CONVERGENCE AND LIMITS OF FINITE TREES

Gábor Elek[a,b] and Gábor Tardos[b,c]

[a] Department of Mathematics and Statistics, Fylde College,
Lancaster University, Lancaster, United Kingdom
[b] Alfréd Rényi Institute of Mathematics, Budapest, Hungary
[c] Moscow Institute of Physics and Technology, Dolgoprudny, Russia

g.elek@lancaster.ac.uk, tardos@renyi.hu

ABSTRACT. Motivated by the work of Lovász and Szegedy on the convergence and limits of dense graph sequences [10], we investigate the convergence and limits of finite trees with respect to sampling in normalized distance. We introduce dendrons (a notion based on separable real trees) and show that the sampling limits of finite trees are exactly the dendrons. We also prove that the limit dendron is unique.

Keywords. convergence of finite trees, real trees, ultraproducts

1. INTRODUCTION

The main motivation of our paper is the (sampling) limit theory of dense graphs introduced by Lovász and Szegedy [10]. Let us recall very briefly the most important definitions. Let G be a finite simple graph and $r \geq 1$ be an integer. Let us pick r distinct vertices $v_1, v_2, \ldots v_r$ of G uniformly at random and consider the graph H induced by the chosen vertices. (Here we assume that the graph has at least r vertices. Alternatively, one can use sampling with repetition and define H as the r vertex graph obtained by possible duplication of the vertices in the induced subgraph.) Then, H will be isomorphic to one of the $2^{\binom{r}{2}}$ graphs on r labeled vertices. Thus, the random choice of the vertices v_i defines a probability distribution p_r^G on the finite set A_r of these labeled graphs. We say that the sequence of finite graphs $(G_n)_{n \in \mathbb{N}}$ is **convergent** if $\lim_{n \to \infty} p_r^{G_n}(K)$ exists for all $r \in \mathbb{N}$ and $K \in A_r$. (Throughout this paper $\mathbb{N}$ stands for the set of positive integers.) Lovász and Szegedy constructed a universal limit object for such convergent graph sequences, the **graphons**. A graphon is a measurable symmetric function $W : [0,1]^2 \to [0,1]$. For $K \in A_r$ (with the vertices of K denoted by the integers 1 through r), $p_r^W(K)$ is defined

as

$$p_r^W(K) = \int_0^1 \int_0^1 \cdots \int_0^1 \prod_{i<j,(i,j)\in E(K)} W(x_i,x_j) \prod_{i<j,(i,j)\notin E(K)} (1-W(x_i,x_j))dx_1 dx_2 \ldots dx_r\,.$$

The graphon W is the limit of the sequence $(G_n)_{n\in\mathbb{N}}$ if for all $r \geq 1$ and $K \in A_r$, $\lim_{n\to\infty} p_r^{G_n}(K) = p_r^W(K)$. It has been proved in [10] that for any convergent sequence $(G_n)_{n\in\mathbb{N}}$ there exists a limit graphon and all graphons are limits of convergent sequences of finite graphs. The uniqueness problem was considered in [2]. The goal of our paper is to introduce and study the sampling limit theory of finite trees. This aim seems to be rather contradictory since the sampling limit theory of Lovász and Szegedy gives non-trivial limit objects only if the graphs are dense, and the trees are very sparse graphs. We solve this problem by regarding trees as dense objects using their natural metric structure.

We identify each finite tree with its vertex set. We call a finite tree **non-trivial** if it has at least two vertices. Let T be a nontrivial finite tree. In order to have more sequences of finite trees that are convergent we need to normalize the shortest path metric $\overline{d}_T$ on T. Here we choose to normalize it by its diameter $\mathrm{diam}_{\overline{d}_T}(T)$. We consider the following metric space structure on T.

$$d_T(x,y) = \frac{\overline{d}_T(x,y)}{\mathrm{diam}_{\overline{d}_T}(T)}.$$

Throughout the paper we use this normalization making every non-trivial finite tree into a metric space of diameter exactly 1. This is for convenience only; see Remark 2 on the possibility for other normalizations. We consider the uniform probability measure μ_T on T as well. So, we obtain a **metric measure space** structure on our tree. The notion of sampling convergence for metric measure spaces has been introduced by Gromov in the famous Section $3\frac{1}{2}$ of his treatise "Metric structures for Riemannian and Non-Riemannian Spaces" [6]. Let us recall the formal definitions.

Definition 1.1. For $r \geq 1$, let M_r be the space of real matrices $(d_{ij})_{1\leq i,j\leq r}$. We identify M_r with the product space $\prod_{1\leq i,j\leq r}\mathbb{R}$.

For a set X and a function $d : X^2 \to \mathbb{R}$ we define the map $\rho_r = \rho_r^{X,d} : X^r \to M_r$ with $\rho_r(x_1,x_2,\ldots,x_r) = (d_{ij})_{1\leq i,j\leq r}$, where $d_{ij} = d(x_i,x_j)$ for $i \neq j$ and $d_{ii} = 0$. Treating d_{ii} separately is needed because we will later use ρ_r (and the sampling measure τ_r below) not only for distance functions d but also for functions not satisfying $d(x,x) = 0$ for all $x \in X$.

A **metric measure space** is a triple (X,d,μ), where (X,d) is a complete separable metric space and μ is a probability Borel measure on X.

For a metric measure space, or more generally for a triple (X,d,μ), where μ is a probability measure on X and $d : X^2 \to \mathbb{R}$ is μ^2-measurable we define the sampling measure $\tau_r(X,d,\mu) \in \mathrm{Prob}(M_r)$ as the push-forward of the product measure μ^r along ρ_r, that is, for a Borel set H in M_r we set $\tau_r(X,d,\mu)(H) = \mu^r(\rho_r^{-1}(H))$.

For a nontrivial finite tree T and $r \geq 1$ we write $\tau_r(T)$ as a shorthand for $\tau_r(T, d_T, \mu_T)$, where d_T is the normalized distance and μ_T is the uniform measure. Our normalization makes the diameter 1, so the image of ρ_r^{T, d_T} is contained in $M_r^1 = \prod_{1 \leq i,j \leq r}[0, 1]$ and thus $\tau_r(T)$ is concentrated on M_r^1.

Definition 1.2. The sequence of finite trees $(T_n)_{n \in \mathbb{N}}$ is convergent in metric sampling if for any $r \geq 1$ the sequence of measures $(\tau_r(T_n))_{n \in \mathbb{N}}$ converges weakly in the space $\mathrm{Prob}(M_r)$ of the Borel probability measures on M_r.

Since $\tau_r(T)$ is concentrated on the compact space M_r^1 for any nontrivial finite tree T, one can pick a convergent subsequence from any sequence of finite trees. The main goal of our paper is to identify the limit objects of such convergent sequences of finite trees. First, we need to recall the notion of a real tree, our key topological notion (see [1] for a survey).

Definition 1.3. We say that the non-empty complete metric space (T, d) is a **real tree** or $\mathbb{R}$-**tree** if for any pair of distinct points $p, q \in T$ one has an isometric embedding of an interval $\alpha : [a, b] \to X$ such that $\alpha(a) = p$, $\alpha(b) = q$ and $\alpha(c)$ separates p from q for any $a < c < b$, that is, p and q are in distinct connected components of $T \setminus \{\alpha(c)\}$.

We write $[x, y]$ to denote $\alpha([a, b])$, which is uniquely determined in a real tree. We let $[x, x] = \{x\}$ and call $[x, y]$ a **segment** in the real tree (T, d). If $x \neq y$ we call $[x, y]$ a **proper segment**. All the points of $[x, y]$ other than its end points x and y are called the **intermediate points** of the segment $[x, y]$. An **intermediate point** of a real tree is an intermediate point of a proper segment in the tree.

For $p \in T$ we call the connected components of $T \setminus \{p\}$ the p-**branches**. Note that p is an intermediate point if and only if there are at least two p-branches. A **branch** of T is a p-branch for some $p \in T$.

A natural limit object would be a **measured real tree**, that is a separable real tree equipped with a probability measure making it a metric measure space. In fact, there are several known metrics for metric measure spaces, like the Gromov-Prohorov metric [5] and Gromov's $\square_1$ metric [6]. Both of these define Gromov's weak topology (given by the sampling) [9], see also Section 10 of [7]. This means that a sequence of measured real trees tend to another metric measure space (necessarily a measured real tree) in one of these metrics if and only if they tend to it in metric sampling. Nevertheless, these metrics are substantively different from the metric obtained from sampling that we consider in this paper. In particular, Gromov-Prohorov metric and the $\square_1$ metric is complete as opposed to the metric obtained from metric samplings, which is pre-compact.

We will introduce *dendrons* as limit objects for sequences of finite trees that are convergent in metric sampling. If a sequence of finite trees (considered as metric measure spaces after normalization) tend to another metric measure space in Gromov's weak topology, then the finite trees form a convergent sequence also in metric sampling and the limit dendron can be easily obtained from the limit in the Gromov's weak topology, see Example 1 below. But

there are many sequences of finite trees that are convergent in metric sampling, but the corresponding measured real trees are not convergent in the Gromov's weak topology, see examples 2–6 below. Hence, our dendrons can be considered to be the compactification of the space of measured real trees of maximum diameter 1.

It is natural to construct the limit objects as some generalization of measured real trees. One way to do it is to drop the requirement of a measured real tree to be separable (and at the same time allow some Borel sets not to be measurable). See the definition of quasi measured real trees in the next section and Remarks 5 and 6. Here we take another route.

Definition 1.4. A **long dendron** $D = (T, d, \nu)$ is a real tree (T, d) together with a probability Borel measure ν on $A_D = T \times [0, \infty)$ satisfying $\nu(B \times [0, \infty)) > 0$ for all branches B of T. We define $d_D : A_D^2 \to \mathbb{R}$ by $d_D((u, a), (v, b)) = d(u, v) + a + b$. We say that D is a **dendron** if d_D is almost surely bounded by 1, that is, if $\nu^2(\{(x, y) \in A_D^2 \mid d_D(x, y) > 1\}) = 0$.

For a long dendron $D = (T, d, \nu)$ and $r \geq 1$, we write $\tau_r(D)$ as the shorthand for the sampling measure $\tau_r(A_D, d_D, \nu)$.

We will use the notation A_D and d_D for (long) dendrons D in the above sense. We say that two long dendrons $D = (T, d, \nu)$ and $D' = (T', d', \nu')$ are **isomorphic** if there exists an isometry f from (T, d) to (T', d') such that $f' : A_D \to A_{D'}$ defined as $f'(p, a) = (f(p), a)$ is measure preserving.

Remark 1. In the definition of a (long) dendron we did not explicitly require the real tree (T, d) to be separable, but this is implicit in the definition. We will state this explicitly in Corollary 2.1.

Note that d_D defined above for a (long) dendron $D = (T, d, \nu)$ is not a distance on A_D as $d_D(x, x) = 2a > 0$ for all $x = (u, a) \in A_D$ with $a > 0$.

Long dendrons can formally be considered marked metric measure spaces as introduced by Depperschmidt, Greven and Pfaffelhuber, see [3], where the infinite interval $[0, \infty)$ is the space of possible marks. This connection is superficial though, as the topology is rather different. We will see that finite trees correspond to dendrons where the mark is 0 with probability one, but their limit points include all dendrons, among them those where the mark is separated from 0 with probability one. See Examples 2, 3 and 6 below.

Our main results are the following three theorems that are analogues of the main results of Lovász and Szegedy in [10] and Borgs, Chayes and Lovász in [2].

Theorem 1. *For any convergent sequence of finite trees $(T_n)_{n \in \mathbb{N}}$ there exists a dendron D (the sampling limit of $(T_n)_{n \in \mathbb{N}}$) such that the sampling measures $\tau_r(T_n)$ weakly converge to $\tau_r(D)$ for all positive integers r.*

Theorem 2. *Any dendron is the sampling limit of a convergent sequence of finite trees.*

Theorem 3. *The sampling limit is unique up to isomorphism. More generally, if two long dendrons D and D' satisfy $\tau_r(D) = \tau_r(D')$ for all r, then D and D' are isomorphic.*

Remark 2. These theorems establish dendrons as the sampling limits of finite trees with respect to distance normalized by the diameter. Long dendrons can be considered an extension suitable for other normalizations. It is possible to consider finite trees with explicitly given normalization factors. In this case they correspond to metric measure spaces with an arbitrary diameter. We formulate Theorem 3 for long dendrons to be more general and capture those limits with an unbounded diameter too. The single main difference between this more general limit theory of finite trees and the theory discussed in this paper is that with no bound on the diameter we lose compactness: we will not be able to find convergent subsequences of any sequence of arbitrarily normalized finite trees. Janson in a recent paper, [7] found other normalization factors more suitable in various (often probabilistic) scenarios.

For the purposes of this paper we encourage the reader to concentrate on the case of dendrons. For this case it is instructive to note that one could define the domain A_D of a dendron $D = (T, d, \nu)$ as $A_D = T \times [0, 1/2]$ instead of the definition $A_D = T \times [0, \infty)$ above. Indeed, for the complementary set $B = T \times (1/2, \infty)$ we have $d_D(x, y) > 1$ for all $x, y \in B$, hence if D is a dendron, we have $\nu(B) = 0$.

1.1. **Examples.** A wealth of interesting examples are also presented in the recent paper of Janson [7]. Here we include a very short list of examples to serve as illustration. Some of these examples are taken from Janson's paper. All these examples are easy to work out, we leave the simple calculations to the reader. For more examples of limits of **random finite trees** using metric sampling see Section 9-14 of [7].

Example 1 (Example 7.1 [7])**.** Let P_n be the **path** on n vertices. These paths converge in metric sampling to the measured real tree (I, d, μ), where $I = [0, 1]$ is the unit interval with the Euclidean metric d and Lebesgue measure μ on the Borel sets of I. This can be realized as a dendron (I, d, ν), where ν is the push forward of the measure μ along the inclusion $\iota : I \to I \times [0, \infty)$ defined by $f(x) = (x, 0)$.

