ALGEBRAIC TWO-LEVEL MEASURE TREES

JOSUÉ NUSSBAUMER, VIET CHI TRAN, AND ANITA WINTER

ABSTRACT. With the algebraic trees, Löhr and Winter (2021) introduced a generalization of the notion of graph-theoretic trees to account for potentially uncountable structures. The tree structure is given by the map which assigns to each triple of points their branch point. No edge length or distance is considered. One can equip a tree with a natural topology and a probability measure on the Borel-σ-field, defining in this way an algebraic measure tree. The main result of Löhr and Winter is to provide with the sample shape convergence a compact topology on the space of binary algebraic measure trees. This was proved by encoding the latter with triangulations of the circle. In the present paper, we extend this result to a two level setup. Motivated by the study of hierarchical systems with two levels in biology, such as host-parasite populations, we equip algebraic trees with a probability measure on the set of probability measures. To show the compactness of the space of binary algebraic two-level measure trees, we enrich the encoding of these trees by triangulations of the circle, by adding a two-level measure on the circle line. As an application, we define the two-level algebraic Kingman tree, that is the random algebraic two-level measure tree obtained from the nested Kingman coalescent.

1. INTRODUCTION

Because of their applications to biology and computer science, trees have received significant interest in the mathematical literature of the last decades. In probability theory in particular, many random tree structures have been introduced to model genealogical (or phylogenetic) trees and their evolutions over time. The simplest of these models are defined on state spaces of trees with a finite number of vertices. The size of the space of all trees with a given number of vertices grows exponentially, and it becomes hard to study qualitative statistics of the structures. To overcome this issue, it seems reasonable to consider continuum limits of tree models and study their properties. We are thus interested in a setup which unifies discrete and continuum trees, and where the focus is put on the structure of the trees. Following Löhr and Winter [LW21], we will consider here algebraic trees with a particular attention to two-level (possibly uncountable) trees. The latter, and more generally trees with nested or hierarchical structures, can serve for the modelling of host-parasite or individual-species systems. The introduction motivates the consideration of algebraic two-level measure trees. We will prove in this paper that the subspace of binary algebraic two-level measure trees can be embedded with a compact topology, allowing for limit theorems. We use this to construct the Kingman algebraic two-level measure tree from finite nested Kingman coalescent trees.
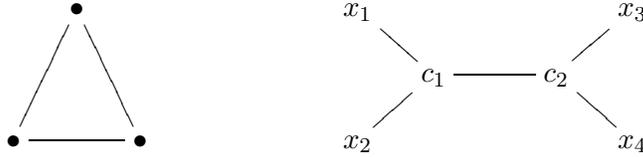
FIGURE 1. *On the left:* the discrete triangle equipped with the graph distance satisfies the four-point condition (1.1), but does not admit branch points. *On the right:* The only possible tree shape spanned by four points separates them into two pairs. Here, $c_1 = c(x_1, x_2, x_3) = c(x_1, x_2, x_4)$ and $c_2 = c(x_1, x_3, x_4) = c(x_2, x_3, x_4)$.

The common approach by now is to encode trees as metric spaces. For a metric space (T, r) to have a tree-like structure, it is necessary to satisfy the so-called *four-point condition* (see Figure 1): for all $x_1, x_2, x_3, x_4 \in T$,

$$(1.1) \qquad r(x_1, x_2) + r(x_3, x_4) \leq \max \left\{ r(x_1, x_3) + r(x_2, x_4), r(x_1, x_4) + r(x_2, x_3) \right\}.$$

However, not all metric spaces satisfying the four-point condition are tree-like, as for example the discrete triangle 3-graph (see Figure 1, left). For this reason, *metric trees* were introduced in [ALW17] as are metric spaces (T, r) satisfying the four-point condition (1.1) and admitting branch points, i.e., for all $x_1, x_2, x_3 \in T$, there exists a (necessarily unique) $c_{(T,r)}(x_1, x_2, x_3) \in T$ such that

$$(1.2) \qquad r\big(x_i, c_{(T,r)}(x_1, x_2, x_3)\big) + r\big(c_{(T,r)}(x_1, x_2, x_3), x_j\big) = r(x_i, x_j) \quad \forall i, j \in \{1, 2, 3\}, i \neq j.$$

Trees defined in the sense of graph theory can be equipped with the graph distance to obtain such metric trees. By letting the sizes of the trees grow to infinity, one might get, after a suitable rescaling of the distance, as a continuum limit a path-connected metric tree, which is referred to as $\mathbb{R}$-tree (see for example [Tit77, Dre84, MO90, MNO92, DMT96, DT96, Ter97, Chi01, Eva08, ALW17]).

A new notion of potentially continuum trees, namely the algebraic trees, was introduced in [LW21]. The focus is here shifted from the metric to the tree structure given by the so-called *branch point map* which assigns to each triple of points their branch point. The algebraic trees are defined axiomatically by several conditions on the branch point map:

Definition 1.1 (Algebraic tree). *An* algebraic tree *is a non-empty set T together with a symmetric map $c \colon T^3 \to T$ satisfying the following:*

(2pc) *For all $x_1, x_2 \in T$, $c(x_1, x_2, x_2) = x_2$.*
(3pc) *For all $x_1, x_2, x_3 \in T$, $c(x_1, x_2, c(x_1, x_2, x_3)) = c(x_1, x_2, x_3)$.*
(4pc) *For all $x_1, x_2, x_3, x_4 \in T$,*

$$(1.3) \qquad c(x_1, x_2, x_3) \in \{c(x_1, x_2, x_4), c(x_1, x_3, x_4), c(x_2, x_3, x_4)\}.$$

We call c the branch point map.

As explained, this approach forgets the notion of metric (hence the denomination of *algebraic trees* in contrast with *metric trees*) and puts the emphasis on the tree structure. In order to sample leaves from an algebraic tree, we equip it with a Borel measure. For this, we consider on

an algebraic tree (T, c) the topology generated by so-called *subtree components*, which are sets of the form

$$(1.4) \qquad \mathcal{S}_x(y) := \{z \in T \setminus \{x\} : c(x, z, y) \neq x\},$$

with $x, y \in T$. An *algebraic measure tree* (T, c, μ) consists of a separable algebraic tree (T, c) together with a probability measure μ on the Borel σ-algebra $\mathcal{B}(T, c)$. Associating each algebraic measure tree to the metric measure space where the intrinsic metric is defined with the help of the distribution of branch points (see Definition 3.12), we can use the Gromov-weak topology (see [Gro99, GPW09, Löh13]) to define a metrizable topology on the set $\mathbb{T}$ of (equivalence classes of) algebraic measure trees. Let $\mathbb{T}_2$ be the space of binary algebraic measure trees with no atoms on the skeleton:

$$(1.5) \qquad \mathbb{T}_2 := \big\{(T, c, \mu) \in \mathbb{T} : \text{degrees at most 3, atoms of } \mu \text{ only at leaves}\big\}.$$

Due to the algebraic nature of the concept one might desire a more combinatorial notion of convergence. This can be done by considering the *sample shape convergence* which, similarly to the Gromov-weak convergence, is based on a sampling procedure but evaluates the shape spanned by the sample rather than the matrix of mutual intrinsic distances. It was shown in [LW21] that on $\mathbb{T}_2$ the sample shape convergence is equivalent to the (branch point distribution distance) Gromov-weak convergence and that for this topology, $\mathbb{T}_2$ is a compact subspace.

