathfrak{T}_h$ are ghp-isometric via $\sigma \mapsto \pi_h(m(\sigma))$, where π_h is the canonical surjection from $[0,1]$ onto $\mathcal{T}_h$. From the point of view of measurability, the important observation is that the map which assigns to $u \in \mathcal{U}$ the function h constructed from the order induced by u is continuous. This follows from the fact that, for any $\varepsilon > 0$, there exist only finitely many points $\sigma \in \mathscr{B}^*$ such that two distinct elements of $\mathcal{C}_\sigma$ have mass or diameter exceeding ε. We omit a formal proof. Let $\tilde{\eta}$ be the distribution of the random variable h when choosing u uniformly at random, that is, following the distribution $\text{Leb}^{\otimes \mathcal{D}}$. Note that, formally, $\tilde{\eta}$ may depend on $\mathcal{T}$ and the choice of the bijections $\pi_\sigma, \sigma \in \mathscr{B}^*$. Thanks to the symmetry of the distribution $\text{Leb}^{\otimes \mathcal{D}}$ under permuting entries, however, it follows that $\tilde{\eta}$ only depends on the ghp-isometry class of $\mathfrak{T}$. Therefore, we shall denote this distribution by $\tilde{\eta}(\mathfrak{T})$. The corresponding map $\tilde{\eta} : \mathbb{T}_{\text{f}}^{\text{ghp}} \to \mathcal{M}_1(\mathbb{D}_{\text{ex}})$ can be shown to be continuous, hence measurable. Now, for a random variable $\mathfrak{T}$ on $\mathbb{T}_{\text{f}}^{\text{ghp}}$ with distribution ν, the annealed measure $\eta(\cdot) = \mathbf{E}[\tilde{\eta}(\mathfrak{T})(\cdot)]$ is the desired probability distribution. □

The following three lemmas concern continuity of functions arising in the definition of Ψ in Section 2.3 and in the proof of Proposition 3. The proofs rely on the concept of correspondences and have many ideas in common. In our presentation, we focus on a detailed proof of Lemma 27 and sketch the arguments needed to prove continuity of ψ_1 in Lemma 28. The remaining statements have simpler proofs, and we omit them.

Lemma 26. Let $\mathbf{r}, \mathbf{s} \in \Sigma_K$ and $\mathfrak{X}_1, \ldots, \mathfrak{X}_K$ be compact rooted metric measured spaces. Let $\chi : \mathcal{X}_1 \times \ldots \times \mathcal{X}_K \to \mathbb{K}^{\mathrm{ghp}}$ be the map which assigns to each $(x_1, \ldots, x_k) \in \mathcal{X}_1 \times \ldots \times \mathcal{X}_K$ the space $\mathfrak{X}$ when the construction described on page 6 is carried out with these values $\mathbf{r}, \mathbf{s}$ and the glue points $\eta_i = x_i$ for $1 \leq i \leq K$. Then χ is continuous.

Lemma 27. The map $\psi : (\mathbb{K}^{\mathrm{ghp}})^K \times \Sigma_K^2 \to \mathcal{M}_1(\mathbb{K}^{\mathrm{ghp}})$ defined in Section 2.3 is continuous.

Lemma 28. The functions $\chi_1 : \mathcal{X}^{\mathbb{N}} \times c_0^+ \to \mathbb{K}^{\mathrm{ghp}}$ and $\psi_1 : \mathbb{K}^{\mathrm{ghp}} \to \mathcal{M}_1(\mathbb{K}^{\mathrm{ghp}})$ defined in the proof of Proposition 3 are continuous.