Example 2 (Example 7.2 [7])**.** For the simplest example where the limit is *not* a measured real tree, consider the sequence of **stars**, consisting of the n-vertex star $K_{1,n-1}$ for all n. These stars converge in metric sampling to the dendron (Υ, d, ν), where (Υ, d) is the one point metric space with point p and ν is concentrated on the single point $(p, 1/2)$.

Example 3 (Example 7.3 [7])**.** Let B_n be the **binary tree** of height $n - 1$ with $2^n - 1$ vertices. These binary trees also converge in metric sampling to the same dendron (Υ, d, ν) we saw in the previous example.

Example 4. We obtain a more complicated limit dendron if we replace the edges of the binary tree with paths of variable lengths. Let us obtain B'_n by replacing each edge e of the binary tree B_n from the preceding example with a path of length $\lfloor n^2/k^2 \rfloor$, where e connects a vertex of distance k from the root with one at distance $k - 1$. These finite trees form a convergent sequence in metric sampling. The limit dendron is (B, d, ν), where (B, d) is a

real tree obtained from the infinite binary tree by making each edge e of the combinatorial tree in distance $k - 1$ from the root into an interval of length C/k^2 where $C = \frac{1}{2} \sum_{i=1}^{\infty} 1/i^2$. The distribution ν is concentrated on the set $H \times \{0\}$ for the Cantor set H formed by the points of B at distance $1/2$ from the root (the "limit points" of the infinite branches) and it is uniform there.

Example 5. A limit dendron with a "more uniform" distribution can be obtained from **combs**. Let C_n be the n^2-vertex tree obtained from the n-vertex path P_n by attaching pairwise disjoint n-vertex paths at every vertex. These combs form a convergent sequence in metric sampling. The limit dendron is (I, d, ν), where (I, d) is the interval of length $1/3$ and ν is concentrated on the rectangle $I \times [0, 1/3]$ and is uniform there.

Example 6. The underlying real trees in the limit dendrons of the previous examples are all compact. Here is a simple way to get a non-compact underlying tree. Let D_n be the depth two finite tree where the root has n children: $P_1, \ldots P_n$ and P_i has 2^i children for $i = 1, \ldots, n$. The trees D_n also form a convergent sequence in metric sampling with the limit dendron being (S, d, ν), where the real tree (S, d) is a star consisting of the infinitely many intervals (r, p_i), each of length $1/4$. The distribution ν is concentrated on the points $(p_i, 1/4)$ with $\nu((p_i, 1/4)) = 2^{-i}$ for $i = 1, 2 \ldots$.

In Section 2 we introduce semi-measured real trees, a technical relaxation of measured real trees. In Sections 3 and 4 we recall the metric ultraproduct and the ultraproduct of measure spaces, respectively, especially as they apply to semi-measured real trees. We prove Theorems 1, 2 and 3 in Sections 5, 7 and 6, respectively.

2. Semi-measured real trees

Definition 2.1. We denote the open ball of radius r around a point x in a metric space by $B_x(r)$.

We call the triple (T, d, μ) a **semi-measured metric space** if (T, d) is a metric space, (T, μ) is a probability measure space and all the balls in (T, d) are μ-measurable. A semi-measured metric space (T, d, μ) is a **quasi metric measure space** (or **quasi mms** for short) if d is μ^2-measurable.

Remark 3. Note that both semi-measured metric spaces and a quasi metric measure spaces are relaxations of the well established notion of metric measure spaces. For metric measure spaces one first requires that the underlying metric space is separable and complete, second that the distribution is a probability Borel measure on this metric space. Here we dropped the first requirement and relaxed the second, allowing some Borel sets to be non-measurable. The requirement for the balls to be measurable in a semi-measured metric space is equivalent to requiring that the single variable distance function $d_x(y) = d(x, y)$ is μ-measurable for all points $x \in T$ and as such, it is weaker even than requiring that the bivariate distance function d is μ^2-measurable as needed for a quasi mms.

Note also, that the sampling measures τ_r are defined for a quasi mms but not (in general) for semi-metric measure spaces.

Definition 2.2. A **semi-measured real tree** is a semi-measured metric space (T, d, μ), where (T, d) is an $\mathbb{R}$-tree. Similarly, a **quasi measured real tree** is a quasi mms (T, d, μ), where (T, d) is a real tree.

If Y is non-empty, closed, connected subset of T in a real tree (T, d) we define the retraction $\pi_Y = \pi_Y^{T,d} : T \to Y$ by setting $\pi_Y(t)$ be the unique closest point to t in Y (the existence of which is stated in part (2) of the following lemma).

Lemma 2.1. *Let (T, d) be a real tree and Y be a non-empty, closed, connected subset of T.*

(1) *Y with the restriction of d is a real tree.*
(2) *The retraction π_Y is well defined and we have $\pi_Y(t) \in [t, y]$ for all $t \in T$ and $y \in Y$.*
(3) *If B is a connected component of $T \setminus Y$ and $x \in B$, then B is a $\pi_Y(x)$-branch of T.*
(4) *We have $d(x, y) = d(x, \pi_Y(x)) + d(\pi_Y(x), \pi_Y(y)) + d(\pi_Y(y), y)$ for all $x, y \in T$ unless x and y are in the same connected component of $T \setminus Y$. In particular, we have $d(\pi_Y(x), \pi_Y(y)) \leq d(x, y)$ for all $x, y \in T$ making π_Y continuous.*
(5) *Any branch B of T is in the σ-algebra generated by the balls of T, so it is μ-measurable for any semi-measured real tree (T, d, μ).*

Proof. A closed subspace Y of T is a complete metric space. If $[x, y] \subseteq Y$ for all $x, y \in Y$ then Y is a real tree, otherwise Y is not connected. This proves part (1).

Let us fix $t \in T$ and for a point $y \in Y$ define p_y to be the unique closest point to t in $Y \cap [y, t]$. As the segment $[y, t]$ is isometric to an interval and Y is closed, this exists. If $p_{y_1} \neq p_{y_2}$ for $y_1, y_2 \in Y$, then p_{y_1} and p_{y_2} can be connected inside Y (as Y is connected) and also outside (through t), a contradiction. Thus, all the points p_y coincide defining $\pi_Y(t)$ and proving part (2).

For part (3) consider the $\pi_Y(x)$-branch containing x. By part (2) it is disjoint from Y, hence it is B.

For $x, y \in T$, the union of the segments $[x, \pi_Y(x)]$, $[\pi_Y(x), \pi_Y(y)]$ and $[\pi_Y(y), y]$ connect x to y. If some two of these three segments intersect in more than their end points, then it must be the first and last ones (as the middle segment is contained in Y), and then x and y are in the same connected component of $T \setminus Y$. If no such non-trivial intersection occurs, then the union of the three segments is homeomorphic to an interval, so the union must be $[x, y]$ itself, proving the formula for $d(x, y)$ in part (4). If the formula applies, it implies the bound $d(\pi_Y(x), \pi_Y(y)) \leq d(x, y)$. Otherwise $\pi_Y(x) = \pi_Y(y)$ by part (3) and the bound holds again.

Let B be a p-branch of the real tree (T, d) and $x \in B$. Consider a sequence of intermediate points p_n of $[x, p]$ tending to p. It is easy to see that $B = \{y \in$

$T \mid \exists n : d(y, p_n) < d(y, p)\}$. This makes $B = \bigcup_{n,r}(B_{p_n}(r) \setminus B_p(r))$, where the union is taken for all n and all rational numbers $r > 0$. The formula proves part (5). $\qquad\square$

The goal of this section is to study semi-measured real trees (T, d, μ) and to define the associated long dendrons for them. Note that the associated long dendron is, in fact, a dendron whenever the essential diameter of (T, d, μ) is at most one, that is, when $\mu(\{x \in T \mid \mu(\{y \in T \mid d(x, y) > 1\}) > 0\}) = 0$.

Definition 2.3. Let (T, d, μ) be a semi-measured real tree. The point $p \in T$ is an **inner point** of T if there is no p-branch B of T such that $\mu(B) = 1$. The **core** of T, $\mathrm{Core}(T)$ is the closure of the set of inner points of T. We write π^T as a shorthand for the retraction $\pi_{\mathrm{Core}(T)}$.

For a semi-measured real tree we define the **associated long dendron** D and **associated projection** $\alpha : T \to A_D$ as follows. For $p \in T$ we set $\alpha(p) = (\pi^T(p), d(p, \pi^T(p)))$. We set $D = (\mathrm{Core}(T), d, \nu)$. Here we slightly abuse notation by denoting the restriction of d to the $\mathrm{Core}(T)$ by d again. We define the Borel probability measure ν as the push-forward of the measure μ along the map α.

The definition of inner point makes sense because the branches are measurable by Lemma 2.1(5).

In order to show that the definition of the associated long dendron makes sense we need to prove a series of lemmas. In Lemmas 2.2 and 2.3 we will show that $\mathrm{Core}(T)$ is non-empty and connected, respectively. As the core is closed by definition it is a real tree by Lemma 2.1(1) and so the retraction π^T exists and the associated projection α is also defined. For the definition of the measure ν in the associated long dendron to make sense we further need that the associated projection α is measurable if considered as a map from (T, μ) to the Borel space on A_D. This is stated in Lemma 2.5 below. Finally, we show that the associated long dendron is indeed a long dendron (that is, it satisfies the positivity condition) in Lemma 2.6 below.

Lemma 2.2. *The set* $\mathrm{Core}(T)$ *is non-empty. Furthermore, a branch B of T intersects* $\mathrm{Core}(T)$ *if and only if* $\mu(B) > 0$.

Proof. It is enough to prove the second statement as if there are no positive measure branches, then all points of T must be inner points. (This scenario is only possible if T has a single point only and thus it has no branches at all.)

Let $p \in T$ and let B be a p-branch. For any $q \in B$ the q-branch B' containing p satisfies $B' \supseteq T \setminus B$. Thus, if $\mu(B) = 0$, then $\mu(B') = 1$ and no point in B is an inner point. As B is open, this implies that $\mathrm{Core}(T)$ is disjoint from B as claimed. Now, let us assume that $\mu(B) > 0$. We will proceed by contradiction. Suppose that B contains no inner points. For $t \in B$ let B_t be the unique t-branch of measure 1, and let C_t be the t-branch containing p. Let $S = \{t \in B \mid B_t \neq C_t\}$. Note that $\mu(C_t) = 1$ if $t \in B \setminus S$ and $\mu(C_t) = 0$ if $t \in S$. Let $z = \sup\{d(p, t) \mid t \in S \cup \{p\}\}$. We claim that z is a maximum, that is, there exists $c \in S \cup \{p\}$ with $d(p, c) = z$. Otherwise we would have $c_i \in S$ such that $d(p, c_i)$ tend to $z > 0$ as i tends to infinity.

We have $\mu(\bigcap_{i=1}^{\infty} B_{c_i}) = 1$. Let $t \in \bigcap_{i=1}^{\infty} B_{c_i}$. All the points c_i are in the segment $[p, t]$, so they must tend to a point $c \in [p, t]$ with $d(p, c) = z$. We have $C_c = \bigcup_{i=1}^{\infty} C_{c_i}$ so we must have $\mu(C_c) = 0$ and $c \in S$ as claimed. Now, let $B' = B$ if $c = p$ (that is $S = \emptyset$) and $B' = B_c$ otherwise. Take an arbitrary point $y \in B'$ and a sequence of intermediate points y_i of the segment $[c, y]$ tending to c. Since $d(p, y_i) > d(p, c) = z$, we have that $\mu(C_{y_i}) = 1$. Notice that $\bigcap_{i=1}^{\infty} C_{y_i}$ is disjoint from the positive measure set B', therefore we cannot have $\mu(C_{y_i}) = 1$ for all i, leading to a contradiction $\qquad \square$

Lemma 2.3. *The set* $\mathrm{Core}(T)$ *is connected. Moreover, all intermediate points of the segment* $[p, q]$ *are inner points if* $p, q \in \mathrm{Core}(T)$.

Proof. Suppose now that x is an intermediate point in $[p, q]$ for some $p, q \in \mathrm{Core}(T)$. The points p and q lie in distinct x-branches and these x-branches have positive measure by Lemma 2.2. Thus, x must be an inner point as stated. $\qquad \square$

Let us fix a point $p_0 \in \mathrm{Core}(T)$ and let Q be the set of inner points q such that $d(p_0, q)$ is a rational number.

Lemma 2.4. *The real tree* $\mathrm{Core}(T)$ *is separable. In particular,* Q *is a countable dense set in* $\mathrm{Core}(T)$ *and furthermore* Q *is dense in any proper segment in* $\mathrm{Core}(T)$.

Proof. For $p_0 \neq q \in Q$, let B_q be the q-branch containing p_0 and $C_q = T \setminus B_q$. As q is inner we have $\mu(C_q) > 0$. Notice that for $p, q \in Q$ distinct with $d(p_0, p) = d(p_0, q)$ the sets C_p and C_q are disjoint. Therefore, Q contains a countable number of points in any fixed distance from p_0 and must itself be a countable set. Consider a pair of distinct points $p, q \in \mathrm{Core}(T)$ and let $t = \pi_{[p,q]}(p_0)$. All intermediate points $s \in [p, q]$ are inner points by Lemma 2.3. By Lemma 2.1 we have $d(p_0, s) = d(p_0, t) + d(t, s)$, and thus $Q \cap [p, q]$ is dense in $[p, q]$ as claimed.

The fact that Q is dense in any proper segment in $\mathrm{Core}(T)$ implies that Q is dense in $\mathrm{Core}(T)$ itself unless $\mathrm{Core}(T)$ is a singleton set. In this latter case $Q = \mathrm{Core}(T)$. $\qquad \square$

Remark 4. The same argument (with "inner" replaced with "intermediate") shows that every separable real tree has a countable subset that is dense in every proper segment.