This state space $\mathbb{T}_2$ has served to construct and study in [LMW20] the Aldous diffusion ([Ald00]), and more generally in [NW] the α-Ford diffusion (see [For]), which are Markov processes on binary trees without edge lengths. The compactness of $\mathbb{T}_2$ allows to get around tightness issues in these construction. Also, the sample shape convergence gives rise to a family of convergence determining classes of functions which are very useful when one wants to study tree-valued stochastic processes.

More recently, genealogical trees with two-level dynamics have been investigated. They are motivated by the study of two-level systems in biology, such as host-parasite, cell-virus or individual-species systems, where individuals of the first level are grouped together in clusters to form the second level and both levels are subject to resampling mechanisms. The phylogenies of such systems bring an important light on the paths and dynamical phenomenas that are interplaying. For example in anthropology or ecology these models have served as basis for statistical studies and understand collected data (e.g. [BJ10, JBA19, LBP+21, VAE+09]). For cell models, we can refer for example to [Ban08, BT11, Kim97, MR13] among others. For extend the notion of algebraic measure trees to hierarchical structures, one can equip metric spaces with two-level measures, that are measures on the set of measures. In [Mei19], the space of metric two-level measure spaces was introduced and equipped with the two-level Gromov-weak topology. It is a sampling topology which is consistent with a two-stage sampling procedure in host-parasite systems, where one would first sample hosts and then parasites within the hosts. The idea of representing a population with hierarchical structure by such a two-level measure is not new (see e.g. [DHW90, Wu91, Wu94, GHW95, DHV95, DHV96, GH00, DGW04, Daw18]). Metric two-level measure spaces allow for instance to study two-level tree-valued processes describing the evolution of ancestral relationships in hierarchical populations, where the metric encodes the genealogical distances between individuals. We applied this theory to define the two-level tree-valued Fleming-Viot dynamics in [Nus].
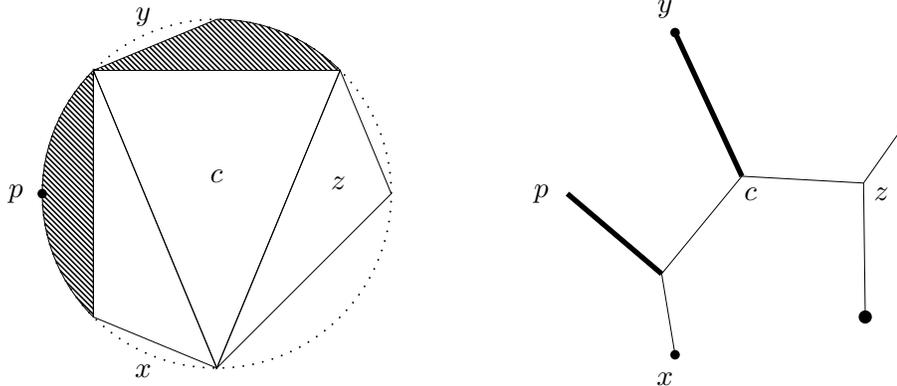
FIGURE 2. *On the left:* A sub-triangulation of the circle with three empty triangles, four empty circular segments and two filled areas. *On the right:* The tree coded by the sub-triangulation. The leaf p does not carry an atom, but the four other leaves carry a weight given by the lengths of the corresponding arcs. The two thick segments carry a non-atomic mass.

Our goal here is to adapt the approach of [Mei19] for two-level trees in the metric setup to the algebraic one. An *algebraic two-level measure tree* (T, c, ν) is thus defined as a separable algebraic tree (T, c) together with a two-level measure $\nu \in \mathcal{M}_1(\mathcal{M}_1(T, c))$, i.e. a Borel probability measure on the set of Borel probability measures on (T, c). Once more, we want to draw our attention away from the distances and focus more on the tree-shapes to define a combinatorial version of the two-level Gromov-weak topology. In particular, we are interested in extending the results in [LW21] to the space $\mathbb{T}_2^{(2)}$ of (equivalence classes of) binary algebraic two-level measure trees (see (4.7) for the proper definition).

When dealing with a two-level measure, say ν, an important ingredient is the intensity measure $M_\nu \in \mathcal{M}_1(T, c)$ defined by

$$M_\nu(\cdot) := \int \nu(\mathrm{d}\mu)\mu(\cdot).$$

Note that if (T, c, ν) is an algebraic two-level measure tree, (T, c, M_ν) is an algebraic measure tree. An algebraic two-level measure tree thus naturally defines a probability measure on $\mathbb{T}_2$ and we know that $\mathcal{M}_1(\mathbb{T}_2)$ is compact since $\mathbb{T}_2$ is compact. However, it seems that the compactness of $\mathbb{T}_2$ cannot be directly used to prove the compactness of $\mathbb{T}_2^{(2)}$. Notice that the information contained in the two-level setup can not be summarized by a mean behaviour, in particular:

- the map $(T, c, \nu) \mapsto (T, c, M_\nu)$ is continuous from $\mathbb{T}_2^{(2)}$ to $\mathbb{T}_2$, but it is not injective. To see this, consider for example, $T = \{x, y\}$, $c_{\{x,y\}}$ the only branch point map on T and the two-level measures $\frac{1}{2}(\delta_{\delta_x} + \delta_{\delta_y})$ and $\delta_{\frac{1}{2}(\delta_x + \delta_y)}$ which both have the same intensity measure $\frac{1}{2}(\delta_x + \delta_y)$.
- the map that sends an a2m tree to the corresponding measure on $\mathbb{T}_2$ is not injective either. Indeed, $(\{x, y\}, c, \delta_x)$, $(\{x, y\}, c, \delta_y)$ and $(\{x\}, c, \delta_x)$ are equivalent in $\mathbb{T}_2$, so the measures on $\mathbb{T}_2$ associated to $(\{x, y\}, c, \frac{1}{2}(\delta_{\delta_x} + \delta_{\delta_y}))$ and $(\{x\}, c, \delta_{\delta_x})$ are equal.
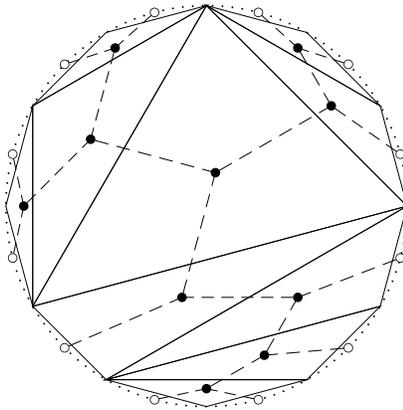
FIGURE 3. A triangulation of the 12-gon. Here, the coded tree is the dual graph, with uniform distribution on the leaves.

For this reason, we show the compactness of $\mathbb{T}_2^{(2)}$ by adapting a key ingredient used in [LW21] to prove the compactness of $\mathbb{T}_2$, namely the coding of binary algebraic measure trees by *sub-triangulations* of the circle. A similar encoding was first introduced by Aldous [Ald94a, Ald94b]. In [LW21], a sub-triangulation of the circle $\mathbb{S}$ is defined as a closed, non-empty subset C of the disc satisfying the following two conditions:

(1) The complement of C in the convex hull of C consists of open interiors of triangles.
(2) C is the union of non-crossing (non-intersecting except at endpoints), possibly degenerate closed straight line segments with endpoints in $\mathbb{S}$.