Proof of Lemma 27. Let $\varepsilon > 0$ and $\mathfrak{X}_1, \ldots, \mathfrak{X}_K, \mathfrak{X}'_1, \ldots, \mathfrak{X}'_K$ be compact rooted measured metric spaces. Let $(Z_1, d^{Z_1}), \ldots, (Z_k, d^{Z_k})$ be compact metric spaces such that $\mathfrak{X}_i, \mathfrak{X}'_i \subseteq Z_i$ and $d^{Z_i} = d_i$ on $\mathcal{X}_i$ and $d^{Z_i} = d'_i$ on $\mathcal{X}'_i$ for all $i = 1, \ldots, K$. (Such spaces exist. One can, for instance, choose $Z_i = \mathcal{X}_i \sqcup \mathcal{X}'_i$.) Further, assume that,

i) $\mathrm{d_h^{Z_i}}(\mathcal{X}_i, \mathcal{X}'_i) < \varepsilon$, $d^{Z_i}(\rho_i, \rho'_i) < \varepsilon$, and $\mathrm{d_p^{Z_i}}(\mu_i, \mu'_i) < \varepsilon$ for all $i = 1, \ldots, K$,

ii) $\mathbf{r}, \mathbf{r}', \mathbf{s}, \mathbf{s}' \in \Sigma_K$ with $\max\{\|\mathbf{r} - \mathbf{r}'\|, \|\mathbf{s} - \mathbf{s}'\|\} < \varepsilon$.

(Σ_K is endowed with the Euclidean distance.) Finally, assume that the constructions of the spaces $\mathfrak{X} = (\mathcal{X}, d, \mu, \rho)$, and $\mathfrak{X}' = (\mathcal{X}', d', \mu', \rho')$ described on page 6 are carried out with fixed glue points η_i, η'_i satisfying $d^{Z_i}(\eta_i, \eta'_i) < \varepsilon$ for all $i = 1, \ldots, K$. Below, we show that

$$\mathrm{d_{ghp}}(\mathfrak{X}, \mathfrak{X}') \leq (K+1)\left(\varepsilon + \frac{\varepsilon^\alpha}{2} \max_{1 \leq i \leq K} \|\mathfrak{X}_i\|\right). \tag{A.1}$$

Taking this deterministic statement for granted, we can argue as follows to conclude the proof: First, we keep the compact rooted measured metric spaces fixed and choose the glue points randomly. Then, by the previous lemma, (the equivalence classes of) $\mathfrak{X}$ and $\mathfrak{X}'$ are $\mathbb{K}^{\mathrm{ghp}}$-valued random variables. Further, by a coupling theorem due to Strassen [48, Page 438], we may sample the pairs of gluepoints (η_i, η'_i) on $\mathcal{X}_i \times \mathcal{X}'_i$ in such a way that $\mathfrak{L}(\eta_i) = \mu_i$, $\mathfrak{L}(\eta'_i) = \mu'_i$ and $\mathbf{P}(d^{Z_i}(\eta_i, \eta'_i) \geq \varepsilon) \leq \varepsilon$ for all $i = 1, \ldots, K$. Using this coupling and writing γ for the right-hand side of (A.1) yields $\mathbf{P}(\mathrm{d_{ghp}}(\mathfrak{X}, \mathfrak{X}') \geq \gamma) \leq K\varepsilon$. Hence, again using Strassen's theorem, it follows that $\mathrm{d_p}(\mathfrak{L}(\mathfrak{X}), \mathfrak{L}(\mathfrak{X}')) \leq \gamma$. Since γ can be made arbitrarily small by choice of $\varepsilon > 0$, this implies the claimed continuity. (The remainder of the proof also shows that, for fixed $\mathbf{r}, \mathbf{s}$, the map ψ with domain $(\mathbb{K}^{\mathrm{ghp}})^K$ is uniformly continuous.)

It remains to show (A.1). To this end, we recall the well-known characterization of the Gromov–Hausdorff distance using correspondences: for two sets S, T, a set $\mathfrak{R} \subseteq S \times T$ is called a *correspondence* if for all $s \in S$ there exists $t \in T$ with $(s, t) \in \mathfrak{R}$ and vice versa. For metric spaces $(S, d_S), (T, d_T)$, the *distorsion* of a correspondence $\mathfrak{R}$ is defined by

$$\mathrm{dis}(\mathfrak{R}) = \sup\{|d_S(s, s') - d_T(t, t')| : (s, t), (s', t') \in \mathfrak{R}\}.$$

Define the Gromov–Hausdorff distance between compact rooted metric spaces (S, d_S, ρ_S), and (T, d_T, ρ_T) using the notation from Eq. (2.1) by

$$\mathrm{d_{gh}}((S, d_S, \rho_S), (T, d_T, \rho_T)) = \inf_{Z, \phi, \phi'} \max\left\{\mathrm{d_h^Z}(\phi(S), \phi'(T)), d^Z(\phi(S), \phi'(T))\right\}.$$