Corollary 2.1. *If* $D = (T, d, \nu)$ *is a long dendron, then the real tree* (T, d) *is separable.*

Proof. Let μ be push forward of the measure ν along the projection from A_D to T. This makes (T, d, μ) a semi-measured real tree. By the definition of the long dendron, every branch of this real tree has positive measure. By Lemma 2.2, this implies that every branch intersects the core. But as the core is non-empty (by Lemma 2.2 again) and closed (by definition), this implies that the core is T itself. Now Lemma 2.4 proves the claim of the corollary. $\quad \square$

Lemma 2.5. *The associated projection* $\alpha : T \to A_D$ *is measurable.*

9

Proof. Recall that for $p \in T$ we have $\alpha(p) = (\pi^T(p), d(\pi^T(p), p))$. We prove that both coordinate functions are measurable. Here $d(\pi^T(p), p) = \min_{q \in \text{Core}(T)} d(p, q) = \inf_{q \in Q} d(p, q)$ by Lemma 2.4. The function $f_q(p) = d(p, q)$ is measurable by definition, therefore so is $d(\pi^T(p), p)$.

We further need that the map π^T considered from (T, μ) to the Borel space on $\text{Core}(T)$ is measurable. As $\text{Core}(T)$ is separable (Lemma 2.4) it is enough to check that the inverse image of the complement of a closed ball, namely the set $H = \{x \in T \mid d(\pi^T(x), p) > r\}$ is measurable for each $p \in \text{Core}(T)$ and $r > 0$. For $x \in H$ there exists $q \in Q \cap [p, \pi^T(x)]$ with $d(q, p) > r$ by Lemma 2.4. This makes $d(x, p) > d(x, q) + r$. Clearly, for points x outside H and $q \in Q$ we have $d(x, p) \leq d(x, \pi^T(x)) + d(\pi^T(x), p) \leq d(x, q) + r$. This proves $H = \bigcup_{q \in Q} \{x \in T \mid d(x, p) > d(x, q) + r\} = \bigcup_{q \in Q, s} (B_q(s) \setminus B_p(r + s))$, where s ranges over the positive rational numbers in the last expression. As the balls are μ-measurable by definition, so is H. $\square$

Lemma 2.6. *The associated long dendron $D = (\text{Core}(T), d, \nu)$ is indeed a long dendron.*

Proof. We have already seen that $(\text{Core}(T), d)$ is a real tree and ν is a Borel probability measure on A_D. It remains to show that for any branch B of $\text{Core}(T)$ we have $\nu(B \times [0, \infty)) > 0$. By the definitions of ν and α we have $\nu(B \times [0, \infty)) = \mu(\{x \in T \mid \pi^T(x) \in B\})$. Let B be a p-branch of $\text{Core}(T)$, then $\{x \in T \mid \pi^T(x) \in B\}$ is the p-branch B' of T containing B. B' intersects $\text{Core}(T)$, so by Lemma 2.2 we have $\mu(B') > 0$. $\square$

We call the connected components of $T \setminus \text{Core}(T)$ the **feathers** of T.

Lemma 2.7. *Let (T, d, μ) be a semi-measured real tree with associated long dendron $D = (\text{Core}(T), d, \nu)$ and associated projection α. The feathers of T are exactly the measure zero p-branches for points $p \in \text{Core}(T)$. We have $d(x, y) = d_D(\alpha(x), \alpha(y))$ for $x, y \in T$ unless x and y are in the same feather of T.*

Proof. Both statements follow directly from Lemmas 2.1 and 2.2. $\square$

A **measured real tree** is a metric measure space (T, d, μ) where (T, d) is a real tree. As we saw, a measured real tree is also a semi-measured real tree, so the associated long dendron and the associated projection are defined.

For a function $\alpha : A \to B$ and $r \in \mathbb{N}$ we write α^r for the function that acts coordinate-wise on A^r, that is $\alpha^r : A^r \to B^r$, $\alpha^r(x_1, \ldots, x_r) = (\alpha(x_1), \ldots, \alpha(x_r))$.

Lemma 2.8. *For a quasi mms (T, d, μ) and associated long dendron D we have $\tau_r(T, d, \mu) = \tau_r(D)$ for all $r \geq 1$.*

Proof. Let $D = (\text{Core}(T), d, \nu)$ and let $\alpha : T \to A_D$ be the associated projection. By Lemma 2.7 we have $d(x, y) = d_D(\alpha(x), \alpha(y))$ for $x, y \in T$ unless x and y are in the same feather of T. Also by the same lemma the feathers have zero measure, so we have $\mu(\{y \in T \mid d(x, y) \neq d_D(\alpha(x), \alpha(y))\}) = 0$ for all $x \in T$ and hence $\mu^2(\{(x, y) \in T^2 \mid d(x, y) \neq d_D(\alpha(x), \alpha(y))\}) = 0$. As a consequence, we have $\mu^r(X) = 0$ for the set $X = \{x \in T^r \mid \rho_r^{T,d}(x) \neq \rho_r^{A_D, d_D}(\alpha^r(x))\}$.

For a Borel set $H \subseteq M_r$ we have $\tau_r(T, d, \mu)(H) = \mu^r(\{x \in T^r \mid \rho_r^{T,d}(x) \in H\})$ and $\tau_r(D)(H) = \nu^r(\{x \in A_D^r \mid \rho_r^{A_D, d_D}(x) \in H\}) = \mu^r(\{x \in T^r \mid \rho_r^{A_D, d_D}(\alpha^r(x)) \in H\})$. As the symmetric difference of the sets $\{x \in T^r \mid \rho_r^{T,d}(x) \in H\}$ and $\{x \in T^r \mid \rho_r^{A_D, d_D}(\alpha^r(x)) \in H\}$ is contained in the zero measure set X, their μ^r measures agree, so $\tau_r(T, d, \mu)(H) = \tau_r(D)(H)$ as needed. $\qquad\square$

Remark 5. Given a (long) dendron D it is natural to look for a quasi measured real tree T_D with D as its associated long dendron. By Lemma 2.8 we could use T_D in place of D for the sampling limit in Theorems 1 and 2. If the construction of T_D is canonical we can also ensure the limit is unique as in Theorem 3.

Let $D = (T, d, \nu)$ be a long dendron. Recall that $d_D : A_D^2 \to \mathbb{R}$ is not a metric as $d_D(x, x)$ is often positive. As a first attempt to fix this one could consider the distance function d' as defined in Equation (10) in Section 6. This is indeed a distance and (A_D, d') is a real tree. It is not separable (unless T is trivial), so (A_D, d', ν) is typically not a metric measure space, but it is a quasi measured real tree so both $\tau_r(D)$ and $\tau_r(A_D, d', \nu)$ are well defined. If $\nu(\{u\} \times (0, \infty)) = 0$ for all $u \in T$, then d' and d_D differ in a measure zero subset of A_D^2. In this case we have $\tau_r(D) = \tau_r(A_D, d', \nu)$ and further the associated long dendron for the measured real tree (A_D, d', ν) is D and the associated projection is the identity on A_D. Note however, that ν is a Borel measure on A_D with its product structure, but unless $|T| = 1$ it is *not* a Borel measure on the real tree (A_D, d') as the latter space is non-separable with many not ν-measurable open sets.

The simple approach above does not work if $\nu(\{u\} \times (0, \infty)) > 0$ for some points $u \in T$. The following, slightly more involved method always works. We define the quasi measured real tree $T_D = (T^*, d^*, \mu)$ as follows.

$$T^* = \{(x, y, z) \mid x \in T, \ (y, z) \in \{(0,0)\} \cup ((0, 1) \times (0, \infty))\}$$

$$d^*((x_1, y_1, z_1), (x_2, y_2, z_2)) = \begin{cases} |z_1 - z_2| & \text{if } (x_1, y_1) = (x_2, y_2) \\ d(x_1, x_2) + z_1 + z_2 & \text{otherwise.} \end{cases}$$

We define μ as a Borel probability measure on T^* considered as subset of the product $T \times [0, 1) \times [0, \infty)$. As above, this is *not* a Borel measure on the real tree (T^*, d^*) because many open sets in the real tree are not measurable. For $H \subseteq T^*$ which is Borel in the former sense we define $\mu(H)$ through the function $f_H : A_D \to [0, 1]$ as follows. We denote the Lebesgue measure on the reals by λ.

$$f_H(x, z) = \begin{cases} 0 & \text{if } z = 0 \text{ and } (x, 0, 0) \notin H \\ 1 & \text{if } z = 0 \text{ and } (x, 0, 0) \in H \\ \lambda(\{y \in (0, 1) \mid (x, y, z) \in H\}) & \text{if } z > 0 \end{cases}$$

$$\mu(H) = \int_{A_D} f_H \, d\nu$$

Now T_D is a quasi measured real tree which has D as its associated long dendron and $\alpha : T^* \to A_D$ given by $\alpha(x, y, z) = (x, z)$ as its associated projection. In particular, we have $\tau_r(T_D) = \tau_r(D)$ for all r.

3. The ultraproduct and the metric ultraproduct

Let us recall the definitions of the ultralimit and the ultraproduct. We will use both the set theoretical and metric ultraproducts as well as the ultraproduct of probability measures (see next section). When referring to ultralimits or ultraproducts we will always use the same fixed nontrivial ultrafilter ω on the set $\mathbb{N}$ of positive integers. We use ω-**few** to refer to any subset of $\mathbb{N}$ that is not in ω. We allow sequences $(x_n)_{n \in \mathbb{N}}$ to be undefined for ω-few indices n.

For any Hausdorff space H and points $x_n \in H$ for $n \in \mathbb{N}$ we define the ultralimit $\lim_\omega x_n$ to be $x \in H$ such that x_n is outside any fixed neighborhood of x for only ω-few indices n. If the ultralimit exists it is clearly unique, and in case H is compact, it does exist for every sequence x_n. Also, if $\lim_{n \to \infty} x_n$ exists, then so does $\lim_\omega x_n$ and they agree. Further, if $\lim_\omega x_n$ exists and $g : H \to H'$ is a continuous function to another Hausdorff space H', then $\lim_\omega g(x_n) = g(\lim_\omega x_n)$.

The **set-theoretical ultraproduct** $\mathbf{T} = \prod_\omega T_n$ of the sequence of sets $(T_n)_{n \in \mathbb{N}}$ consists of the equivalence classes of sequences $(x_n)_{n \in \mathbb{N}}$, $x_n \in T_n$, where $(x_n)_{n \in \mathbb{N}}$ and $(y_n)_{n \in \mathbb{N}}$ are equivalent if $x_n \neq y_n$ for ω-few indices n. We denote the class of the sequence $(x_n)_{n \in \mathbb{N}}$ by $[(x_n)_{n \in \mathbb{N}}]$.

Let $A_n \subseteq T_n$ for each n. Clearly, each element $\mathbf{x}$ of $\mathbf{A} = \prod_\omega A_n$ (as an equivalence class) is contained in a distinct element of $\mathbf{T} = \prod_\omega T_n$. We identify $\mathbf{x}$ with the element of $\mathbf{T}$ containing it making $\mathbf{A} \subseteq \mathbf{T}$.

For sets S_n, T_n we identify the ultraproduct of their direct products $\prod_\omega (S_n \times T_n)$ with the direct product $\prod_\omega S_n \times \prod_\omega T_n$ by identifying $[(x_n, y_n)_{n \in \mathbb{N}}]$ in the former space with $([(x_n)_{n \in \mathbb{N}}], [(y_n)_{n \in \mathbb{N}}])$ in the latter. We make the same identification in ultraproducts of products of finitely many factors. In particular, we identify $\prod_\omega (S_n^k)$ with $(\prod_\omega S_n)^k$ for $k \geq 2$.

Let H be a Hausdorff space and $f_n : T_n \to H$ be arbitrary functions. We write $\lim_\omega f_n$ for the function $\mathbf{f}$ defined by $\mathbf{f}([(x_n)_{n \in \mathbb{N}}]) = \lim_\omega f_n(x_n)$. The ultralimit does not depend on the choice of the sequence $(x_n)_{n \in \mathbb{N}}$ representing the class $[(x_n)_{n \in \mathbb{N}}]$, but it may be undefined for some classes if H is not compact. In this case $\mathbf{f}$ is only defined on a subset of $\prod_\omega T_n$.

The metric ultraproduct of metric spaces (T_n, d_n) was introduced by van den Dries and Wilkie in [11]. First we define $\mathbf{d} = \lim_\omega d_n$. For this we consider the distance functions d_n with values in the compact space $\mathbb{R}^\infty = \mathbb{R} \cup \{-\infty, \infty\}$. This makes $\mathbf{d} : \prod_\omega (T_n^2) \to \mathbb{R}^\infty$. We identified $\prod_\omega (T_n^2)$ with $\mathbf{T}^2$, where $\mathbf{T} = \prod_\omega T_n$, so $\mathbf{d} : \mathbf{T}^2 \to \mathbb{R}^\infty$. Clearly, $\mathbf{d}$ is symmetric, non-negative and satisfies the triangle inequality, that is, $\mathbf{d}$ is a **pseudometric** possibly containing zero and infinite distances. The **metric ultraproduct** $\hat{\mathbf{T}} = \hat{\prod}_\omega T_n$ is obtained from $\mathbf{T}$ by factoring out the equivalence relation of having zero distance. Let $\hat{\pi} : \mathbf{T} \to \hat{\mathbf{T}}$ be the **natural projection** mapping a point in $\mathbf{T}$ to its equivalence class in $\hat{\mathbf{T}}$. We write $[\widehat{(p_n)_{n \in \mathbb{N}}}]$ to denote $\hat{\pi}([(p_n)_{n \in \mathbb{N}}])$. Clearly, the pseudodistance $\mathbf{d}$ defines a pseudodistance $\hat{\mathbf{d}}$ on $\hat{\mathbf{T}}$ by setting $\hat{\mathbf{d}}(\hat{\pi}(\mathbf{x}), \hat{\pi}(\mathbf{y})) = \mathbf{d}(\mathbf{x}, \mathbf{y})$. Here $\hat{\mathbf{d}}$ is a pseudometric in which distinct points have a positive distance but infinite distances may still appear. Clearly, $\mathrm{diam}(\hat{\mathbf{T}}, \hat{\mathbf{d}}) = \lim_\omega \mathrm{diam}(T_n, d_n)$. If this value is finite, we call the sequence $(T_n, d_n)_{n \in \mathbb{N}}$ **uniformly bounded**.