In this coding, branch points correspond to empty triangles, leaves carrying atoms to empty circular segments, and line segments with non-atomic mass to "filled areas" (see Figure 2). Moreover, the arc lengths play an important role as they encode the way the mass is distributed in the algebraic measure tree. For example, a triangulation of an n-gon encodes the dual graph, equipped with the measure on the leaves given by the Lebesgue measure of the corresponding arcs of the circle line (see Figure 3).

We extend this coding in the two-level case. For that, one cannot simply rely on the Lebesgue measure of arcs anymore to encode all the information on the distribution of the random μ-mass in the tree. We rather need to replace the Lebesgue measure by a two-level measure $K \in \mathcal{M}_1(\mathcal{M}_1(\mathbb{S}))$ on the circle and we formally construct the *coding map* that associates an algebraic two-level measure tree in $\mathbb{T}_2^{(2)}$ to a pair (C, K) where C is a sub-triangulation of the circle, and K is a two-level measure on the circle line.

We show that the coding map is continuous and surjective when the set of sub-triangulations is equipped with the Hausdorff metric topology and the set of two-level measures on the circle line with the weak topology. Using that both of these topologies are compact, our main result (Theorem 2) states that the space $\mathbb{T}_2^{(2)}$ is again compact. As an application, we finish with the construction of the Kingman algebraic two-level measure tree, which is the nested Kingman coalescent measure tree without branch length. This tree is the algebraic tree associated with the metric nested Kingman coalescent tree of [Mei19].

Outline. We recall in Section 2 some definitions and results on metric one and two-level measure spaces. In Section 3, we introduce the space of (equivalence classes of) order separable algebraic measure trees, and equip it with a separable, metrizable topology based on the two-level Gromov-weak convergence of metric reprensatives. Section 4 is devoted to the coding of binary algebraic two-level measure trees by sub-triangulations of the circle together with a two-level measure on the circle line. In Section 5, we introduce with the two-level sample shape another topology on the subspace of binary algebraic two-level measure trees and we use the encoding of Section 4 to show that both topologies are equivalent and compact. Finally, we apply this theory in Section 6 to define the Kingman algebraic two-level measure tree.

2. METRIC ONE-LEVEL AND TWO-LEVEL MEASURE SPACES

We recall the basics on the theory of metric measure spaces in Section 2.1 and shortly explain how it is extended to a two-level setup in Section 2.2. This serves as a basis for constructing algebraic two-level measure trees.

2.1. **Metric measure spaces.** We introduce in this section the set of equivalence classes of metric measure spaces (see [Gro99, GPW09]). Equipped with the topology induced by so-called distance polynomials, it is a Polish space. Denoting by $\mathcal{M}_1(X)$ the set of Borel probability measure on a metric space X, we start with the definition of a metric measure space.

Definition 2.1 (Metric measure space). *A metric measure space (mm space) (X, r, μ) is a non-empty Polish space (X, r) together with a Borel probability measure $\mu \in \mathcal{M}_1(X, r)$.*

Two metric measure spaces are equivalent if there exists a measure-preserving isometry between the supports of their respective measures is enough. Thus, by considering the equivalence class of a metric measure space (X, r, μ), one focuses on the structure of μ and on the restriction of r to the support of μ denoted by $\operatorname{supp}(\mu)$, rather than the whole set X.

Definition 2.2 (Equivalence of metric measure spaces).   (1) *Two mm spaces (X_i, r_i, μ_i), $i = 1, 2$, are called* mm-isomorphic *if there exists a measurable function $f \colon X_1 \to X_2$ such that $\mu_2 = f_* \mu_1$ and f is isometric on the set $\operatorname{supp}(\mu_1)$ (but not necessarily on the whole space X_1). The function f is called an* mm-isomorphism.
   (2) *The relation of being mm-isomorphic is an equivalence relation on the set of metric measure spaces.*
   The set of equivalence classes of metric measure spaces is denoted by $\mathbb{M}$.

The following notion of test functions is based on the idea of sampling finite spaces of (X, r) by means of μ, that is, we sample a finite number of points and evaluate the matrix of their mutual distances.

Definition 2.3 (Distance polynomials). *A distance polynomial is a function $\Phi \colon \mathbb{M} \to \mathbb{R}$ of the form*

$$(2.1) \qquad \Phi(\chi) := \int_{X^m} \mu^{\otimes m}(\mathrm{d}\underline{u}) \varphi\left(\left(r(u_i, u_j)_{i,j}\right)\right),$$

where $\chi = (X, r, \mu)$, $m \in \mathbb{N}$, and $\varphi \in \mathcal{C}_b(\mathbb{R}^{m^2})$. We write Π_r for the set of all distance polynomials.

The space $\mathbb{M}$ is then equipped with the coarsest topology such that all distance polynomials are continuous.

Definition 2.4 (Gromov-weak topology). *The* Gromov-weak topology *is the initial topology on $\mathbb{M}$ induced by the test functions in Π_r. That is, a sequence of mm spaces $(\chi_n)_{n \in N}$ is said to* converge Gromov-weakly *to χ in $\mathbb{M}$ if and only if $\Phi(\chi_n)$ converges to $\Phi(\chi)$ in $\mathbb{R}$, for all polynomials $\Phi \in \Pi_r$.*

With the following result, the space of metric measure spaces is suitable to do probability theory on it (see [GPW09, Theorem 1]).

Proposition 2.5 ($\mathbb{M}$ is Polish). *Equipped with the Gromov-weak topology, $\mathbb{M}$ is a Polish space.*

2.2. **Metric two-level measure spaces.** We here introduce the two-level analogues of the metric measure spaces from [Mei19]. Equipped with the topology induced by two-level distance polynomials, the set of equivalence classes of metric two-level measure spaces is also Polish.

Definition 2.6 (Metric two-level measure space). *A metric two-level measure space (m2m space) (X, r, ν) is a non-empty Polish space (X, r) together with a* two-level measure *$\nu \in \mathcal{M}_1(\mathcal{M}_1(X, r))$, i.e. a Borel probability measure on the set of Borel probability measures on (X, r).*

In order to define equivalence classes of m2m spaces, we introduce the so-called two-level push-forward operator. Let (X, r) and $(\widehat{X}, \widehat{r})$ be Polish metric spaces and g be a Borel measurable function from X to $\widehat{X}$. As usual, $g_*\mu$ denotes the push-forward measure $\mu \circ g^{-1}$ for a Borel probability measure $\mu \in \mathcal{M}_1(X)$. We regard g_* as an operator

$$(2.2) \qquad \begin{aligned} g_* \quad : \quad \mathcal{M}_1(X) \quad &\longrightarrow \quad \mathcal{M}_1(\widehat{X}) \\ \mu \quad &\longmapsto \quad g_*\mu. \end{aligned}$$

and call g_* the *(one-level) push-forward operator* of g. This enables us to define the *two-level push-forward operator g_{**} of g* by

$$(2.3) \qquad \begin{aligned} g_{**} \quad : \quad \mathcal{M}_1(\mathcal{M}_1(X)) \quad &\longrightarrow \quad \mathcal{M}_1(\mathcal{M}_1(\widehat{X})) \\ \nu \quad &\longmapsto \quad g_{**}\nu := \nu \circ (g_*)^{-1}. \end{aligned}$$

We need yet another notion to define equivalence classes of m2m spaces. In the one-level case, the equivalence of two metric measure spaces is given by a measure-preserving isometry between the supports of their respective measures. For an m2m space, the measure μ is replaced by a two-level measure ν and the analog of $\mathrm{supp}(\mu)$ is the support of the *intensity measure* $M_\nu \in \mathcal{M}_1(X)$, also called *first moment measure*, which is defined by

$$(2.4) \qquad M_\nu(\cdot) = \int \nu(\mathrm{d}\mu)\mu(\cdot).$$

This notion allows to easily adapt some results on metric measure spaces to m2m spaces by replacing the (one-level) measure μ by the intensity measure of the two-level measure ν (compare for example [GPW09, Proposition 7.1] to [Mei19, Theorem 7.2] in the case $\nu \in \mathcal{M}_1(\mathcal{M}_1(X))$). We will also exploit this idea in the case of algebraic trees below (e.g. in Definition 3.12).