It is standard ([see, e.g., 15, Theorem 7.3.5] in the unrooted set-up) that, for compact rooted metric spaces (S, d_S, ρ_S) and (T, d_T, ρ_T), we have

$$\begin{aligned} &\mathrm{d_{gh}}((S, d_S, \rho_S), (T, d_T, \rho_T)) \\ &= \frac{1}{2} \inf\{\mathrm{dis}(\mathfrak{R}) : \mathfrak{R} \text{ is a correspondence with } (\rho_S, \rho_T) \in \mathfrak{R}\}. \end{aligned} \tag{A.2}$$

It easy to construct optimal correspondences explicitly: For $i = 1, \dots, K$, we can set

$$\mathfrak{R}_i = \{(x, y) : x \in \mathcal{X}_i, y \in \mathcal{X}'_i, d^{Z_i}(x, y) < \varepsilon\}.$$

By the triangle inequality, we have $\mathrm{dis}(\mathfrak{R}_i) \le 2\varepsilon$. Recall the projections $\varphi^\circ, (\varphi^\circ)'$ in step *ii)* of the construction and define a correspondence for $\mathcal{X}$ and $\mathcal{X}'$ by

$$\mathfrak{R} = \bigcup_{i=1}^{K} \{(\varphi^\circ(x), \varphi^\circ(x')) : (x, x') \in \mathfrak{R}_i\}.$$

(Note that the roots are already identified in the $\mathfrak{R}_i$.) From (A.2), we see that $\mathrm{d_{gh}}((\mathcal{X}, d, \rho), (\mathcal{X}', d', \rho')) \le \frac{1}{2}\mathrm{dis}(\mathfrak{R})$. In order to extend the result to the Gromov–Hausdorff–Prokhorov distance, we consider a specific embedding of the spaces. The following construction is standard, see, e.g. the proof of Theorem 7.3.5 in [15]. Let $Z = \mathcal{X} \sqcup \mathcal{X}'$ and set $d^Z = d$ on $\mathcal{X}^2$, $d^Z = d'$ on $\mathcal{X}' \times \mathcal{X}'$ while, for $x \in \mathcal{X}, y \in \mathcal{X}'$, define

$$d^Z(x, y) = \inf\{d(x, x_1) + d'(x_2, y) : (x_1, x_2) \in \mathfrak{R}\} + \frac{1}{2}\mathrm{dis}(\mathfrak{R}).$$

Finally, for $x \in \mathcal{X}', y \in \mathcal{X}$, set $d^Z(x, y) = d^Z(y, x)$. A straightforward computation shows that d^Z is a metric on Z. Further, by construction, $\mathrm{d_h^Z}((\mathcal{X}, d, \rho), (\mathcal{X}', d', \rho')) = d^Z(\rho, \rho') = \frac{1}{2}\mathrm{dis}(\mathfrak{R})$. To study the Prokhorov distance between μ and μ', let $(\sigma_i, \sigma'_i), i = 1, \dots, K$ be pairs of random variables on $\mathcal{X}_i \times \mathcal{X}'_i$ with $\mathfrak{L}(\sigma_i) = \mu_i, \mathfrak{L}(\sigma'_i) = \mu'_i$ and $\mathbf{P}(|\sigma_i - \sigma'_i| \ge \varepsilon) \le \varepsilon$. By our assumptions on s, s', there exists a pair of random variables (J, J') with $\mathbf{P}(J = i) = s_i$ and $\mathbf{P}(J' = i) = s'_i$ for all $i = 1, \dots, K$ and $\mathbf{P}(J \ne J') \le K\varepsilon$. Note that $\mathfrak{L}(\varphi^\circ(\sigma_J)) = \mu$ and $\mathfrak{L}((\varphi^\circ)'(\sigma'_{J'})) = \mu'$. Hence, as $d^{Z_i}(x, y) < \varepsilon$ for $x \in \mathcal{X}_i, y \in \mathcal{X}'_i$ implies $d^Z(\varphi^\circ(x), (\varphi^\circ)'(y)) \le \frac{1}{2}\mathrm{dis}(\mathfrak{R})$, it follows that $\mathbf{P}(|\varphi^\circ(\sigma_J) - (\varphi^\circ)'(\sigma'_{J'})| > \frac{1}{2}\mathrm{dis}(\mathfrak{R})) \le (K + 1)\varepsilon$. Thus, $\mathrm{d_p}(\mu, \mu') \le \max\{\frac{1}{2}\mathrm{dis}(\mathfrak{R}), (K + 1)\varepsilon\}$ and the same bound applies to $\mathrm{d_{ghp}}(\mathfrak{X}, \mathfrak{X}')$.