In this case $(\hat{\mathbf{T}}, \hat{\mathbf{d}})$ is a metric space (no infinite distances). In general, "having finite distance" is an equivalence relation on $\hat{\mathbf{T}}$ and each equivalence class $\hat{\mathbf{X}}$ is made into a metric space by the restriction of $\hat{\mathbf{d}}$. We call these metric spaces the **clusters** of the metric ultraproduct $\hat{\mathbf{T}}$.

Let $A_n \subseteq T_n$ for each n. As we identified $\prod_\omega A_n$ with a subset of $\mathbf{T}$ we similarly identify $\hat{\prod}_\omega A_n$ with the corresponding subset of $\hat{\mathbf{T}}$.

Note that the ultraproduct objects will always be typeset in bold and we will put the metric ultraproduct objects under the ˆ sign. Recall that we use normalized distance on finite trees and all those distances are bounded by 1. Therefore, sequences of finite trees are uniformly bounded. We could have restricted attention to uniformly bounded sequences of metric spaces in this paper to avoid dealing with several clusters. We allow unbounded diameter for more generality and because having several clusters does not significantly increase complexity.

Lemma 3.1. *Any cluster $\hat{\mathbf{X}}$ of the metric ultraproduct $(\hat{\mathbf{T}}, \hat{\mathbf{d}})$ of real trees (T_n, d_n) is a real tree. For two points $\hat{\mathbf{x}} = [\widehat{(x_n)_{n\in\mathbb{N}}}]$ and $\hat{\mathbf{y}} = [\widehat{(y_n)_{n\in\mathbb{N}}}]$ in $\hat{\mathbf{X}}$ we have $[\hat{\mathbf{x}}, \hat{\mathbf{y}}] = \hat{\prod}_\omega [x_n, y_n]$.*

Proof. As $\hat{\mathbf{X}}$ is obtained as a metric ultraproduct, it is complete, see [11]. For $\hat{\mathbf{x}}, \hat{\mathbf{y}} \in \hat{\mathbf{X}}$ the segments $[x_n, y_n]$ in T_n are isometric to a real interval of length $d_n(x_n, y_n)$ and we have $\lim_\omega d_n(x_n, y_n) = \hat{\mathbf{d}}(\hat{\mathbf{x}}, \hat{\mathbf{y}}) < \infty$. Therefore, the metric ultraproduct $\hat{\mathbf{l}} = \hat{\prod}_\omega [x_n, y_n]$ of these segments is isometric to a real interval of length $\hat{\mathbf{d}}(\hat{\mathbf{x}}, \hat{\mathbf{y}})$. It is contained in $\hat{\mathbf{X}}$ and contains $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$, so all we need to establish to prove the lemma is that all the intermediate points of $\hat{\mathbf{l}}$ separate $\hat{\mathbf{x}}$ from $\hat{\mathbf{y}}$.

So let $\hat{\mathbf{p}} = [\widehat{(p_n)_{n\in\mathbb{N}}}] \in \hat{\mathbf{l}} \setminus \{\hat{\mathbf{x}}, \hat{\mathbf{y}}\}$ with $p_n \in [x_n, y_n]$. Let A_n be the p_n-branch in T_n containing x_n and $B_n = T_n \setminus A_n$. (For the ω-few indices n where $p_n = x_n$ the sets A_n and B_n are not defined.) Let $\hat{\mathbf{A}} = \hat{\prod}_\omega A_n$ and $\hat{\mathbf{B}} = \hat{\prod}_\omega B_n$. We have $\hat{\mathbf{A}} \cup \hat{\mathbf{B}} = \hat{\mathbf{T}}$ as for any $\hat{\mathbf{z}} = [\widehat{(z_n)_{n\in\mathbb{N}}}] \in \hat{\mathbf{T}}$ we have either $z_n \notin A_n$ for ω-few indices making $\hat{\mathbf{z}} \in \hat{\mathbf{A}}$ or $z_n \notin B_n$ for ω-few indices making $\hat{\mathbf{z}} \in \hat{\mathbf{B}}$. Take any $\hat{\mathbf{a}} \in \hat{\mathbf{A}}$ and $\hat{\mathbf{b}} \in \hat{\mathbf{B}}$. From the similar equations in T_n we have $\hat{\mathbf{d}}(\hat{\mathbf{a}}, \hat{\mathbf{b}}) = \hat{\mathbf{d}}(\hat{\mathbf{a}}, \hat{\mathbf{p}}) + \hat{\mathbf{d}}(\hat{\mathbf{p}}, \hat{\mathbf{b}})$. This means that any point $\hat{\mathbf{a}} \in \hat{\mathbf{A}} \setminus \{\hat{\mathbf{p}}\}$ is separated from $\hat{\mathbf{B}}$ by the positive distance of $\hat{\mathbf{d}}(\hat{\mathbf{a}}, \hat{\mathbf{p}})$ and similarly any point $\hat{\mathbf{b}} \in \hat{\mathbf{B}} \setminus \{\hat{\mathbf{p}}\}$ is separated from $\hat{\mathbf{A}}$ by the positive distance of $\hat{\mathbf{d}}(\hat{\mathbf{p}}, \hat{\mathbf{b}})$. Therefore the sets $(\hat{\mathbf{A}} \cap \hat{\mathbf{X}}) \setminus \{\hat{\mathbf{p}}\}$ and $(\hat{\mathbf{B}} \cap \hat{\mathbf{X}}) \setminus \{\hat{\mathbf{p}}\}$ are disjoint and open and form a partition of $\hat{\mathbf{X}} \setminus \{\hat{\mathbf{p}}\}$. The point $\hat{\mathbf{x}}$ is in the former set, while $\hat{\mathbf{y}}$ is in the latter, so $\hat{\mathbf{p}}$ separates them in $\hat{\mathbf{X}}$ as claimed. $\square$

4. The ultraproduct of probability spaces

Let us recall the ultraproduct of probability measures. Let μ_n be a probability measure on a σ-algebra $\mathcal{A}_n$ over a set T_n for each $n \in \mathbb{N}$. The ultraproduct sets $\prod_\omega A_n$ with $A_n \in \mathcal{A}_n$ form a Boolean algebra $\mathcal{P}$ on $\mathbf{T} = \prod_\omega T_n$. Additionally, we have a finitely additive measure $\mu_{\mathcal{P}}$ on $\mathcal{P}$ given by $\mu_{\mathcal{P}}(\prod_\omega A_n) =$

$\lim_\omega \mu_n(A_n)$. This finitely additive measure can be extended to a σ-algebra containing $\mathcal{P}$ (see [4]) as follows. We call $\mathbf{N} \subset \mathbf{T}$ a **nullset** if for any $\varepsilon > 0$ there exists an element $\mathbf{A} \in \mathcal{P}$ such that $\mathbf{N} \subseteq \mathbf{A}$ and $\mu_\mathcal{P}(\mathbf{A}) < \varepsilon$. A set $\mathbf{M} \subset \mathbf{T}$ is called **measurable** if there exists $\mathbf{P} \in \mathcal{P}$ such that the symmetric difference $\mathbf{P} \triangle \mathbf{M}$ is a nullset. The family of measurable sets $\mathcal{M}_\mathbf{T}$ form a σ-algebra with a probability measure $\mu = \prod_\omega \mu_n$ (the ultraproduct of the measures μ_n), where we define $\mu(\mathbf{M}) = \mu_\mathcal{P}(\mathbf{P})$. Hence, we made the ultraproduct space $\mathbf{T}$ into a probability measure space $(\mathbf{T}, \mu)$. Note that the paper [4] only dealt with ultraproducts of finite spaces with uniform probability measures, so the following lemma is only stated there for this special case. However, the proof of the lemma does not use this restriction on the factors of the ultraproduct and the same proof applies verbatim for the more general case stated here.

Lemma 4.1 (Lemma 5.1 [4]). *Let (T_n, μ_n) be probability measure space and $f_n : T_n \to \mathbb{R}^\infty$ be a μ_n-measurable function for $n \in \mathbb{N}$. Let $\mathbf{T} = \prod_\omega T_n$, $\mu = \prod_\omega \mu_n$. In this case the function $\mathbf{f} = \lim_\omega f_n : \mathbf{T} \to \mathbb{R}^\infty$ is μ-measurable. If $(f_n)_{n \in \mathbb{N}}$ is uniformly bounded (that is $\sup_{n,x} |f_n(x)| < \infty$), then we also have*

$$\lim_\omega \int_{T_n} f_n \, d\mu_n = \int_\mathbf{T} \mathbf{f} \, d\mu \,.$$

Let (S_n, μ_n) and (T_n, ν_n) be probability measure spaces for each $n \in \mathbb{N}$. Let $\mathbf{S} = \prod_\omega S_n$, $\mathbf{T} = \prod_\omega T_n$. Recall that we have identified $\mathbf{S} \times \mathbf{T}$ with $\prod_\omega (S_n \times T_n)$. We have two probability measures on this set. First we have $\mu_\mathbf{S} \times \mu_\mathbf{T}$, where $\mu_\mathbf{S} = \prod_\omega \mu_n$ and $\mu_\mathbf{T} = \prod_\omega \nu_n$. But we also have $\mu_{\mathbf{S} \times \mathbf{T}} = \prod_\omega (\mu_n \times \nu_n)$. It is easy to see that the letter measure extends the former, that is, any $\mu_\mathbf{S} \times \mu_\mathbf{T}$-measurable set H is also $\mu_{\mathbf{S} \times \mathbf{T}}$-measurable and we have $\mu_{\mathbf{S} \times \mathbf{T}}(H) = (\mu_\mathbf{S} \times \mu_\mathbf{T})(H)$.

In general, a $\mu_{\mathbf{S} \times \mathbf{T}}$-measurable set is not necessarily $\mu_\mathbf{S} \times \mu_\mathbf{T}$-measurable, but we still have the following form of Fubini's theorem (see [8] and also [12])

Lemma 4.2. *In the setting above for any $\mu_{\mathbf{S} \times \mathbf{T}}$-measurable set H the sections $H_\mathbf{x} = \{\mathbf{y} \mid (\mathbf{x}, \mathbf{y}) \in H\}$ are $\mu_\mathbf{T}$-measurable unless $\mathbf{x} \in \mathbf{Z}$ for some $\mathbf{Z} \subseteq \mathbf{S}$ with $\mu_\mathbf{S}(\mathbf{Z}) = 0$ and we have*

$$\mu_{\mathbf{S} \times \mathbf{T}}(H) = \int_{\mathbf{S} \setminus \mathbf{Z}} \mu_\mathbf{T}(H_\mathbf{x}) \, d\mu_\mathbf{S}(x).$$

Let (T_n, d_n, μ_n) be semi-measured metric spaces. Let $\mathbf{T} = \prod_\omega T_n$ be the set theoretic ultraproduct, let $(\hat{\mathbf{T}}, \hat{\mathbf{d}})$ be the metric ultraproduct of the metric spaces (T_n, d_n) with $\hat{\pi} : \mathbf{T} \to \hat{\mathbf{T}}$ being the natural projection.

Let μ be the ultraproduct of the measures μ_n. We call the sets $H \subseteq \hat{\mathbf{T}}$ **proper** if $\hat{\pi}^{-1}(H)$ is μ-measurable and define $\hat{\mu}(H) = \mu(\hat{\pi}^{-1}(H))$ in this case. Clearly, the proper sets form a σ-algebra on $\hat{\mathbf{T}}$ and $\hat{\mu}$ is a probability measure on it, the push-forward of μ along $\hat{\pi}$.

Note that if $S_n \subseteq T_n$ is μ_n-measurable for $n \in \mathbb{N}$ and $\hat{\mathbf{S}} = \hat{\prod}_\omega S_n \subseteq \hat{\mathbf{T}}$ is proper, then we have $\hat{\mu}(\hat{\mathbf{S}}) \geq \lim_\omega \mu_n(S_n)$. Indeed, $\lim_\omega \mu_n(S_n) = \mu(\mathbf{S})$, where $\mathbf{S} = \prod_\omega S_n$ while $\hat{\mu}(\hat{\mathbf{S}}) = \mu(\{\mathbf{p} \in \mathbf{T} \mid \hat{\pi}(\mathbf{p}) \in \hat{\pi}(\mathbf{S})\})$ and the set $\{\mathbf{p} \in \mathbf{T} \mid \hat{\pi}(\mathbf{p}) \in \hat{\pi}(\mathbf{S})\}$ clearly contains $\mathbf{S}$.

Let $\mathbf{x} = [(x_n)_{n \in \mathbb{N}}] \in \mathbf{T}$. By definition, all the single variable distance functions $f_n : T_n \to \mathbb{R}$ given by $f_n(y) = d_n(x_n, y)$ are μ_n-measurable. By Lemma 4.1 the function $\mathbf{f} = \lim_\omega f_n$ is μ-measurable. Here $\mathbf{f} : \mathbf{T} \to \mathbb{R}^\infty$ is the function $\mathbf{f}(\mathbf{y}) = \mathbf{d}(\mathbf{x}, \mathbf{y})$. Hence its factor $\hat{\mathbf{f}} : \hat{\mathbf{T}} \to \mathbb{R}^\infty$ given by $\hat{\mathbf{f}}(\hat{\mathbf{y}}) = \hat{\mathbf{d}}(\hat{\pi}(\mathbf{x}), \hat{\mathbf{y}})$ is $\hat{\mu}$-measurable.