We are now able to define a notion of equivalence for m2m spaces.

Definition 2.7 (Equivalence of m2m spaces).   (1) *Two m2m spaces* (X_i, r_i, ν_i), $i = 1, 2$, *are called* m2m-isomorphic *if there exists a measurable function* $f \colon X_1 \to X_2$ *such that* $\nu_2 = f_{**}\nu_1$ *and* f *is isometric on the set* $\operatorname{supp}(M_{\nu_1})$ *(but not necessarily on the whole space* X_1*). The function* f *is called an* m2m-isomorphism.

   (2) *The relation of being m2m-isomorphic is an equivalence relation on the set of m2m spaces.*

   The set of equivalence classes of m2m spaces is denoted by $\mathbb{M}^{(2)}$.

The following notion of test functions is based on the idea of sampling finite spaces of (X, r) by means of ν, i.e. we first sample measures from $\mathcal{M}_1(X)$ according to ν and then with each sampled measure, we sample finitely many points in X.

Definition 2.8 (Two-level distance polynomials). *A two-level distance polynomial is a function* $\Phi \colon \mathbb{M}^{(2)} \to \mathbb{R}$ *of the form*

$$(2.5) \qquad \Phi(\chi) := \int \nu^{\otimes m}(\mathrm{d}\underline{\mu}) \int \bigotimes_{i=1}^{m} \mu_i(\mathrm{d}\underline{u_i}) \varphi\left(\left(r(u_{ij}, u_{i'j'})_{(i,j),(i',j')}\right)\right),$$

where $\chi = (X, r, \nu)$, $m \in \mathbb{N}$, $\underline{n} \in \mathbb{N}^m$ *and* $\varphi \in \mathcal{C}_b(\mathbb{R}^{|\underline{n}|^2})$. *We write* $\Pi_r^{(2)}$ *for the set of all two-level distance polynomials.*

The space $\mathbb{M}^{(2)}$ is then equipped with the coarsest topology such that all two-level distance polynomials are continuous.

Definition 2.9 (Two-level Gromov-weak topology). *The* two-level Gromov-weak topology *is the initial topology on* $\mathbb{M}^{(2)}$ *induced by the test functions in* $\Pi_r^{(2)}$. *A sequence of m2m spaces* $(\chi_n)_{n \in N}$ *is said to* converge two-level Gromov-weakly *to* χ *in* $\mathbb{M}^{(2)}$ *if and only if* $\Phi(\chi_n)$ *converges to* $\Phi(\chi)$ *in* $\mathbb{R}$, *for all polynomials* $\Phi \in \Pi_r^{(2)}$.

The following result is given by Proposition 4.6 and Theorem 8.1 in [Mei19]:

Proposition 2.10 ($\mathbb{M}^{(2)}$ is Polish). *Equipped with the two-level Gromov-weak topology,* $\mathbb{M}^{(2)}$ *is a Polish space.*

3. The space of algebraic two-level measure trees

In this section, we introduce the notion of algebraic two-level measure spaces. We first recall in Section 3.1 tools on algebraic trees from [LW21] on which we will rely in our set-up. In Section 3.2, we add two-level measures on algebraic trees and equip the space of algebraic two-level measure trees with a separable and metrizable topology.

3.1. **Algebraic trees.** After recalling some basic examples and definitions, we describe the natural topology that exists on an algebraic tree. We then present relations between algebraic trees and $\mathbb{R}$-trees that we will use to exploit results on metric two-level measure spaces. We finish with the notion of struture preserving morphisms to define a notion of equivalence on the space of a2m trees in Section 3.

Recall from Definition 1.1 the notion of algebraic tree given by the branch point map that sends a triplet of 3 points in the tree to its branch point. For an algebraic tree (T, c), one can define for every $x, y \in T$, the *interval* $[x, y]$ by

$$(3.1) \qquad\qquad [x, y] := \{z \in T : c(x, y, z) = z\}.$$

We say that $\{x, y\}$ is an *edge* if $x \neq y$ and $[x, y] = \{x, y\}$, that is, there is "nothing between x and y", and we denote by

(3.2) $$\text{edge}(T, c)$$

the set of edges of (T, c). If T is finite, then an algebraic tree (T, c) simply is an undirected graph and $\text{edge}(T, c)$ is the set of edges in the corresponding graph.

Example 3.1 (Totally ordered spaces as algebraic trees). Let $(T, \leq)$ be a totally ordered space. For all $x, y, z \in T$ such that $x \leq y \leq z$, we define $c_\leq(x, y, z) := y$. It is easy to see that $c_\leq$ is symmetric and satisfies the conditions (2pc)–(4pc), so that $(T, c_\leq)$ is an algebraic tree. Moreover, the interval $[x, y]$ in $(T, c_\leq)$ coincides with the order interval $\{z \in T : x \leq z \leq y\}$.

Conversely, if (T, c) is an algebraic tree and $\rho \in T$ a distinguished point (often called the *root*), we can define a *partial order* $\leq_\rho$ by letting for $x, y \in T$,

(3.3) $$x \leq_\rho y \iff x \in [\rho, y].$$

This partial order allows us to define a notion of completeness of algebraic trees.

Definition 3.2 (Directed order complete trees). *Let (T, c) be an algebraic tree. We call (T, c) (directed) order complete if for all $\rho \in T$, the supremum of every totally ordered, non-empty subset exists in the partially ordered set $(T, \leq_\rho)$.*

The following definition gives the analogs of complete $\mathbb{R}$-trees, i.e., $\mathbb{R}$-trees that are complete as metric spaces.

Definition 3.3 (Algebraic continuum tree). *We call an algebraic tree (T, c) algebraic continuum tree if the following two conditions hold:*

(1) *(T, c) is order complete.*
(2) *$\text{edge}(T, c) = \emptyset$.*

We can then equip an algebraic tree with a natural topology. To this end, we introduce, for each point $x \in T$, an equivalence relation $\sim_x$ on $T \setminus \{x\}$ such that for all $y, z \in T \setminus \{x\}$, $y \sim_x z$ if and only if $c(x, y, z) \neq x$. For $y \in T \setminus \{x\}$, we denote by

(3.4) $$\mathcal{S}_x(y) := \{z \in T \setminus \{x\} : z \sim_x y\}$$

the equivalence class of y for this equivalence relation $\sim_x$. We also call $\mathcal{S}_x(y)$ the *component* of $T \setminus \{x\}$ containing y. The *component topology* is defined as the one generated by the set of all components $\mathcal{S}_x(y)$ with $x \neq y$, $x, y \in T$. In the following, we will suppose that algebraic trees are equipped with this topology.