It remains to find an upper bound on $\mathrm{dis}(\mathfrak{R})$. We have

$$\mathrm{dis}(\mathfrak{R}) = \sup_{1 \le i, j \le K} \sup\{|d(\varphi^\circ(x), \varphi^\circ(y)) - d'((\varphi^\circ)'(x'), (\varphi^\circ)'(y'))| : (x, x') \in \mathfrak{R}_i, (y, y') \in \mathfrak{R}_j\}.$$

For $(x, x'), (y, y') \in \mathfrak{R}_i$, it follows that under our imposed assumptions, we have

$$|d(\varphi^\circ(x), \varphi^\circ(y)) - d'(\varphi^\circ(x'), \varphi^\circ(y'))| \le 2\varepsilon + \varepsilon^\alpha \|\mathfrak{X}_i\|.$$

An application of the triangle inequality shows that

$$\mathrm{dis}(\mathfrak{R}) \le K\Big(2\varepsilon + \varepsilon^\alpha \max_{i=1,\dots,K} \|\mathfrak{X}_i\|\Big),$$

which completes the proof. □

Proof of Lemma 28. We only sketch the arguments necessary to prove of continuity of ψ_1 which go beyond the details presented in the previous proof. Fix $\mathfrak{X} \in \mathbb{K}^{\mathrm{ghp}}$. Let $\varepsilon > 0$ and $\mathfrak{X}' \in \mathbb{K}^{\mathrm{ghp}}$ with $\mathrm{d_{ghp}}(\mathfrak{X}, \mathfrak{X}') \le \varepsilon$. Let $(\mathcal{X}, d, \mu, \rho)$ and $(\mathcal{X}', d', \mu', \rho')$ be representatives of these classes embedded in the same compact metric space (Z, d^Z). For $n \in \mathbb{Z}$, let $I_n = [2^{-(n+1)}, 2^{-n})$. We can construct coupled Poisson processes $\mathcal{P}$ on $\mathcal{X} \times [0, \infty)$ and $\mathcal{P}'$ on $\mathcal{X}' \times [0, \infty)$ by superposing coupled independent Poisson processes $\mathcal{P}_n$ on $\mathcal{X} \times I_n$ and $\mathcal{P}'_n$ on $\mathcal{X}' \times I_n$ as follows: first, independently for different values of n, let $(U_1^{(n)}, S_1^{(n)}), (U_2^{(n)}, S_2^{(n)}), \dots,$ be a sequence of independent and identically distributed $(\mathcal{X} \times I_n)$-valued random variables, where $U_i^{(n)}, S_i^{(n)}$ are independent, $U_i^{(n)}$ is distributed