This, in particular, means that each cluster of $\hat{\mathbf{T}}$ is proper. We call a cluster **dominant** if its measure is 1. If a dominant cluster exists, then we say that the sequence (T_n, d_n, μ_n) is **essentially bounded**. In this case we simply disregard the rest of the metric ultraproduct and consider $(\hat{\mathbf{X}}, \hat{\mathbf{d}}_{\hat{\mathbf{X}}}, \hat{\mu}_{\hat{\mathbf{X}}})$ as the ultraproduct of the semi-measurable metric spaces (T_n, d_n, μ_n), where $\hat{\mathbf{X}}$ is the dominant cluster and $\hat{\mathbf{d}}_{\hat{\mathbf{X}}}$ and $\hat{\mu}_{\hat{\mathbf{X}}}$ are the restrictions of $\hat{\mathbf{d}}$ and $\hat{\mu}$, respectively. By the above observation the ultraproduct of essentially bounded semi-measured metric spaces is a semi-measured metric space. Note that if (T_n, d_n) is uniformly bounded, then (T_n, d_n, μ_n) is essentially bounded, but the converse often fails.

Remark 6. The ultraproduct of an essentially bounded sequence of metric measure spaces (or even of a uniformly bounded sequence of metric measure spaces) is typically not a metric measure space, not even a quasi mms. A simple example of this phenomenon is obtained from random graphs. Let us associate with the n vertex graph G_n the metric measure space X_n on the n vertices of G_n with the uniform distribution, where the distance between adjacent vertices is 1 and between non-adjacent vertices it is 2. The diameter of these spaces is bounded by 2, so the metric measure spaces X_n form a uniformly bounded sequence. Their ultraproduct is clearly non-separable. Furthermore, with probability 1 for uniform random graphs G_n, the ultraproduct is not even a quasi mms.

We introduced semi-measured metric spaces because the ultraproduct of an essentially bounded sequence of semi-measured metric spaces is always a semi-measured metric space. However, this detour is not necessary. It is not hard to see that the ultraproduct of an essentially bounded sequence of measured real trees (or even quasi measured real trees) is always a quasi measured real tree. We could simply use this ultraproduct as the sampling limit (as opposed to the associated long dendron) by Lemma 2.8. But this approach does not give a unique sampling limit and the non-separability also complicates things. If one still finds a quasi measured real tree as a more esthetically pleasing sampling limit, this can be achieved without losing uniqueness through the process explained in Remark 5.

5. The proof of Theorem 1

Let T be a nontrivial finite tree. Consider the real tree (T', d) obtained from T by turning its edges into equal length real segments and letting d be the normalized distance such that $\operatorname{diam}(T', d) = 1$. Let μ be the Borel probability measure on T' that is concentrated on the vertices of T and is uniform there. Now (T', d, μ) is a measured real tree and we clearly have $\tau_r(T', d, \mu) = \tau_r(T)$ for each $r \geq 1$.

Theorem 1 is a consequence of Theorem 4 below. Indeed, if $(T_n)_{n\in\mathbb{N}}$ is a convergent sequence of finite trees, then the corresponding sequence (T'_n, d_n, μ_n) of measured real trees is uniformly bounded, so Theorem 4 finds their limit long dendron. This long dendron D is, in fact, a dendron, since $\mathrm{diam}(T'_n, d_n, \mu_n) = 1$ for all $n \in \mathbb{N}$, so $\tau_2(D)$ is concentrated on matrices with no entry exceeding 1.

Theorem 4. *Let $(T_n, d_n, \mu_n)_{n\in\mathbb{N}}$ be an essentially bounded sequence of measured real trees. For any $r \in \mathbb{N}$ we have*

$$\lim_\omega \tau_r(T_n, d_n, \mu_n) = \tau_r(D),$$

where D is the long dendron associated with the semi-measured real tree $(\hat{\mathbf{X}}, \hat{\mathbf{d}}, \hat{\mu})$ obtained as the ultraproduct of the measured real trees (T_n, d_n, μ_n) and the ultralimit is understood in the space $\mathrm{Prob}(M_r)$ with the weak topology.

Proof. Let $\hat{\pi} : \mathbf{T} \to \hat{\mathbf{T}}$ be the natural projection from the set theoretic ultraproduct $\mathbf{T} = \prod_\omega T_n$ to the metric ultraproduct $\hat{\mathbf{T}} = \hat{\prod}_\omega T_n$. Note that $\hat{\mathbf{X}}$ is the dominant cluster of $\hat{\mathbf{T}}$. Let $D = (X^*, \hat{\mathbf{d}}, \nu)$, and let the associated projection be $\alpha : \hat{\mathbf{X}} \to A_D$. For the proof we need to show that for any bounded continuous function $g : M_r \to \mathbb{R}$ we have

$$(1) \qquad \int_{M_r} g\, d\tau_r(D) = \lim_\omega \int_{M_r} g\, d\tau_r(T_n, d_n, \mu_n).$$

Let us set $\mu = \prod_\omega \mu_n$ and $\mathbf{X} = \hat{\pi}^{-1}(\hat{\mathbf{X}})$. For the left hand side of Equation (1) we use the fact that we defined several of our measures as push-forwards of other measures.

$$
\begin{aligned}
(2) \qquad \int_{M_r} g(z)\, d\tau_r(D)(z) &= \int_{A_D^r} g(\rho_r^{A_D, d_D}(x))\, d\nu^r(x) \\
(3) \qquad &= \int_{\hat{\mathbf{X}}^r} g(\rho_r^{A_D, d_D}(\alpha(\hat{\mathbf{x}})))\, d\hat{\mu}^r(\hat{\mathbf{x}}) \\
(4) \qquad &= \int_{\mathbf{X}^r} g(\rho_r^{A_D, d_D}(\alpha(\hat{\pi}(\mathbf{x}))))\, d\mu^r(\mathbf{x}).
\end{aligned}
$$

Namely, the equation in line (2) holds as $\tau_r(D)$ is defined as the push-forward of ν^r along $\rho_r^{A_D, d_D}$. Line (3) follows as ν is defined as a push-forward of $\hat{\mu}$ along α. Finally, line (4) follows as $\hat{\mu}$ was itself defined as the push-forward of μ along $\hat{\pi}$. Here x is an r-tuple from A_D, $\hat{\mathbf{x}}$ is an r-tuple from $\hat{\mathbf{X}}$ and $\mathbf{x}$ is an r-tuple from $\mathbf{X}$. The functions α and $\hat{\pi}$ act on r-tuples coordinate-wise.

For the right hand side of Equation (1) we have a longer sequence of equations. We use some of the same notations as above and will explain each line

16

separately together with any additional notation introduced there.

$$(5)\,\lim_\omega \int_{M_r} g(z)d\tau_r(T_n, d_n, \mu_n)(z) \;=\; \lim_\omega \int_{T_n^r} g(\rho_r^{T_n, d_n}(x_n))\, d\mu_n^r(x_n)$$

$$(6)\qquad\qquad\qquad\qquad =\; \int_{\mathbf{T}^r} \lim_\omega g(\rho_r^{T_n, d_n}(x_n))\, d\mu^{(r)}([(x_n)_{n\in\mathbb{N}}])$$

$$(7)\qquad\qquad\qquad\qquad =\; \int_{\mathbf{X}^r} \lim_\omega g(\rho_r^{T_n, d_n}(x_n))\, d\mu^{(r)}([(x_n)_{n\in\mathbb{N}}])$$

$$(8)\qquad\qquad\qquad\qquad =\; \int_{\mathbf{X}^r} g(\rho_r^{\mathbf{X}, \mathbf{d}}(\mathbf{x}))\, d\mu^{(r)}(\mathbf{x})$$

$$(9)\qquad\qquad\qquad\qquad =\; \int_{\mathbf{X}^r} g(\rho_r^{A_D, d_D}(\alpha(\hat{\pi}(\mathbf{x}))))\, d\mu^{(r)}(\mathbf{x})$$

Line (5) follows again from the definition of $\tau_r(T_n, d_n, \mu_n)$ as push-forward measure.

To obtain line (6) we apply Lemma 4.1 to the measure spaces (T_n^r, μ_n^r). We have $\prod_\omega T_n^r = \mathbf{T}^r$. We wrote $\mu^{(r)}$ to denote $\prod_\omega \mu_n^r$. Recall that this measure is an extension of μ^r.

We have $\mu(\mathbf{X}) = \hat{\mu}(\hat{\mathbf{X}}) = 1$ and therefore $\mathbf{X}^r$ is a full measure subset of $\mathbf{T}^r$, so restricting the domain to it, as done in line (7), does not affect the integral.

Let $\mathbf{d}$ be the restriction of $\lim_\omega d_n$ to $\mathbf{X}^2$. Note that $\mathbf{d}$ does not take infinite values. Let $\mathbf{x} = (\mathbf{x}^1, \ldots, \mathbf{x}^r) \in \mathbf{X}^r$ with $\mathbf{x}^i = [(x_n^i)_{n\in\mathbb{N}}]$ and $x_n^i \in T_n$. Let us write $x_n = (x_n^1, \ldots, x_n^r)$. We have

$$\begin{aligned}
\lim_\omega \rho_r^{T_n, d_n}(x_n) &= \lim_\omega (d_n((x_n)_i, (x_n)_j))_{1\leq i,j\leq r} \\
&= (\lim_\omega d_n((x_n)_i, (x_n)_j))_{1\leq i,j,\leq r} \\
&= (\mathbf{d}(\mathbf{x}_i, \mathbf{x}_j))_{1\leq i,j\leq r} \\
&= \rho_r^{\mathbf{X}, \mathbf{d}}(\mathbf{x})\,.
\end{aligned}$$

As g is continuous, this implies

$$g(\rho_r^{\mathbf{X}, \mathbf{d}}(\mathbf{x})) = g(\lim_\omega \rho_r^{T_n, d_n}(x_n)) = \lim_\omega g(\rho_r^{T_n, d_n}(x_n))\,,$$

so we have the exact same integral in line (8) as we had in line (7).

Consider a point $\mathbf{x} \in \mathbf{X}^r$ where the integrands of lines (8) and (9) differ. Let $\hat{\mathbf{x}} = \hat{\pi}(\mathbf{x}) \in \hat{\mathbf{X}}^r$. We must have $\rho_r^{\mathbf{X}, \mathbf{d}}(\mathbf{x}) \neq \rho_r^{A_D, d_D}(\alpha^r(\hat{\mathbf{x}}))$. Both sides are matrices in M_r with zero entries in the diagonal. The off-diagonal entry in position (i, j) is $\mathbf{d}(\mathbf{x}_i, \mathbf{x}_j)$ and $d_D(\alpha(\hat{\mathbf{x}}_i), \alpha(\hat{\mathbf{x}}_j))$, respectively. Consider the set $\mathbf{H}_{i,j} \subseteq \mathbf{X}^r$ consisting of points where these values differ and project it to coordinates i and j. This projection from $(\mathbf{T}^r, \mu^{(r)})$ to $(\mathbf{T}^2, \mu^{(2)})$ is clearly measure preserving and the image is the zero measure set $\mathbf{H}$ described in Lemma 5.1 below. Thus, $\mu^{(r)}(\mathbf{H}_{i,j}) = 0$.

We saw that the integrands in lines (8) and (9) agree outside the zero measure set $\bigcup_{i\neq j} \mathbf{H}_{i,j}$, so the two integrals agree.

Finally, notice that we integrate the exact same function on the same domain in lines (4) and (9). We use different measures but one of them is an

extension of the other. As both integrals are defined, they must agree. This shows that Equation (1) holds and finishes the proof of the theorem assuming Lemma 5.1 below. $\qquad\square$

Lemma 5.1. *Using the notation of Theorem 4 and its proof we have* $\mu^{(2)}(\mathbf{H}) = 0$ *for the set* $\mathbf{H} = \{(\mathbf{x}, \mathbf{y}) \in \mathbf{X}^2 \mid \mathbf{d}(\mathbf{x}, \mathbf{y}) \neq d_D(\alpha(\hat{\pi}(\mathbf{x})), \alpha(\hat{\pi}(\mathbf{y})))\}$.

Proof. The right hand side of the inequality defining $\mathbf{H}$ is μ^2-measurable. The left hand side is $\mathbf{d} = \lim_\omega d_n$, so it is $\mu^{(2)}$-measurable by Lemma 4.1. This makes $\mathbf{H}$ also $\mu^{(2)}$-measurable. Its slices are $\mathbf{H_x} = \{\mathbf{y} \in \mathbf{T} \mid (\mathbf{x}, \mathbf{y}) \in \mathbf{H}\}$. By Lemma 2.7 and since $\mathbf{d}(\mathbf{x}, \mathbf{y}) = \hat{\mathbf{d}}(\hat{\pi}(\mathbf{x}), \hat{\pi}(\mathbf{y}))$, we have $(\mathbf{x}, \mathbf{y}) \in \mathbf{H}$ if and only if $\hat{\pi}(\mathbf{x})$ and $\hat{\pi}(\mathbf{y})$ are in the same feather of $(\hat{\mathbf{X}}, \hat{\mathbf{d}}, \hat{\mu})$. This makes the slice $\mathbf{H_x}$ either empty (if $\mathbf{x} \in \mathbf{T} \setminus \mathbf{X}$ or $\hat{\pi}(\mathbf{x}) \in \mathrm{Core}(\hat{\mathbf{X}})$) or $\hat{\pi}^{-1}(\hat{\mathbf{F}})$ for a feather $\hat{\mathbf{F}}$. We have $\hat{\mu}(\hat{\mathbf{F}}) = 0$ by Lemma 2.7, so $\mu(\mathbf{H_x}) = 0$ in both cases. We have $\mu^{(2)}(\mathbf{H}) = \int_{\mathbf{T}} \mu(\mathbf{H_x}) \, d\mu(\mathbf{x}) = \int_{\mathbf{T}} 0 \, d\mu = 0$ by Lemma 4.2. $\qquad\square$

6. The proof of Theorem 3

We will use the following simple observations of compact subtrees exhausting some real trees.