Furthermore, we will often assume the topology to be separable. But for many of our results, a condition on the number of edges will be crucial, so we rely on the following notion of separability (see also [LW21, Example 2.23]).

Definition 3.4 (Order separability). *We call an algebraic tree (T, c) order separable if it is separable w.r.t. the component topology and has at most countably many edges.*

Remark 3.5 (Sufficient condition for metrizability). The component topology of any algebraic tree is Hausdorff [LW21, Lemma 2.18]. Thus, by Propositions 2.19 and 2.20 in [LW21], any order complete, order separable algebraic tree is a compact, second countable Hausdorff space. In particular, the component topology is metrizable and thus Polish.
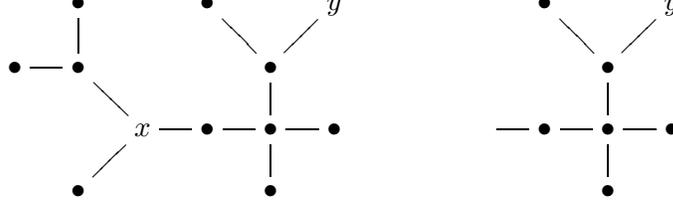
FIGURE 4. On the left: x and y are two points of the tree. On the right: the component $\mathcal{S}_x(y)$.

We also recall a definition of structure-preserving morphisms of algebraic trees which will allow to define equivalence classes of trees in Section 3.

Definition 3.6 (Morphisms). *Let (T, c) and $(\widehat{T}, \widehat{c})$ be algebraic trees. A map $f \colon T \to \widehat{T}$ is called a* tree homomorphism *(from T into $\widehat{T}$) if for all $x, y, z \in T$,*

$$(3.5) \qquad f\big(c(x, y, z)\big) = \widehat{c}\big(f(x), f(y), f(z)\big),$$

or equivalently, if for all $x, y \in T$,

$$(3.6) \qquad f([x, y]) \subseteq [f(x), f(y)].$$

We refer to a bijective tree homomorphism as tree isomorphism.

There exists a connection between algebraic trees and *metric trees*, which are defined in [ALW17] as metric spaces (T, r) satisfying the 4-point condition (1.1) and admitting branch points, i.e., for all x_1, x_2, x_3, there exists a (necessarily unique) $c_{(T,r)}(x_1, x_2, x_3) \in T$ such that

$$(3.7) \qquad r\big(x_i, c_{(T,r)}(x_1, x_2, x_3)\big) + r\big(c_{(T,r)}(x_1, x_2, x_3), x_j\big) = r(x_i, x_j) \quad \forall i, j \in \{1, 2, 3\}, i \neq j.$$

Obviously, $c_{(T,r)}$ satisfies the conditions of Definition 1.1 and we call $(T, c_{(T,r)})$ the algebraic tree *induced by* (T, r), and (T, r) a *metric representation* of $(T, c_{(T,r)})$.

Remark 3.7 (Homeomorphisms are tree homomorphisms). Let (T, r) and $(\widehat{T}, \widehat{r})$ be two $\mathbb{R}$-trees. Since the branch point map can be expressed in terms of intervals, a homeomorphism f between (T, r) and $(\widehat{T}, \widehat{r})$ is also a tree homomorphism between the corresponding induced algebraic trees (see [LW21, Lemma 2.35]).

Conversely, under the assumption of order separability, one can build a metric representation of an algebraic tree (T, c) as follows. Equip (T, c) with the Borel σ-algebra $\mathcal{B}(T, c)$ of the component topology. For any measure ξ on $(T, \mathcal{B}(T, c))$ such that ξ is finite on every interval, we consider the following pseudometric

$$(3.8) \qquad r_\xi(x, y) := \xi([x, y]) - \frac{1}{2}\xi(\{x\}) - \frac{1}{2}\xi(\{y\}), \quad x, y \in T.$$

For such a measure ξ, denote by T_ξ the set of equivalence classes of points in T that are at distance zero for r_ξ. More precisely, x and y are equivalent if ξ puts weight zero on the interval $[x, y]$. Denoting again the quotient metric on T_ξ by r_ξ, (T_ξ, r_ξ) is thus the quotient metric space. Let $\pi_\xi \colon T \to T_\xi$ be the canonical projection. We have the following [LW21, Lemma 2.29].

Lemma 3.8 ((T_ξ, r_ξ) is a metric tree). *Let (T,c) be an algebraic tree, and ξ a measure on (T,c) with $\xi([x,y]) < \infty$ for all $x, y \in T$. Then the quotient space (T_ξ, r_ξ) is a metric tree, and the canonical projection π_ξ is a tree homomorphism.*

Furthermore, if the measure ξ satisfies $\xi[x,y] > 0$ for all $x, y \in T$, then r_ξ is a metric on T. Such a measure ξ always exists in the case of order separable algebraic trees (see [LW21, Lemma 2.32]).

3.2. **Algebraic two-level measure trees.** In this section, we define algebraic two-level measure trees (a2m trees for short) and equip the space of (equivalence classes of) a2m trees with a topology related to the two-level Gromov-weak topology for m2m spaces. We then give a result that allows to construct a2m trees from a partial knowledge of the two-level mass distribution.

Algebraic measure trees are introduced in [LW21] as order separable algebraic trees equipped with probability measures. We extend this idea by equipping algebraic trees with two-level probability measures. The order separability condition is crucial in the sequel.

Definition 3.9 (Algebraic two-level measure tree). *An* algebraic two-level measure tree (a2m tree) (T, c, ν) *is an order separable algebraic tree (T,c) together with a* two-level measure $\nu \in \mathcal{M}_1(\mathcal{M}_1(T,c))$, *i.e. a Borel probability measure on the set of Borel probability measures on (T,c).*

Similarly to m2m spaces, we will consider equivalence classes of a2m trees. Recall from (2.4) the intensity measure M_ν and define a *subtree* of T as a set $A \subseteq T$ such that $c(A^3) = A$.

Definition 3.10 (Equivalence of a2m trees). (1) *Two a2m trees (T_i, c_i, ν_i), $i = 1, 2$, are called* a2m-isomorphic *if there exist subtrees A_i of T_i with $M_\nu(A_i) = 1$ and a tree isomorphism $f\colon A_1 \to A_2$ such that $\nu_2 = f_{**}\nu_1$. The function f is called an* a2m-isomorphism.
  (2) *The relation of being a2m-isomorphic is an equivalence relation on the set of a2m trees. The set of equivalence classes of a2m trees is denoted by $\mathbb{T}^{(2)}$.*

For an a2m tree (T, c, ν), let $A := c((\mathrm{supp}(M_\nu))^3)$. By [Mei19, Corollary 2.3], the support of ν is a subset of $\{\mu \in \mathcal{M}_1(T) \; : \; \mathrm{supp}(\mu) \subseteq \mathrm{supp}(M_\nu)\}$. Thus the subtree A of (T,c) is such that (T, c, ν) is equivalent to (A, c', ν'), where c' is the restriction of c to A^3 and ν' is the restriction of ν to $\mathcal{M}_1(A, c')$. Therefore, we define for an a2m tree $\chi := (T, c, \nu)$,

$$\tag{3.9} \mathrm{supp}(\chi) := c((\mathrm{supp}(M_\nu))^3).$$

With this in mind, note that by considering equivalence classes of a2m trees in the latter, we will only focus on the restriction of c and ν to the support of M_ν.