according to μ and $S_i^{(n)}$ follows the distribution of a non-negative random variable with density $s^{-1-1/\alpha}$ conditioned on taking a value in I_n. As $\mathrm{d}_{\mathrm{p}}^{\mathrm{Z}}(\mu,\mu') \le \varepsilon$, for any $n \in \mathbb{Z}$, we can construct a sequence of independent and identically distributed $\mathcal{X}'$-valued random variables $V_1^{(n)}, V_2^{(n)}, \dots$ such that $\mathbf{P}(d^Z(U_i^{(n)}, V_i^{(n)}) \ge \varepsilon) \le \varepsilon$ for all $i \ge 1$. Let $N_n, n \in \mathbb{Z}$, be a family of independent random variables which is independent of all previously defined quantities, where N_n follows the Poisson distribution with parameter $\int_{I_n} s^{-1-1/\alpha} ds = \alpha 2^{n/\alpha}(2^{1/\alpha}-1)$. The sets of points $\{(U_i^{(n)}, S_i^{(n)}) : 1 \le i \in N_n\}$ and $\{(V_i^{(n)}, S_i^{(n)}) : 1 \le i \le N_n\}$ constitute Poisson processes $\mathcal{P}_n$ on $\mathcal{X} \times I_n$, and $\mathcal{P}'_n$ on $\mathcal{X}' \times I_n$, respectively. Upon superposing the processes for different values of n, we obtain the sought coupled processes $\mathcal{P}$ and $\mathcal{P}'$. It should be clear and can be shown using correspondences that, for any $\delta > 0$, $\sup_{\mathfrak{X}'} \mathbf{P}\left(\mathrm{d}_{\mathrm{ghp}}(f(\mathcal{P}), f(\mathcal{P}')) \ge \delta\right) \to 0$ as $\varepsilon \to 0$, where the supremum is taken over all $\mathfrak{X}' \in \mathbb{K}^{\mathrm{ghp}}$ satisfying $\mathrm{d}_{\mathrm{ghp}}(\mathfrak{X}, \mathfrak{X}') \le \varepsilon$. (Recall that $\mathfrak{X}$ is kept fixed.) Thus, $\sup_{\mathfrak{X}'} \mathrm{d}_{\mathrm{p}}(\mathfrak{L}(f(\mathcal{P})), \mathfrak{L}(f(\mathcal{P}')) \to 0$ as $\varepsilon \to 0$ showing the claimed continuity. □

Lemma 29. For any $s > 0$, the map H^s is measurable for d_{gh}, and thus $\dim_{\mathrm{h}}$ as well.

Proof. For fixed $\delta > 0$, let

$$H_\delta^s(A) := \inf\left\{\sum_{i\ge1} |U_i|^s : A \subseteq \bigcup_{i\ge1} U_i \text{ and } |U_i| \le \delta \text{ for all } i \ge 1\right\}.$$

Then, $H^s = \lim_{\delta\downarrow 0} H_\delta^s$ and it is enough to prove the measurability of $H_\delta^s : \mathbb{K}^{\mathrm{gh}} \to [0,\infty]$ for fixed $\delta > 0$. To this end, we show that the function is upper-semicontinuous. Let $\delta > 0, s \ge 0$ and $\mathfrak{X}, \mathfrak{X}_1, \mathfrak{X}_2, \dots$ be compact metric spaces with $d_{\mathrm{gh}}(\mathfrak{X}_n, \mathfrak{X}) \to 0$. For $\varepsilon > 0$, by compactness, there exists a natural number N and sets $U_1, \dots, U_n$ with $|U_i| < \delta$ for all $i = 1, \dots, N$ such that $\mathcal{X} \subseteq \bigcup_{i=1}^N U_i$ and $H_\delta^s(\mathfrak{X}) \ge \sum_{i=1}^N |U_i|^s - \varepsilon$. Now, let $0 < \varepsilon' < \varepsilon$ be sufficiently small such that $|U_i^{\varepsilon'}| < \delta$ for all $i = 1, \dots N$. Choose n large enough such that $d_{\mathrm{gh}}(\mathfrak{X}_n, \mathfrak{X}) < \varepsilon'/2$. We may assume that $\mathcal{X}_n$ and $\mathcal{X}$ are embedded in a compact metric space (Z, d^Z) such that $\mathrm{d}_{\mathrm{h}}^{\mathrm{Z}}(\mathfrak{X}_n, \mathfrak{X}) < \varepsilon'/2$. Then, $H_\delta^s(\mathfrak{X}_n) \le \sum_{i=1}^N |U_i^{\varepsilon'}|^s \le (1+\varepsilon)H_\delta^s(\mathfrak{X}) + 2\varepsilon$. As this inequality is true for sufficiently large n, we can take the limit superior on the left-hand side. Then, letting $\varepsilon \to 0$ on the right-hand side shows the claim. Finally, it is easy to deduce measurability of $\dim_{\mathrm{h}}$, e.g. from the representation