Lemma 6.1. *Let* (T, d, μ) *be a semi-measured real tree with* $\mu(B) > 0$ *for all branches* B *of* T. *Then there exists a countable set* $Q \subseteq T$ *that is dense in any proper segment of* T. *For all* $\varepsilon > 0$ *there exists a compact subtree* Y *of* T *with* $\mu(Y) > 1 - \varepsilon$.

Proof. Any intermediate point p has more than one p-branches, each of positive measure, making p an inner point. The intermediate points are dense in T unless $|T| = 1$, so we have $\mathrm{Core}(T) = T$. This trivially holds even if $|T| = 1$. The separability comes from Lemma 2.4.

Let us fix $p_0 \in T$ and for a rational number $r > 0$ let Q_r be the set of inner points in T of distance r from p_0. Recall that the set $Q = \bigcup_r Q_r$ is a countable set that is dense in every proper segment of T by Lemma 2.4.

To find the compact subtree with large measure let us fix constants $\varepsilon_r > 0$ for the positive rational numbers r and $\varepsilon' > 0$ with $\varepsilon' + \sum_r \varepsilon_r < \varepsilon$. Let us find a closed ball A around p_0 with $\mu(A) > 1 - \varepsilon'$. Let us define B_p for $p \in Q$ to be the p-branch of T containing p_0 and $C_p = T \setminus (B_p \cup \{p\})$. Clearly, the open sets C_p for $p \in Q_r$ are pairwise disjoint, so we have $\sum_{p \in Q_r} \mu(C_p) \leq 1$. We find finite subsets $Q'_r \subseteq Q_r$ with $\sum_{p \in Q_r \setminus Q'_r} \mu(C_p) < \varepsilon_r$.

Consider the sets $C = \bigcup C_p$, where the union is taken for $p \in \bigcup_r (Q_r \setminus Q'_r)$ and let $Y = A \setminus C$. We have $\mu(C) < \sum_r \varepsilon_r$ and $\mu(Y) \geq \mu(A) - \mu(C) > 1 - \varepsilon$. As a closed and connected subset of T, Y is subtree. It remains to prove that Y is compact.

As a subtree Y is a complete metric space, so it is enough to show that for any $\varepsilon^* > 0$ and for any infinite sequence of points $y_i \in Y$ we can find $i \neq j$ with $d(y_i, y_j) < \varepsilon^*$. As the sequence $(d(p_0, y_i))_{i \in \mathbb{N}}$ is bounded we can find a short interval containing infinitely many of these distances, namely for some $a \geq 0$ the subsequence consisting of the points y_i satisfying $a \leq d(p_0, y_i) \leq a + \varepsilon^*/4$ is

infinite. If $a = 0$, then the distance of any two of the points in the subsequence is at most $\varepsilon^*/2$ and we are done. Otherwise, we find a positive rational number r with $a - \varepsilon^*/4 < r < a$. For a point y_i in the subsequence the unique point z_i in $[p_0, y_i] \cap Q_r$ satisfies $z_i \in Q_r'$ and $d(z_i, y_i) < \varepsilon^*/2$. As Q_r' is finite we can find two distinct elements of the subsequence y_i and y_j with $z_i = z_j$. But then $d(y_i, y_j) \leq d(y_i, z_i) + d(y_j, z_j) < \varepsilon^*$ as needed. $\qquad\square$

The next lemma extends the concept of compact exhaustion to the domains of long dendrons. The similar statement for dendrons is even more immediate as we can always choose $a = 1/2$.

Lemma 6.2. *For a long dendron $D = (T, d, \nu)$ the real tree (T, d) is separable, moreover it has a countable subset that is dense in every proper segment. For every $\varepsilon > 0$ we have a compact subtree Y of T and $a > 0$ such that $\nu(Y \times [0, a]) > 1 - \varepsilon$.*

Proof. Recall that ν is a Borel probability measure on $A_D = T \times [0, \infty)$, and consider its marginal ν_T on T. In other words, ν_T is the push-forward of ν along the projection of A_D to T. This (T, d, ν_T) a measured real tree with each branch having positive measure. We have a countable set in T that is dense in every proper segment by Lemma 6.1. The lemma further implies that there is a compact subtree Y of T such that $\nu_T(Y) > 1 - \varepsilon/2$. We have $\lim_{a \to \infty} \nu(Y \times [0, a]) = \nu(Y \times [0, \infty)) = \nu_T(Y) > 1 - \varepsilon/2$. Therefore $a > 0$ can be chosen as required. $\qquad\square$

Lemma 6.3. *The unique minimal subtree of a real tree T containing the points $x_1, \ldots, x_n \in T$ is $\bigcup_{i=1}^n [x_1, x_i]$.*

Proof. Any subset of T containing x_1 and x_i must also contain $[x_1, x_i]$ to be connected. But $\bigcup_{i=1}^n [x_1, x_i]$ is connected, closed and non-empty, so by Lemma 2.1(1) it is indeed a subtree. $\qquad\square$

Definition 6.1. Let $D = (T, d, \nu)$ be a long dendron and $r \geq 1$ integer. We call $x \in A_D^n$ an n-**sample** of D. For an n-sample $x = ((p_1, a_1), \ldots, (p_n, a_n))$, we define the measured real tree $T_x^D = (T', d', \mu)$ as follows. Let T_0 be the minimal subtree of T containing the points $p_1, \ldots, p_n$, whose existence is given by Lemma 6.3. We obtain T' from T_0 by appending the segments $[p_i, q_i]$ of length a_i to T_0 for $i = 1, \ldots, n$ such that T_0 and all the sets $[p_i, q_i] \setminus \{p_i\}$ are pairwise disjoint. We set d' to be the shortest path metric. This makes (T', d') a real tree. The measure μ is defined by $\mu(H) = |\{1 \leq i \leq n \mid q_i \in H\}|/n$ for any Borel subset $H \subseteq T'$. In other words, μ is the distribution of q_i with a uniform random i. We also define the map $\alpha_x^D : T' \to A_D$ by setting $\alpha_x^D(p) = (p, 0)$ for $p \in T_0$ and $\alpha_x^D(p) = (p_i, d'(p, p_i))$ if $p \in [p_i, q_i]$.

We call the sequence $(x^n)_{n \in \mathbb{N}}$ an **infinite sample** of D if x^n is an n-sample for all n. An infinite sample obtained by independently selecting an n-sample x^n for all n according the distribution ν^n is called an **infinite random sample** of D. We say that the infinite sample $(x^n)_{n \in \mathbb{N}}$ **obeys** a Borel set $H \subseteq A_D$ if $\nu(H) = \lim_{n \to \infty} |\{1 \leq i \leq n \mid x_i^n \in H\}|/n$, where $x^n = (x_1^n, \ldots, x_n^n)$.

Lemma 6.4. *For a long dendron D and every single Borel subset $H \subseteq A_D$ an infinite random sample of D almost surely obeys H.*

19

Proof. This is a form of the law of large numbers. $\qquad\square$

Our main result in this section is the following theorem. Theorem 3 will be a simple consequence.

Theorem 5. *Let $(x^n)_{n\in\mathbb{N}}$ be an infinite random sample of the long dendron $D = (T, d, \nu)$. The measured real trees $T_{x^n}^D$ almost surely form an essentially bounded sequence. Furthermore, the long dendron associated with their ultraproduct is almost surely isomorphic to D.*

The proof is through a series of lemmas. For this proof we fix $D = (T, d, \nu)$, and the infinite sample $(x^n)_{n\in\mathbb{N}}$ of D with $x^n = (x_1^n, \ldots, x_n^n)$. We say that a Borel set in A_D is **obeyed** if the sequence $(x^n)_{n\in\mathbb{N}}$ obeys it. Throughout the proof we will assume that various Borel sets in A_D are obeyed. We can do that as long as we make the assumption for a countable family of sets by Lemma 6.4.

We introduce some notation. Let $T_{x^n}^D = (T_n, d_n, \mu_n)$ and let $\alpha_n = \alpha_{x^n}^D$. Let $(\hat{\mathbf{T}}, \hat{\mathbf{d}})$ be the metric ultraproduct of (T_n, d_n) and let $\hat{\mu}$ be the push-forward measure of $\prod_\omega \mu_n$ along the natural projection $\hat{\pi} : \mathbf{T} \to \hat{\mathbf{T}}$ from the set theoretic ultraproduct $\mathbf{T} = \prod_\omega T_n$ to the metric ultraproduct $\hat{\mathbf{T}}$.

Let $\alpha = \lim_\omega \alpha_n$. Here $\alpha_n : T_n \to A_D$ and A_D (considered always with the product topology) is not compact, so α is defined on a subset of $\mathbf{T}$, namely for $[(x_n)_{n\in\mathbb{N}}] \in \mathbf{T}$, where $\lim_\omega \alpha_n(x_n)$ exists.

Consider the distance function d' on A_D defined as

$$(10) \qquad d'((u, a), (v, b)) = \begin{cases} |a - b| & \text{if } u = v \\ d(u, v) + a + b & \text{otherwise.} \end{cases}$$

Note that this is indeed a distance function making (A_D, d') a real tree as mentioned in Remark 5.

For any n and $p, q \in T_n$ we have

$$d'(\alpha_n(p), \alpha_n(q)) \leq d_n(p, q) \leq d_D(\alpha_n(p), \alpha_n(q)).$$

Note that d_D is continuous on A_D^2 and d' is lower semicontinuous there. Thus, for points $\mathbf{p}$ and $\mathbf{q}$ in the domain of α we have

$$d'(\alpha(\mathbf{p}), \alpha(\mathbf{q})) \leq \mathbf{d}(\mathbf{p}, \mathbf{q}) \leq d_D(\alpha(\mathbf{p}), \alpha(\mathbf{q})),$$

where $\mathbf{d} = \lim_\omega d_n$. This means, in particular, that $\alpha(\mathbf{p}) = \alpha(\mathbf{q})$ if $\mathbf{d}(\mathbf{p}, \mathbf{q}) = 0$. One can also see that if $\mathbf{d}(\mathbf{p}, \mathbf{q}) = 0$ and one of $\alpha(\mathbf{p})$ and $\alpha(\mathbf{q})$ is defined then so is the other. Therefore, we can define the function $\hat{\alpha}$ by setting $\hat{\alpha}(\hat{\pi}(\mathbf{x})) = \alpha(\mathbf{x})$ if $\alpha(\mathbf{x})$ is defined and keeping $\hat{\alpha}(\hat{\pi}(\mathbf{x}))$ undefined if $\alpha(\mathbf{x})$ is not defined. Let $\hat{\mathbf{D}}$ stand for the domain of $\hat{\alpha}$. For $\hat{\mathbf{p}}, \hat{\mathbf{q}} \in \hat{\mathbf{D}}$ we have

$$(11) \qquad d'(\hat{\alpha}(\hat{\mathbf{p}}), \hat{\alpha}(\hat{\mathbf{q}})) \leq \hat{\mathbf{d}}(\hat{\mathbf{p}}, \hat{\mathbf{q}}) \leq d_D(\hat{\alpha}(\hat{\mathbf{p}}), \hat{\alpha}(\hat{\mathbf{q}})).$$

We call a point $\hat{\mathbf{p}} \in \hat{\mathbf{D}}$ a **base point** for $p \in T$ if $\hat{\alpha}(\hat{\mathbf{p}}) = (p, 0)$. The **base** of $\hat{\mathbf{T}}$ is the set of base points in $\hat{\mathbf{D}}$.

Let us fix a countable set $Q \subseteq T$ that is dense in every proper segment of T. We can do this by Lemma 6.2. In the sequel we assume that $B \times [0, \infty)$ is obeyed if B is a q-branch of T for some $q \in Q$. Note that T is separable and

Q is countable, therefore we are making this assumption about a countable family of sets.

Lemma 6.5. *There is a unique base point for every $p \in T$. The function $\beta : T \to \hat{\mathbf{T}}$ mapping $p \in T$ to the base point for p is an isometry from T to the base of $\hat{\mathbf{T}}$.*

Proof. Notice that $d'((p,0),(q,0)) = d_D((p,0),(q,0)) = d(p,q)$ for $p,q \in T$, so if $\hat{\mathbf{p}}$ is a base point for $p \in T$ and $\hat{\mathbf{q}}$ is a base point for $q \in T$, then

$$\hat{\mathbf{d}}(\hat{\mathbf{p}}, \hat{\mathbf{q}}) = d(p,q)$$

by our bound (11). In particular, this means that the base point for $p \in T$ is unique if exists and if they exist for all $p \in T$, then β is an isometry as claimed. It remains to show the existence.

Notice that $T \cap T_n$ is a subtree of T. Let p be an arbitrary point of T and let $\hat{\mathbf{p}} = [\widehat{(p_n)_{n\in\mathbb{N}}}] \in \hat{\mathbf{T}}$ with $p_n = \pi^{T,d}_{T \cap T_n}(p)$. Assume $\hat{\mathbf{p}}$ is not a base point for p. We have $\alpha_n(p_n) = (p_n, 0)$, so this means that there is $\varepsilon > 0$ such that $d(p, p_n) < \varepsilon$ for ω-few indices n. In this case we can take a segment $[p, p']$ of length ε in T, an intermediate point $q \in Q$ in that segment and notice that $q \in T_n$ for ω-few indices n. As q is an intermediate point, there are at least two distinct q-branches B and B'. The sets $B \times [0, \infty)$ and $B' \times [0, \infty)$ are both obeyed and have positive ν-measure. Therefore either is avoided by the sample x^n for finitely many indices n. If neither is avoided, then $q \in T_n$. So $q \in T_n$ for ω-few indices n, but $q \notin T_n$ for finitely many indices n. The contradiction shows that $\hat{\mathbf{p}}$ is a base point for p. $\qquad\square$

Let us choose compact subtrees Y_s of T and reals $a_s > 0$ for $s \in \mathbb{N}$ and form the compact product spaces $Z_s = Y_s \times [0, a_s] \subseteq A_D$. We can do this such that $\nu(Z_s) > 1 - 1/s$ by Lemma 6.2. From now on we assume that the sets Z_s are obeyed.