The example below is the analogue of [LW21, Example 3.4], which states that there is only one equivalence class of algebraic measure trees without branch points nor atoms. For an algebraic tree (T, c), we call $v \in T$ a *branch point* if there exist $x_1, x_2, x_3 \in T \setminus \{v\}$ such that $v = c(x_1, x_2, x_3)$. We denote by $\mathrm{br}(T, c)$ the set of branch points and if $\chi := (T, c, \nu)$,

$$\tag{3.10} \mathrm{br}(\chi) := \mathrm{br}(T, c) \cap \mathrm{supp}(\chi).$$

We will also denote by $\mathrm{at}(\mu)$ the set of atoms of a measure μ on (T,c). In particular, $\mathrm{at}(M_\nu)$ are the atoms of the intensity measure M_ν defined in (2.4).

Example 3.11 (A2m trees without branch points nor atoms). Let $\chi := (T, c, \nu)$ be an a2m tree such that $\mathrm{br}(\chi) = \emptyset$ and $\mathrm{at}(M_\nu) = \emptyset$. By [LW21, Theorem 1], there is a tree isomorphism from T into $[0, 1]$ and we may thus assume $T \subseteq [0, 1]$. Let $F_{M_\nu}\colon [0, 1] \to [0, 1]$ be the cumulative

distribution function of the intensity measure M_ν. Define $\widetilde{\nu} := (F_{M_\nu})_{**}\nu$. Then, $M_{\widetilde{\nu}} = F_{M_\nu*}M_\nu$ is the Lebesgue measure on $[0,1]$. Indeed since $\mathrm{at}(M_\nu) = \emptyset$, the function F_{M_ν} is continuous and for $a \in [0,1]$, the set $\{x : F_{M_\nu} \leq a\}$ is of the form $[0, x_a]$ with $F_{M_\nu}(x_a) = a$, so that $F_{M_{\widetilde{\nu}}}(a) = M_\nu(\{x : F_{M_\nu} \leq a\}) = M_\nu([0, x_a]) = F_\nu(x_a) = a$.

Let $A := \{x \in \mathrm{supp}(M_\nu) \; : \; \text{there is no } (y_n)_n \in ([0,1] \setminus \mathrm{supp}(M_\nu))^{\mathbb{N}} \text{such that} y_n < x, y_n \to x\}$ be the support of M_ν with left boundary points removed. Then F_{M_ν} restricted to A is bijective and thus an a2m-isomorphism from (A, c, ν) onto $([0,1], c_\leq, \widetilde{\nu})$ where $c_\leq$ is defined in Example 3.1.

We showed that if we consider the equivalence class of an a2m tree (T, c, ν) such that $\mathrm{br}(\chi) = \mathrm{at}(M_\nu) = \emptyset$, we can always assume that $T = [0,1]$ and M_ν is the Lebesgue measure $\lambda_{[0,1]}$. More generally, if there exists an interval $(v, w) \subseteq \mathrm{supp}(M_\nu)$ such that $(v, w) \cap \mathrm{at}(M_\nu) = \emptyset$ and $(v, w) \cap \mathrm{br}(T, c) = \emptyset$, we can assume without loss of generality that M_ν restricted to (v, w) is the Lebesgue measure $\lambda_{(v,w)}$.

We now equip the space $\mathbb{T}^{(2)}$ with a topology that relies on the two-level Gromov-weak topology on $\mathbb{M}^{(2)}$. To do so, we use that due to the order separability assumption, a2m trees allow for metric representations (see [LW21, Theorem 1]). But in order to get a useful topology on $\mathbb{T}^{(2)}$, we consider a particular metric representation of an a2m tree (T, c, ν) by using the metric r_ξ defined in (3.8) with ξ being the so-called *branch point distribution*.

Definition 3.12 (Branch point distribution). *We call* branch point distribution *of an a2m tree* (T, c, ν) *the push-forward of* $M_\nu^{\otimes 3}$ *under the branch point map,*

$$(3.11) \qquad\qquad \xi := c_* M_\nu^{\otimes 3}.$$

Recall from Lemma 3.8 that the quotient space (T_ξ, r_ξ) is a metric space. Therefore the following map associates a particular metric representation to each a2m tree.

Definition 3.13 (Selection map ι). *Define the* selection map $\iota: \mathbb{T}^{(2)} \to \mathbb{M}^{(2)}$ *by*

$$(3.12) \qquad\qquad \iota(T, c, \nu) := (T_\xi, r_\xi, \nu_\xi),$$

where $\xi = c_* M_\nu^{\otimes 3}$ *is the branch point distribution of* (T, c, ν), (T_ξ, r_ξ) *is the quotient metric space, and* $\nu_\xi := \pi_{\xi_{**}}\nu$ *is the two-level push-forward measure of* ν *under the canonical projection* π_ξ.

It is easy to see that if two a2m trees are equivalent, then their images under ι are also equivalent, with the same isomorphism (see Remark 3.7). Thus the selection map is well-defined. The following result states that ι indeed selects metric representations, and is injective.

Proposition 3.14 (ι is an embedding). *For all* $\chi \in \mathbb{T}^{(2)}$, $\iota(\chi)$ *is a metric representation of* χ. *Moreover, the selection map* $\iota: \mathbb{T}^{(2)} \to \mathbb{M}^{(2)}$ *is injective.*

Proof. It is enough to show that for all $\chi \in \mathbb{T}^{(2)}$, the algebraic tree induced by $\iota(\chi)$ is χ, because by Remark 3.7, if $\iota(\chi)$ and $\iota(\chi')$ are equivalent in $\mathbb{M}^{(2)}$, the measure preserving bijective homeomorphism $f: \iota(\chi) \to \iota(\chi')$ is a tree homomorphism and thus yields an a2m-isomorphism on the corresponding a2m trees χ and χ'.

Fix $\chi = (T, c, \nu) \in \mathbb{T}^{(2)}$. We can assume w.l.o.g that for all $v \in \mathrm{br}(T)$, $\xi\{v\} > 0$. By Lemma 3.8, the canonical projection $\pi_\xi: T \to T_\xi$ is a (surjective) tree homomorphism. Therefore, to show equivalence of (T, c, ν) and $(T_\xi, c_{(T_\xi, r_\xi)}, \nu_\xi)$, it is sufficient to show that π_ξ is injective on

a subtree $A \subseteq T$ such that $M_\nu(A) = 1$. We take $A := T \setminus N$ with $N := \{v \in T : \pi_\xi(v) \neq \{v\}\}$, and π_ξ is injective on A.

Let us first show that $M_\nu(A) = 1$. If $\pi_\xi(v) \neq \{v\}$ for some $v \in T$, then we can find some $u \neq v$ in T such that $r_\xi(u, v) = 0$, i.e., $\xi([u, v]) - \frac{1}{2}\xi\{u\} - \frac{1}{2}\xi\{v\} = 0$. Thus, $\xi\{v\} = 0$ and $M_\nu\{v\} = 0$. Moreover, since π_ξ is a tree isomorphism, $w \in \pi_\xi(v)$ implies $[v, w] \subseteq \pi_\xi(v)$. But due to order separability, there are at most countably many non-degenerate, disjoint closed intervals in T, which implies that $\pi_\xi(N)$ is countable, and thus $M_\nu(N) = 0$.