$$\{\mathfrak{X} \in \mathbb{K}^{\mathrm{gh}} : \dim_{\mathrm{h}}(\mathfrak{X}) > t\} = \bigcup_{q>t} \bigcap_{\varepsilon>0} \bigcap_{\delta>0} \{\mathfrak{X} \in \mathbb{K}^{\mathrm{gh}} : H_\delta^s(\mathfrak{X}) < \varepsilon\}, \quad t \ge 0,$$

where q, ε, δ take rational values. □

The proof of the next lemma runs along the same lines as the proof of Lemma 29. (Technically, $N_\cdot(\delta)$ plays a very similar role as $H_\delta^s(\cdot)$.) We omit the details.

Lemma 30. Let $\delta > 0$. The map from $\mathbb{K}^{\mathrm{gh}}$ to $\mathbb{N}$ that to $\mathfrak{X} \in \mathbb{K}^{\mathrm{gh}}$ assigns $N_{\mathcal{X}}(\delta)$, the smallest number of open balls of radius δ needed to cover $\mathcal{X}$, is upper-semicontinuous In particular, the maps $\underline{\dim}_{\mathrm{m}}, \overline{\dim}_{\mathrm{m}} : \mathbb{K}^{\mathrm{gh}} \to [0,\infty]$ are measurable.

B Hölder continuity of $\mathscr{H}$

Proof of Proposition 19. The positive result for $\alpha < \varrho$ follows from bounds on the moments of $\mathscr{H}$ provided in Lemma 31 and Kolmogorov's criterion. Next, recall from [14] that $h(t) := \mathbf{E}\left[\mathscr{H}(t)\right] = \kappa'\sqrt{t(1-t)}$ for some $\kappa' > 0$. Let $\varrho < \alpha < \gamma$. With Q_n^ϑ defined in Eq. (4.9), 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 $n \ge 1$, 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,

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

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., [21]) 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. □

The final lemma generalizes Proposition 4.1 in [19] to non-integer values of p.

Lemma 31. 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 [19] 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

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

Thus, for $0 \leq p \leq 1$, the assertion follows immediately from (B.1). 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 [19] 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 Eq. (2.13), we have, in a stochastic sense,

$$\begin{aligned}
&\mathscr{H}(t)^p \\
&\leq \mathbf{1}_{[0,U_1)}(t) W^{p/3} (\mathscr{H}^{(1)})^p \left(\frac{t}{\Delta_1}\right) + \mathbf{1}_{[U_2,1]}(t) W^{p/3} (\mathscr{H}^{(1)})^p \left(\frac{t-\Delta_2}{\Delta_1}\right) \\
&+ \mathbf{1}_{[U_1,U_2)}(t) \left((1-W)^{p/3} (\mathscr{H}^{(2)})^p \left(\frac{t-U_1}{\Delta_2}\right) + W^{p/3} (\mathscr{H}^{(1)})^p (\xi)\right) \\
&+ \mathbf{1}_{[U_1,U_2)}(t) \left(C_1 (1-W)^{(p-1)/3} W^{1/3} (\mathscr{H}^{(2)})^{p-1} \left(\frac{t-U_1}{\Delta_2}\right) \mathscr{H}^{(1)}(\xi)\right) && \text{(B.3)} \\
&+ \mathbf{1}_{[U_1,U_2)}(t) \left(C_1 (1-W)^{1/3} W^{(p-1)/3} (\mathscr{H}^{(2)}) \left(\frac{t-U_1}{\Delta_2}\right) (\mathscr{H}^{(1)})^{p-1}(\xi)\right), && \text{(B.4)}
\end{aligned}$$

with conditions as in Eq. (2.13) on the right hand side. Subsequently, we consider $0 \leq t \leq 1/2$ which suffices by symmetry. With $q = 3/(3+p)$, taking the expectation on both sides of the last display leads to