For $s, n \geq 1$ let $T_{s,n} = \alpha_n^{-1}(Z_s)$. This is a subtree of T_n whenever non-empty. Let $\hat{\mathbf{T}}_s = \hat{\prod_\omega} T_{s,n} \subseteq \hat{\mathbf{T}}$.

Lemma 6.6. *The sequence of metric measure spaces $(T^D_{x^n})_{n\in\mathbb{N}}$ is essentially bounded with $\hat{\mathbf{D}}$ contained in the dominant cluster of $\hat{\mathbf{T}}$.*

The sets $\hat{\mathbf{T}}_s$ are real trees with $\hat{\mu}(\hat{\mathbf{T}}_s) \geq 1 - 1/s$. We have $\hat{\mathbf{T}}_s \subseteq \hat{\mathbf{D}}$ and therefore $\hat{\mu}(\hat{\mathbf{D}}) = 1$.

Proof. We have $\text{diam}(T_{s,n}) \leq \text{diam}(Y_s) + 2a_s$. This makes the sequence of measured real trees $(T_{s,n}, d_n)_{n\in\mathbb{N}}$ uniformly bounded and thus their metric ultraproduct $\hat{\mathbf{T}}_s$ is a real tree or empty by Lemma 3.1. We have $\hat{\mu}(\hat{\mathbf{T}}_s) \geq \lim_\omega \mu_n(T_{s,n})$. Here $\mu_n(T_{s,n}) = |\{1 \leq i \leq n \mid x_i^n \in Z_s\}/n$, so as Z_s is obeyed, we have $\lim_{n\to\infty} \mu_n(T_{s,n}) = \nu(Z_s) > 1 - 1/s$. Therefore $\hat{\mu}(\hat{\mathbf{T}}_s) > 1 - 1/s$. In particular, $\hat{\mathbf{T}}_s$ is not empty, so it is a real tree.

We have $\hat{\mathbf{T}}_s \subseteq \hat{\mathbf{D}}$ as for $\hat{\mathbf{p}} = [\widehat{(p_n)_{n\in\mathbb{N}}}]$ with $p_n \in T_{s,n}$ the ultralimit $\hat{\alpha}(\hat{\mathbf{p}}) = \lim_\omega \alpha_n(p_n)$ exists because $\alpha_n(p_n) \in Z_s$ and Z_s is compact. Thus, $\hat{\mu}(\hat{\mathbf{D}}) = 1$ as claimed.

We have $\hat{\mathbf{d}}(\hat{\mathbf{p}}, \hat{\mathbf{q}}) < \infty$ for $\hat{\mathbf{p}}, \hat{\mathbf{q}} \in \hat{\mathbf{D}}$ by the bound (11), therefore $\hat{\mathbf{D}}$ is contained in a single cluster of $\hat{\mathbf{T}}$. As $\hat{\mu}(\hat{\mathbf{D}}) = 1$, this cluster must be dominant and the sequence $(T_{x^n}^D)_{n \in \mathbb{N}}$ is essentially bounded as claimed. $\qquad \square$

We denote the dominant cluster of $\hat{\mathbf{T}}$ by $\hat{\mathbf{X}}$. We slightly abuse notation by also writing $\hat{\mathbf{X}}$ when referring to the ultraproduct of the measured real trees $T_{x^n}^D$, that is, to $\hat{\mathbf{X}}$ together with the restrictions of $\hat{\mathbf{d}}$ and $\hat{\mu}$ making it a semi-measured real tree.

Lemma 6.7. *The core of $\hat{\mathbf{X}}$ is the base of $\hat{\mathbf{T}}$.*

Proof. First we claim that every inner point $\hat{\mathbf{x}}$ of $\hat{\mathbf{X}}$ belongs to $\hat{\mathbf{D}}$. Indeed, for $\hat{\mathbf{x}} \in \hat{\mathbf{X}} \setminus \hat{\mathbf{D}}$ the real tree $\hat{\mathbf{T}}_s$ does not contain $\hat{\mathbf{x}}$, so it must be contained in a single $\hat{\mathbf{x}}$-branch. This branch has measure over $1 - 1/s$. As such an $\hat{\mathbf{x}}$-branch exists for each s, $\hat{\mathbf{x}}$ is not an inner point.

Consider a point $\hat{\mathbf{x}} = [\widehat{(x_n)_{n \in \mathbb{N}}}] \in \hat{\mathbf{D}}$ with $\hat{\alpha}(\hat{\mathbf{x}}) = (z, a)$. Let $\alpha_n(x_n) = (z_n, a_n)$. We define a subtree T'_n of T_n for $n \in \mathbb{N}$ as follows. Recall that T_n is obtained from a subtree of T by appending n segments. If $x_n \notin T$ we obtain T'_n from T_n by removing the added segment that contains x_n. If $x_n \in T$ (that is, $a_n = 0$) we simply set $T'_n = T_n$. The distance between x_n and T'_n is a_n. Let $\hat{\mathbf{T}}' = \hat{\mathbf{X}} \cap \prod_\omega T'_n$. This is a subtree by Lemma 3.1. The distance between $\hat{\mathbf{x}}$ and $\hat{\mathbf{T}}'$ is $\lim_\omega a_n = a$, so if $a > 0$, then $\hat{\mathbf{x}} \notin \hat{\mathbf{T}}'$ and $\hat{\mathbf{T}}'$ must lie inside a single $\hat{\mathbf{x}}$-branch. As $\mu_n(T'_n) \geq 1 - 1/n$, we have $\hat{\mu}(\hat{\mathbf{T}}') = 1$, so this $\hat{\mathbf{x}}$-branch has full measure and therefore $\hat{\mathbf{x}}$ is not an inner point if $a > 0$.

So far we proved that for an inner point $\hat{\mathbf{x}}$ of $\hat{\mathbf{X}}$, $\hat{\alpha}(\hat{\mathbf{x}})$ must be defined and $\hat{\alpha}(\hat{\mathbf{x}}) = (z, a)$ with $a > 0$ is not possible. Therefore, every inner point of $\hat{\mathbf{X}}$ must be in the base.

Let $q \in Q$ be an intermediate point in T. Let B and B' be two distinct q-branches of T. Let $C = B \times [0, \infty)$ and $C' = B' \times [0, \infty)$. These positive measure subsets of A_D are both obeyed. Therefore the sets $\hat{\mathbf{C}} = \prod_\omega C_n$ and $\hat{\mathbf{C}}' = \prod_\omega C'_n$ are also positive measure subsets, where $C_n = \alpha_n^{-1}(C)$ and $C'_n = \alpha_n^{-1}(C')$. Consider points $\hat{\mathbf{p}} = [\widehat{(p_n)_{n \in \mathbb{N}}}] \in \hat{\mathbf{C}} \cap \hat{\mathbf{X}}$ with $p_n \in C_n$ and $\hat{\mathbf{p}}' = [\widehat{(p'_n)_{n \in \mathbb{N}}}] \in \hat{\mathbf{C}}' \cap \hat{\mathbf{X}}$ with $p'_n \in C'_n$. We have $q \in [p_n, p'_n]$. Note that p_n or p'_n might be undefined for ω-few indices n where C_n or C'_n is empty, but otherwise $q \in T_n$, so we can set $\hat{\mathbf{q}} = [\widehat{(q_n)_{n \in \mathbb{N}}}]$ with $q_n = q$ if $q \in T_n$ and q_n undefined otherwise. Clearly, $\hat{\mathbf{q}}$ is the base point for q. By Lemma 3.1 we have $[\hat{\mathbf{p}}, \hat{\mathbf{p}}'] = \prod_\omega [p_n, p'_n]$, so $\hat{\mathbf{q}} \in [\hat{\mathbf{p}}, \hat{\mathbf{p}}']$. This implies that $\hat{\mathbf{q}}$ is an inner point in $\hat{\mathbf{X}}$. Indeed, otherwise there would be a full measure $\hat{\mathbf{q}}$-branch $\hat{\mathbf{B}}$ and the positive measure sets $\hat{\mathbf{C}}$ and $\hat{\mathbf{C}}'$ would both intersect $\hat{\mathbf{B}}$, but the segment between any two points of $\hat{\mathbf{B}}$ does not contain $\hat{\mathbf{q}}$.

Assume $|T| > 1$. Then the intermediate points in Q are dense in T, so by the isometry, the base points for these, all inner points, are dense in the base. This means that $\text{Core}(\hat{\mathbf{X}})$ contains the base. But we also saw that all inner points are in the base, and the base is complete, therefore closed, so we must have that $\text{Core}(\hat{\mathbf{X}})$ is the base as claimed.

The same conclusion is even easier to obtain if $|T| = 1$. Indeed, we know that the core is not empty (Lemma 2.2) and there is no inner point outside singleton base, so the core of $\hat{\mathbf{X}}$ must again be the base. $\qquad\square$

Let $D^* = (T^*, d^*, \nu^*)$ be the long dendron associated with $\hat{\mathbf{X}}$ and let $\alpha^* : \hat{\mathbf{X}} \to A_{D^*}$ be the associated projection. Recall that β is the function that maps a point in T to the unique base point for it in $\hat{\mathbf{T}}$.

Lemma 6.8. *Let $\hat{\mathbf{x}} \in \hat{\mathbf{D}}$ be a point with $\hat{\alpha}(\hat{\mathbf{x}}) = (p, a) \in A_D$. Then $\alpha^*(\hat{\mathbf{x}}) = (\beta(p), a) \in A_{D^*}$.*

Proof. Let $\hat{\mathbf{q}} \in \mathrm{Core}(\hat{\mathbf{X}})$, so by Lemma 6.7 $\hat{\mathbf{q}}$ is the base point for some $q \in T$. We have
$$d'(\hat{\alpha}(\hat{\mathbf{x}}), \hat{\alpha}(\hat{\mathbf{q}})) \leq \hat{\mathbf{d}}(\hat{\mathbf{x}}, \hat{\mathbf{q}}) \leq d_D(\hat{\alpha}(\hat{\mathbf{x}}), \hat{\alpha}(\hat{\mathbf{q}}))$$
by bound (11). Here $\hat{\alpha}(\hat{\mathbf{x}}) = (p, a)$, $\hat{\alpha}(\hat{\mathbf{q}}) = (q, 0)$ and $d'((p, a), (q, 0)) = d_D((p, a), (q, 0)) = d(p, q) + a$. So we have $\hat{\mathbf{d}}(\hat{\mathbf{x}}, \hat{\mathbf{q}}) = d(p, q) + a$. This implies that $\pi^{\hat{\mathbf{X}}}(\hat{\mathbf{x}})$, the unique closest point to $\hat{\mathbf{x}}$ in $\mathrm{Core}(\hat{\mathbf{X}})$ is $\beta(p)$ with $\hat{\mathbf{d}}(\hat{\mathbf{x}}, \beta(p)) = a$. This proves the lemma. $\qquad\square$

Remark 7. This last lemma gives a nice description of α^* on the domain $\hat{\mathbf{D}}$ of $\hat{\alpha}$. It is somewhat harder to describe α^* on the zero measure set $\hat{\mathbf{X}} \setminus \hat{\mathbf{D}}$. For example for a point $\hat{\mathbf{p}} = [\widehat{(p_n)_{n \in \mathbb{N}}}] \in \hat{\mathbf{X}} \setminus \hat{\mathbf{D}}$ the second coordinate of $\alpha^*(\hat{\mathbf{p}})$ is strictly larger than the ultralimit of the second coordinates of $\alpha_n(p_n)$. Fortunately, knowing α^* almost everywhere in $\hat{\mathbf{X}}$ is enough to determine ν^*.

Proof of Theorem 5. Let $(x^n)_{n \in \mathbb{N}}$ be an infinite random sample of the long dendron D. We will use the notation introduced above. By Lemma 6.4 our assumption on various sets being obeyed holds almost surely, so we can use Lemmas 6.5–6.8 above.

In particular, by Lemma 6.6 the sequence of measured real trees $(T^D_{x^n})_{n \in \mathbb{N}}$ is essentially bounded, so its ultraproduct $\hat{\mathbf{X}}$ exists. For the long dendron $D^* = (T^*, d^*, \nu^*)$ associated with $\hat{\mathbf{X}}$ we have $T^* = \mathrm{Core}(\hat{\mathbf{X}})$, so by Lemmas 6.5 and 6.7 the function $\beta : T \to T^*$ is an isometry. It remains to prove that the function $\beta' : A_D \to A_{D^*}$ defined by $\beta'(p, a) = (\beta(p), a)$ is almost surely measure preserving from (A_D, ν) to (A_{D^*}, ν^*). That is, we need to show that almost surely
$$\nu(\beta'^{-1}(H)) = \nu^*(H)$$
holds for all Borel sets $H \subseteq A_{D^*}$. Here ν^* is defined as the push-forward of $\hat{\mu}$ along α^*. By Lemmas 6.8 and 6.6 we have $\alpha^*(\hat{\mathbf{x}}) = \beta'(\hat{\alpha}(\hat{\mathbf{x}}))$ almost everywhere in $\hat{\mathbf{X}}$. So we have
$$\nu^*(H) = \hat{\mu}((\alpha^*)^{-1}(H)) = \hat{\mu}(\hat{\alpha}^{-1}(\beta'^{-1}(H))).$$
Therefore, it is enough to show that almost surely

(12) $$\nu(H) = \hat{\mu}(\hat{\alpha}^{-1}(H))$$

holds for all Borel sets $H \subseteq A_D$. Both sides of Equation 12 are Borel probability measures on A_D. It is therefore enough to check the equation for an

23

intersection-closed family $\mathcal{A}$ of Borel sets that generate the entire Borel σ-algebra on A_D. We choose such a countable family $\mathcal{A}$ consisting of closed sets. For example if T has more than a single point, then the finite intersections of sets of the form $(T \setminus B) \times [0, a]$ will work if B runs through q-branches for $q \in Q$ and $a \geq 0$ is a rational number. Recall, that Q is a countable set in the separable real tree T, so this is a countable family. Therefore, it is enough to check that Equation 12 holds almost surely for any single closed set $H \subseteq A_D$.