Finally, to see that A is a subtree, let $x, y, z \in A$. If $v := c(x, y, z) \in \{x, y, z\}$, then $v \in A$. Otherwise, $v \in \mathrm{br}(T, c)$, and hence $M_\nu\{v\} > 0$ by assumption, which implies that $\pi_\xi(v) = \{v\}$, i.e. $v \in A$. $\qquad\square$

Since the selection map is an embedding, it is suitable to define a useful topology on $\mathbb{T}^{(2)}$.

Definition 3.15 (Two-level bpdd-Gromov-weak topology). *Let $\mathbb{M}^{(2)}$ be equipped with the two-level Gromov-weak topology. We call the topology induced on $\mathbb{T}^{(2)}$ by the selection map ι two-level branch-point distribution distance Gromov-weak topology (two-level bpdd-Gromov-weak topology).*

An immediate consequence of Proposition 2.6 is the following:

Corollary 3.16 (Separability and metrizability). *$\mathbb{T}^{(2)}$ equipped with the two-level bpdd-Gromov-weak topology is a separable, metrizable space.*

Proof. The two-level Gromov-weak topology on $\mathbb{M}^{(2)}$ is separable and metrizable by the two-level Gromov-Prokhorov metric $d_{2\mathrm{GP}}$ (see [Mei19, Proposition 4.6, Theorem 8.1]). Now define for $\chi_1, \chi_2 \in \mathbb{T}^{(2)}$,

$$(3.13) \qquad\qquad d_{2\mathrm{BGP}}(\chi_1, \chi_2) := d_{2\mathrm{GP}}(\iota(\chi_1), \iota(\chi_2)).$$

Since ι is injective, $d_{2\mathrm{BGP}}$ is a metric on $\mathbb{T}^{(2)}$ and it induces the two-level bpdd-Gromov-weak topology. $\qquad\square$

4. Triangulations of the circle

A key point in proving compactness of $\mathbb{T}_2$ in [LW21] was to encode binary algebraic measure trees as sub-triangulations of the circle, and then exploit the compactness of the space of sub-triangulations equipped with the Hausdorff topology. We are interested in formulating a similar result in the case of binary a2m trees. In order to encode the random masses of the subtree components, we add a two-level measure on the circle to the sub-triangulation.

We first define what we mean by sub-triangulations of the circle and in Subsection 4.2, we construct the map that associates an a2m tree to every sub-triangulation together with a two-level measure on the circle line and show that, as in the one-level case, this coding map is continuous and surjective.

4.1. **The space of sub-triangulations of the circle.** In the whole Section 4, we fix $\mathbb{D}$ a closed disc of circumference 1, and denote by $\mathbb{S} := \partial\mathbb{D}$ the circle. We will repeatedly identify $\mathbb{S}$ with $[0, 1]$ where the endpoints are glued. In this section, we denote by λ_I the Lebesgue measure on an interval $I \subseteq \mathbb{S} \simeq [0, 1]$. For a subset $A \subseteq \mathbb{D}$, we denote by $\overline{A}$, A°, ∂A and $\mathrm{conv}(A)$ the closure, the interior, the boundary and the convex hull of A, respectively. Furthermore, we define

$$(4.1) \qquad\qquad \Delta(A) := \{\text{connected components of } \mathrm{conv}(A) \setminus A\},$$
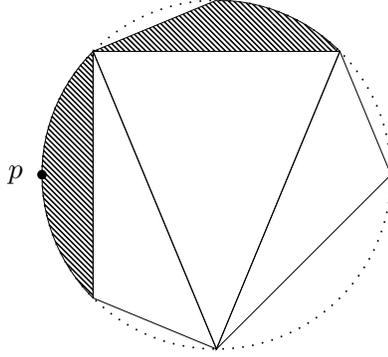
FIGURE 5. A sub-triangulation C of the circle. From the definitions (4.1), (4.2), (4.10) and (4.12), we have in C: $\#\Delta(C) = 3$, $\#\nabla(C) = 4$, $\#\blacksquare(C) = 2$ and $\#\square(C) = 1$, (more precisely, $\square(C) = \{p\}$).

and

$$(4.2) \qquad \nabla(A) := \big\{\text{connected components of } \mathbb{D} \setminus \text{conv}(A)\big\},$$

so that we have the partition of the disc

$$(4.3) \qquad \mathbb{D} = A \uplus \biguplus_{a \in \Delta(A)} a \uplus \biguplus_{b \in \nabla(A)} b,$$

where $\uplus$ denotes the disjoint union.

Definition 4.1 (Sub-triangulations of the circle). *A sub-triangulation of the circle is a closed non-empty subset C of $\mathbb{D}$ satisfying the following two conditions:*

(Tri1) $\Delta(C)$ *consists of open interiors of triangles.*

(Tri2) C *is the union of non-crossing (non-intersecting except at endpoints), possibly degenerate closed straight line segments with endpoints in $\mathbb{S}$.*

We denote the set of sub-triangulations of the circle by $\mathcal{T}$.

Note that given (Tri1), (Tri2) implies that $\nabla(C)$ consists of circular segments with the bounding straight line excluded and the rest of the bounding arc included (see Figure 5).

The set $\mathcal{T}$ can be equipped with a compact metrizable topology in the following way. Let

$$(4.4) \qquad \mathcal{F}(\mathbb{D}) := \big\{F \subseteq \mathbb{D} \ : \ F \neq \emptyset, \ F \text{closed}\big\}.$$

We equip $\mathcal{F}(\mathbb{D})$ with the *Hausdorff metric topology*. That is, we say that a sequence $(F_n)_{n \in \mathbb{N}}$ converges to F in $\mathcal{F}(\mathbb{D})$ if and only if for all $\epsilon > 0$ and all large enough $n \in \mathbb{N}$,

$$(4.5) \qquad F \subseteq F_n^\epsilon \quad \text{and} \quad F_n \subseteq F^\epsilon,$$

where for all $A \in F(\mathbb{D})$, $A^\epsilon := \{x \in \mathbb{D} : d(x, A) < \epsilon\}$. As $\mathbb{D}$ is compact, $\mathcal{F}(\mathbb{D})$ is a compact metrizable space which contains $\mathcal{T}$ and it can be shown that $\mathcal{T}$ is actually a closed subset of $\mathcal{F}(\mathbb{D})$ (see [LW21, Lemma 4.2]).

In order to construct the branch point map from a sub-triangulation in the next subsection, we need another characterization of sub-triangulations of the circle. Roughly speaking, condition

(Tri2) can be replaced by the existence of triangles that separate triples of connected components of $\mathbb{D} \setminus C$.

Proposition 4.2 (Characterization of sub-triangulations). *Let C be a closed non-empty subset of $\mathbb{D}$. Then C is a sub-triangulation of the circle if and only if condition (Tri1) holds, all extreme points of $\operatorname{conv}(C)$ are contained in $\mathbb{S}$ and*

(Tri2)' *For $x, y, z \in \Delta(C) \cup \nabla(C)$ pairwise distinct, there exists a unique $c_{xyz} \in \Delta(C)$ such that x, y, z are subsets of pairwise different connected components of $\mathbb{D} \setminus \partial c_{xyz}$.*

4.2. Coding of binary algebraic two-level measure trees.
We define here a map that associates a binary a2m tree to a sub-triangulation together with a two-level measure on the circle line whose intensity measure is the Lebesgue measure. We show that this coding map is surjective and continuous.