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

where $s_p(t)$ is the sum of the expectation of (B.3) and (B.4). As in the proof of Proposition 4.1 in [19], 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 $(a+b)^p \geq a^p + b^p$ for $a, b \geq 0$, 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 $t \leq t_0$, we have

$$m_p(t) \leq 2(1+\delta)qt^2 \int_t^1 m_p(x)x^{-3}dx + 2s_p(t). \tag{B.5}$$

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

$$m_p(t) \leq 2(1+\delta)qt^2 \int_t^{t_0} m_p(x)x^{-3}dx + 2s_p(t) + C_2 t^2.$$

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

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

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 $1 < p \leq 2$. Then, a direct computation shows that, for some $C_4 = C_4(p)$, we have $s_p(t) \leq C_4 t^{(p+2)/2}$. Thus, by (B.2), $s_p(t)$ is asymptotically negligible compared to $m_p(t)$. Using (B.5), for any $\delta' > 0$, upon decreasing t_0 if necessary, we have

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

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 $s_p(t) \leq C_5 t^{3/2+2(p-1)/(p+2)-\varepsilon}$. Thus, $s_p(t)t^{-2}$ is bounded on $[0,1]$ and the result follows as indicated from inequality (B.5). □

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. MR-3035742

[2] 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, 2015. MR-3399812

[3] D. Aldous. The continuum random tree. I. *Ann. Probab.*, 19(1):1–28, 1991. MR-1085326

[4] D. Aldous. The continuum random tree II: an overview. In M.T. Barlow and N.H. Bingham, editors, *Stochastic Analysis*, pages 23–70. CUP, 1991. MR-1166406

[5] D. Aldous. The continuum random tree III. *Ann. Probab.*, 21:248–289, 1993. MR-1207226

[6] D. Aldous. Recursive self-similarity for random trees, random triangulations and Brownian excursion. *Ann. Probab.*, 22(2):527–545, 1994. MR-1288122

[7] D. Aldous. Triangulating the circle, at random. *Amer. Math. Monthly*, 101(3):223–233, 1994. MR-1264002

[8] D. J. Aldous and A. Bandyopadhyay. A survey of max-type recursive distributional equations. *Ann. Appl. Probab.*, 15(2):1047–1110, 2005. MR-2134098

[9] 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. MR-3522292

[10] J. Bertoin and A. Gnedin. Asymptotic laws for nonconservative self-similar fragmentations. *Electron. J. Probab.*, 9:575–593, 2004. MR-2080610

[11] P. J. Bickel and D. A. Freedman. Some asymptotic theory for the bootstrap. *Ann. Statist.*, 9(6):1196–1217, 1981. MR-0630103

[12] P. Billingsley. *Convergence of Probability Measures*. Wiley, New York, 1968. MR-0233396

[13] N. Broutin and J.-F. Marckert. Asymptotics of trees with a prescribed degree sequence and applications. *Random Structures & Algorithms*, 44:290–316, 2014. MR-3188597

[14] N. Broutin and H. Sulzbach. The dual tree of a recursive triangulation of the disk. *Ann. Probab.*, 43(2):738–781, 2015. MR-3306004

[15] D. Burago, Y. Burago, and S. Ivanov. *A course in metric geometry*. AMS, 2001. MR-1835418

[16] V. V. Chistyakov and O. E. Galkin. On maps of bounded p-variation with $p > 1$. *Positivity*, 2(1):19–45, 1998. MR-1655755

[17] R. Cori and B. Vauquelin. Planar maps are well labeled trees. *Canadian Journal of Mathematics*, 33:1023–1042, 1981. MR-0638363

[18] D. Croydon and B. Hambly. Self-similarity and spectral asymptotics for the continuum random tree. *Stochastic Process. Appl.*, 118(5):730–754, 2008. MR-2411518

[19] N. Curien and J.-F. Le Gall. Random recursive triangulations of the disk via fragmentation theory. *Ann. Probab.*, 39(6):2224–2270, 2011. MR-2932668

[20] N. Curien and Y. Peres. Random laminations and multitype branching processes. *Electron. Commun. Probab.*, 16:435–446, 2011. MR-2831082

[21] A. Dembo and O. Zeitouni. *Large Deviation Techniques and Applications*. Springer, second edition, 1998. MR-1619036

[22] A. Depperschmidt, A. Greven, and P. Pfaffelhuber. Marked metric measure spaces. *Electr. J. Probab.*, 16:174–188, 2011. MR-2783338