First we claim that the inequality

$$(13) \qquad \nu(H) \leq \hat{\mu}(\hat{\alpha}^{-1}(H))$$

holds for any closed set $H \subseteq A_D$ that $(x^n)_{n \in \mathbb{N}}$ obeys. Let us fix such a set H. Consider the sets $H_n = \alpha_n^{-1}(H)$ for $n \in \mathbb{N}$. We have $\mu_n(H_n) = |\{1 \leq i \leq n \mid x_i^n \in H\}|/n$ and $\lim_{n \to \infty} \mu_n(H_n) = \nu(H)$ as H is obeyed. Consider $\hat{\mathbf{H}} = \hat{\prod}_\omega H_n$. We have $\hat{\mu}(\hat{\mathbf{H}}) \geq \lim_\omega \mu_n(H_n) = \nu(H)$. Any point $\hat{\mathbf{p}} \in \hat{\mathbf{H}} \cap \hat{\mathbf{D}}$ can be written as $\hat{\mathbf{p}} = [\widehat{(p_n)_{n \in \mathbb{N}}}]$ with $\alpha_n(p_n) \in H$ and we have $\hat{\alpha}(\hat{\mathbf{p}}) = \lim_\omega \alpha_n(p_n)$. An ultralimit of points in H is also in H as H is closed. So we have $\hat{\mathbf{H}} \cap \hat{\mathbf{D}} \subseteq \hat{\alpha}^{-1}(H)$. Using Lemma 6.6, we have $\nu(H) \leq \hat{\mu}(\hat{\mathbf{H}}) = \hat{\mu}(\hat{\mathbf{H}} \cap \hat{\mathbf{D}}) \leq \hat{\mu}(\hat{\alpha}^{-1}(H))$ as claimed.

Let us now fix a closed set $H \subseteq A_D$. Our goal is to prove that Equation 12 holds almost surely. Let us choose an increasing sequence of closed sets $H_1' \subseteq H_2' \subseteq \cdots$ with $\bigcup_{s=1}^\infty H_s' = A_D \setminus H$. One can, for example, take H_s' to be the complement of the open $1/s$-neighborhood of H in some metrization of A_D. We have $\lim_{s \to \infty} \nu(H_s') = \nu(A_D \setminus H) = 1 - \nu(H)$ and similarly, using that $\hat{\alpha}$ is defined almost everywhere (Lemma 6.6), we have $\lim_{s \to \infty} \hat{\mu}(\hat{\alpha}^{-1}(H_s')) = \hat{\mu}(\hat{\alpha}^{-1}(A_D \setminus H)) = 1 - \hat{\mu}(\hat{\alpha}^{-1}(H))$.

By lemma 6.4 H and all the sets H_s' are almost surely obeyed. If so, we have $\nu(H) \leq \hat{\mu}(\hat{\alpha}^{-1}(H))$ and also $\nu(H_s') \leq \hat{\mu}(\hat{\alpha}^{-1}(H_s'))$ for all s, all from Inequality (13). Taking limit on both sides of the latter inequality we get $1 - \nu(H) \leq 1 - \hat{\mu}(\hat{\alpha}^{-1}(H))$. This makes Equation (12) hold almost surely and finishes the proof of the theorem. □

Definition 6.2. Let $n \geq 1$ and let $A \in M_n$ a n by n real matrix. The measured real tree (T, d, μ) is an A-**tree** if there exist $q = (q_1, \ldots, q_n) \in T^n$ such that (i) $\rho_n^{T,d}(q) = A$, (ii) no subtree of T other than T itself contains all of the points q_i and (iii) μ is a Borel measure on T given by $\mu(H) = |\{1 \leq i \leq n \mid q_i \in H\}|/n$.

Lemma 6.9. *Let $n \geq 1$ and $A \in M_n$. If there exists an A-tree it is unique up to measure preserving isometry. If $n \geq 2$ and x is an n-sample of a long dendron D, then T_x^D is a $\rho_n^{A_D, d_D}(x)$-tree.*

Proof. Assume both (T, d, μ) and (T', d', μ') are A-trees for $A = (d_{i,j})_{1 \leq i,j \leq n}$. Let $q = (q_1, \ldots, q_n) \in T^n$ and $q' = (q_1', \ldots, q_n') \in T'^n$ be the samples that satisfy conditions (i-iii) in the definition. By condition (ii) and Lemma 6.3 we have $T = \bigcup_{i=1}^n [q_1, q_i]$ and $T' = \bigcup_{i=1}^n [q_1', q_n']$. Both the segments $[q_1, q_i]$ and $[q_1', q_i']$ have length $d_{1,i}$ by condition (i), so there is an isometry f_i between

them with $f_i(q_1) = q'_1$, $f_i(q_i) = q'_i$. The intersection of the segments $[q_1, q_i]$ and $[q_1, q_j]$ is a segment starting at q_1, and by condition (i) its length is $(d_{1,i} + d_{1,j} - d_{i,j})/2$. Similarly, $[q'_1, q'_i] \cap [q'_1, q'_j]$ is a segment of the same length starting at q'_1. Therefore, the functions f_i agree on the intersections of their domains, so there is a global function $f : T \to T'$ extending all these functions. It is easy to see that f is an isometry between T and T'. By condition (iii) and since $f(q_i) = q'_i$ for all i, the map f is measure preserving.

Let us recall, that for a long dendron $D = (T, d, \nu)$ and a n-sample $x = ((p_1, a_1), \ldots, (p_n, a_n))$ of D we constructed the measured real tree T_x^D by adding segments $[p_i, q_i]$ to the minimal subtree of T containing the points p_i. The measured real tree T_x^D with $q = (q_1, \ldots, q_n)$ satisfy conditions (i) and (iii) for a $\rho_n^{A_D, d_D}(x)$-tree. They also satisfy condition (ii) as we assumed $n \geq 2$. Note that condition (ii) might be violated for $n = 1$ as in that case T_x^D is a segment of length a_1 and (if $a_1 > 0$) not the trivial subtree $\{q_1\}$. □

Proof of Theorem 3. Using the observations in Lemma 6.9 we can restate Theorem 5 as follows. For a long dendron D take independent samples A_n from the distribution $\tau_n(D)$ and create the corresponding A_n-trees T_n. These exist and are unique up to measure preserving isometry. Almost surely, the sequence $(T_n)_{n \in \mathbb{N}}$ is essentially bounded and the long dendron associated with its ultraproduct is isomorphic to D.

If the two long dendrons D and D' satisfy $\tau_n(D) = \tau_n(D')$ for all n, then the above process is the same for both of them. Almost surely, the process results in an associated long dendron that is isomorphic to both D and D'. Thus, D and D' are isomorphic as claimed. □

7. The proof of Theorem 2

Definition 7.1. Let T be a finite (graph-theoretic) tree. We turn it into a real tree by keeping the vertices and replacing every edge by an arbitrary length interval connecting the corresponding vertices. Finally, we make it into a measured real tree by adding an arbitrary Borel measure concentrated on the original set of vertices. We call the measured real trees obtained this way **finite real trees**. Observe that a measured real tree (T, d, μ) is a finite real tree if and only if (i) there is a finite set of points in T with no subtree containing all of them other than T and (ii) μ is concentrated on a finite set of points. In particular, if x is an n-sample of a long dendron D, then T_x^D is a finite real tree.

Theorem 2 easily follows from the following two lemmas as the weak topology of $\mathrm{Prob}(M_r)$ is metrizable for every $r \in \mathbb{N}$, by Prokhorov's Theorem. Note that Lemma 7.1 has a direct proof using the Azuma Inequality (see 5.3 [10]), but at this point it is easier for us to use Theorems 4 and 5 instead.

Lemma 7.1. *Any long dendron D is the sampling limit of finite real trees, that is, one can find finite real trees (T_n, d_n, μ_n) such that for every $r \geq 1$ the sampling measures $\tau_r(T_n, d_n, \mu_n)$ weakly converge to $\tau_r(D)$.*

If D is a dendron, then the finite real trees can be chosen such that the diameter of each is at most one.

Proof. Let us choose an infinite random sample $(x^n)_{n\in\mathbb{N}}$ of D. By Theorem 5, D can almost surely be obtained as the long dendron associated with the ultraproduct of the essentially bounded sequence of finite real trees $T^D_{x^n}$. In this case we have $\lim_\omega \tau_r(T^D_{x^n}) = \tau_r(D)$ for any $r \in \mathbb{N}$ by Theorem 4. The ultralimit is always the (ordinary) limit of a subsequence in a metrizable space, so we have a sequence of finite real trees (T_n, d_n, μ_n) satisfying that $\tau_r(T_n, d_n, \mu_n)$ weakly converges to $\tau_r(D)$. We can clearly find the same sequence for all r.

If D is a dendron, then $\operatorname{diam}(T^D_{x^n}) \leq 1$ holds almost surely, so the second statement of the lemma also holds. $\square$

Lemma 7.2. *For a finite real tree (T, d, μ) of diameter at most 1 one has a sequence of finite (graph-theoretic) trees T_n such that $\tau_r(T_n)$ weakly converges to $\tau_r(T, d, \mu)$ for every $r \in \mathbb{N}$.*

Proof. Let (T, d) be obtained from the finite tree (V, E) by replacing each edge $e \in E$ with a segment of length a_e. Note here that we have $a_e \leq 1$. For $n \in \mathbb{N}$ we build T_n in three steps. In the first step we replace each edge e of the tree (V, E) with a path of length $\lceil a_e n \rceil$. Then for every vertex $v \in V$ we form a set H^v_n of $\lceil \mu(v) n^2 \rceil$ new vertices and add them to the tree as leaves whose only neighbor is v. Finally, if the diameter of the tree constructed so far is below n, we attach a path to it that brings the diameter to exactly n.

The diameter of the tree T_n so constructed is $n + O(1)$. The distance between any vertex $x \in H^v_n$ and $y \in H^w_n$, is $d(v, w)n + O(1)$, so the normalized distance $d_n(x, y)$ is $d(v, w) + O(1/n)$. T_n has $n^2 + O(n)$ vertices, so the probability that a uniformly random vertex falls in H^v_n is $\mu(v) + O(1/n)$ with the probability of choosing a vertex outside all the sets H^v_n being $O(1/n)$. The hidden constants in the order notation depends only on the size of the vertex set V.

This shows that a μ-random point v can be coupled with a uniform random vertex of T such that a point $v \in V$ is coupled with a vertex $x \in H^v_n$ with probability $1 - O(1/n)$. Doing this in every coordinate one can couple a sample $(v_1, \ldots, v_r)$ of μ^r with a uniform sample $(x_1, \ldots, x_r)$ of $(V(T_n))^r$ such that $x_i \in H^{v_i}_n$ for all $1 \leq i \leq r$ with probability $1 - O(r/n)$. If this happens, then the matrices $\rho^{T,d}_r(v_1, \ldots, v_r)$ and $\rho^{T_n,d_n}_r(x_1, \ldots, x_n)$ differ by $O(1/n)$ in every coordinate. Therefore, their distributions $\tau_r(T, d, \mu)$ and $\tau_r(T_n)$ are close. In particular, this implies the weak convergence stated in the lemma. $\square$

REFERENCES

[1] M. BESTVINA, $\mathbb{R}$-trees in topology, geometry, and group theory. *Handbook of geometric topology*, North-Holland, Amsterdam, (2002) 55–91,

[2] C. BORGS, J. CHAYES AND L. LOVÁSZ, Moments of two-variable functions and the uniqueness of graph limits. *Geom. Funct. Anal.* **19** (2010), no. 6, 1597–1619.

[3] A. DEPPERSCHMIDT, A. GREVEN AND P. PFAFFELHUBER, Marked metric measure spaces. *Elec. Comm. Probab.* **16** (2011), 174–188.

[4] G. ELEK AND B. SZEGEDY, A measure-theoretic approach to the theory of dense hypergraphs. *Adv. Math.* **231** (2012), no. 3-4, 1731–1772.

[5] A. GREVEN, P. PFAFFELHUBER AND A. WINTER, Convergence in distribution of random metric measure spaces (Λ-coalescent measure trees). *Prob. Theo. Rel. Fields* **145** (2009), no. 1-2, 285-322.

[6] M. GROMOV, Metric structures for Riemannian and non-Riemannian spaces. With appendices by M. Katz, P. Pansu and S. Semmes. *Progress in Math.* **152** Birkhauser, Boston MA (2001)

[7] S. JANSON, Tree limits and limits of random trees. (to appear in *Combinatorics, Probability and Computing* doi:10.1017/S0963548321000055)

[8] P. A. LOEB, A nonstandard functional approach to Fubini's theorem. *Proc. Amer. Math. Soc.* bf 93 (1985), no. 2, 343–346.

[9] W. LÕHR, Equivalence of Gromov–Prohorov– and Gromov's $\square_\lambda$-metric on the space of metric measure spaces. *Electron. Commun. Probab.* **18** (2013), no. 17.

[10] L. LOVÁSZ AND B. SZEGEDY, Limits of dense graph sequences. *J. Combin. Theory Ser. B* **96** (2006), no. 6, 933–957.

[11] L. VAN DEN DRIES AND A. J. WILKIE, Gromov's theorem on groups of polynomial growth and elementary logic. *J. Algebra* **89** (1984), no. 2, 349–374.

[12] E. WARNER, Ultraproducts and the foundations of higher order Fourier analysis, *Ph.D. Thesis* (2012)