[23] B. Duplantier, J. Miller, and S. Sheffield. Liouville quantum gravity as a mating of trees. Preprint, arXiv:1409.7055, 2014. MR-3729063

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

[25] T. Duquesne and J.-F. Le Gall. Probabilistic and fractal aspects of Lévy trees. *Probab. Theory Related Fields*, 131(4):553–603, 2005. MR-2147221

[26] T. Duquesne and J.-F. Le Gall. The Hausdorff measure of stable trees. *ALEA Lat. Am. J. Probab. Math. Stat.*, 1:393–415, 2006. MR-2291942

[27] R. Durrett and T. M. Liggett. Fixed points of the smoothing transformation. *Z. Wahrsch. Verw. Gebiete*, 64(3):275–3719, 1983. MR-0716487

[28] S. N. Evans. *Probability and real trees*, volume 1920 of *Lecture Notes in Mathematics*. Springer, Berlin, 2008. MR-2351587

[29] K. J. Falconer. *Fractal geometry: Mathematical Foundations and Applications*. John Wiley & Sons Ltd., Chichester, 1990. MR-1102677

[30] 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. MR-2520129

[31] M. Gromov. *Metric Structures for Riemannian and Non-Riemannian Spaces*. Birkhauser, 1999. MR-1699320

[32] 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. MR-2041829

[33] S. Janson and R. Neininger. The size of random fragmentation trees. *Probability Theory and Related Fields*, 142:399–442, 2008. MR-2438697

[34] A. S. Kechris. *Classical descriptive set theory*, volume 156 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995. MR-1321597

[35] J.-F. Le Gall. Random trees and applications. *Probab. Surveys*, 2:245–311, 2005. MR-2203728

[36] J.-F. Le Gall. Uniqueness and universality of the Brownian map. *Ann. Probab.*, 41:2880–2960, 2013. MR-3112934

[37] W. Löhr. Equivalence of Gromov-Prokhorov- and Gromov's $\square_\lambda$-metric on the space of metric measure spaces. *Electr. J. Probab.*, 18:1–10, 2013. MR-3037215

[38] J.-F. Marckert and G. Miermont. The CRT is the scaling limit of unordered binary trees. *Random Structures & Algorithms*, 38:467–501, 2011. MR-2829313

[39] S.P. Meyn and R.L. Tweedie. *Markov chains and stochastic stability*. Springer, New York, 1993. MR-1287609

[40] G. Miermont. The Brownian map is the scaling limit of uniform random plane quadrangulations. *Acta Math.*, 210:319–401, 2013. MR-3070569

[41] R. Neininger and H. Sulzbach. On a functional contraction method. *Ann. Probab.*, 43(4):1777–1822, 2015. MR-3353815

[42] K. Panagiotou and B. Stufler. Scaling limits of random Pólya trees. *Probab. Theory Related Fields*, 170(3–4):801–820, 2018. MR-3773800

[43] K. Panagiotou, B. Stufler, and K. Weller. Scaling limits of random graphs from subcritical classes. *Ann. Probab.*, 44(5):3291–3334, 2016. MR-3551197

[44] J. Picard. A tree approach to p-variation and to integration. *Ann. Probab.*, 36(6):2235–2279, 2008. MR-2478682

[45] F. Rembart and M. Winkel. Recursive construction of continuum random trees. *Ann. Probab.*, 46(5):2715–2748, 2018. MR-3846837

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

[47] R. Stephenson. General fragmentation trees. *Electron. J. Probab.*, 18(101), 2013. MR-3141802

[48] V. Strassen. The existence of probability measures with given marginals. *Ann. Math. Statist.*, 36:423–439, 1965. MR-0177430

[49] B. Stufler. The continuum random tree is the scaling limit of unlabelled unrooted trees. arXiv:1412.6333, 2014. MR-3983790

[50] A. M. Vershik. The universal Urysohn space, Gromov metric triples and random metrics on the natural numbers. *Russian Mathematical Surveys*, 53(5):921, 1998. MR-1691182

[51] W. Vervaat. On a stochastic difference equation and a representation of non-negative infinitely divisible random variables. *Adv. Appl. Probab.*, 11:750–783, 1979. MR-0544194