We start by defining the set of binary algebraic two-level measure trees with no atoms on the skeleton. For this, define in an algebraic (T, c) the *degree* of $x \in T$ as the number of components of $T \setminus \{x\}$, and we write

$$(4.6) \qquad \deg(x) := \#\{\mathcal{S}_x(y) : y \in T \setminus \{x\}\}.$$

A *leaf* is then a point $u \in T$ such that $\deg(u) = 1$, and we write $\operatorname{lf}(T, c)$ for the set of leaves. We then define

$$(4.7) \qquad \mathbb{T}_2^{(2)} := \left\{ \chi = \overline{(T, c, \nu)} \in \mathbb{T}^{(2)} : \deg(v) \leq 3 \ \forall v \in T, \ \operatorname{at}(\mu) \subseteq \operatorname{lf}(T, c) \ \forall \mu \in \operatorname{supp}(\nu) \right\}.$$

Note that for all $(T, c, \nu) \in \mathbb{T}_2^{(2)}$, we also have $\operatorname{at}(M_\nu) \subseteq \operatorname{lf}(T, c)$.

In Theorem 1, we introduce a coding map which associates an a2m tree $(T, c, \nu) \in \mathbb{T}_2^{(2)}$ to each two-level sub-triangulation, which are defined below:

Definition 4.3 (Two-level sub-triangulation). *A two-level sub-triangulation of the circle (C, K) consists of a sub-triangulation of the circle C together with a two-level measure $K \in \mathcal{M}_1(\mathcal{M}_1(\mathbb{S}))$ on $\mathbb{S}$ such that its intensity measure M_K is $\lambda_{\mathbb{S}}$ the Lebesgue on the circle line.*

We denote by

$$(4.8) \qquad \mathfrak{D} := \left\{ (C, K) \in \mathcal{T} \times \mathcal{M}_1(\mathcal{M}_1(\mathbb{S})) : M_K = \lambda_{\mathbb{S}} \right\}$$

the set of all two-level sub-triangulations of the circle.

In the construction we give below, the algebraic measure tree (T, c, M_ν) is the one associated to the sub-triangulation C given by the coding map in [LW21, Theorem 2]. In particular, a triangle $x \in \Delta(C)$ will correspond to a branch point of the associated a2m tree and a circular segment $x \in \nabla(C)$ to an atom of the intensity measure M_ν of the a2m tree such that the arc $x \cap \mathbb{S}$ has length the M_ν-mass of the corresponding atom.

Example 4.4 (A simple case). Consider a triangulation C of the regular n-gon into $n - 2$ triangles (see Figure 6). In this case, the coded tree is the dual graph. That is, a triangle corresponds to a branch point of the tree, and two branch points of the tree are connected by an edge if and only if the triangles are adjacent. Each circular segment corresponds to a leaf and M_ν will assign to each leaf the length of the circular segment, i.e. n^{-1} since the n-gon is regular.

We then add on the circle line a two-level measure K such that $M_K = \lambda_{\mathbb{S}}$ by defining

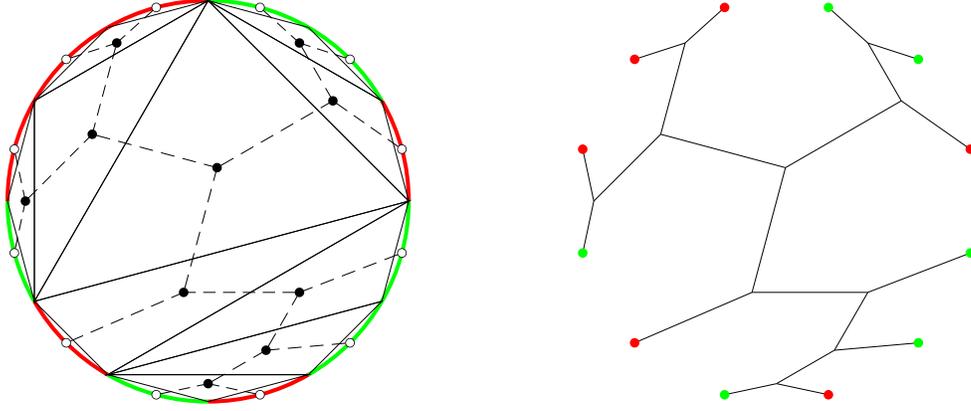$$(4.9) \qquad K = \frac{1}{2}\delta_{\kappa_1} + \frac{1}{2}\delta_{\kappa_2},$$

FIGURE 6. A triangulation of the 12-gon and the tree coded by it. Suppose that we add on the triangulation the two-level measure $K = \frac{1}{2}\delta_{\kappa_1} + \frac{1}{2}\delta_{\kappa_2}$, where κ_1 is the renormalized Lebesgue measure on the union of the red circular segments and κ_2 on the union of the green circular segments. Then the corresponding two-level measure on the tree is given by $\frac{1}{2}\delta_{\mu_1} + \frac{1}{2}\delta_{\mu_2}$ where μ_1 assigns mass $\frac{1}{6}$ to each red leaf and μ_2 to each green leaf.

where $\kappa_1, \kappa_2 \in \mathcal{M}_1(\mathbb{S})$ are defined in Figure 6. Then the associated two-level measure ν on the tree is given by $\frac{1}{2}\delta_{\mu_1} + \frac{1}{2}\delta_{\mu_2}$ where μ_i assigns to each leaf the κ_i-mass carried by the corresponding circular segment.

Remark 4.5 (About the condition $M_K = \lambda_{\mathbb{S}}$)**.** Consider the sub-triangulation C and $K := \delta_{\delta_0}$ as in Figure 7 on the left. We need a condition on K to avoid such a case where M_K puts an atom on an endpoint of a line segment of C separating two circular segments, as we would not know where to assign the mass in the tree. We could have overcome this issue by deciding *a priori* to which of the two circular segment the atom adds mass. However, with this solution, the coding map would not be continuous when $\mathcal{T}$ is equipped with the Hausdorff metric topology and $\mathcal{M}_1(\mathcal{M}_1(\mathbb{S}))$ with the weak topology. Indeed, suppose we decide to add the atoms on endpoints to the mass of the circular segment after the endpoint and let $K_n := \delta_{\frac{1}{2}\delta_0 + \frac{1}{2}\delta_{-n^{-1}}}$ (see Figure 7 on the right). Then the sequence (C, K_n) converges to (C, K) but it is not true of the corresponding a2m trees since the one associated with (C, K_n) has two leaves carrying mass but not the one associated with (C, K).

The condition $M_K = \lambda_{\mathbb{S}}$ prevents such cases from occurring and it seems natural for the specific reason that it also ensures that the algebraic measure tree (T, c, M_ν) corresponding to (C, K) is the one associated to the sub-triangulation C in the one-level case.

Before stating the result, we need further notation. We first define on the sub-triangulation the points that will correspond through the coding construction to leaves without mass. For $x \in \Delta(C)$, let $p_i(x)$, $i = 1, 2, 3$, be the mid-points of the three arcs of $\mathbb{S} \setminus \partial x$ and let $\mathrm{arc}_x(p_i(x))$ be the arc of $\mathbb{S} \setminus \partial x$ that contains $p_i(x)$. Define

$$(4.10) \qquad \Box(C) := \big\{ \{p_i(x)\} : x \in \Delta(C), i \in \{1,2,3\}, \mathrm{arc}_x(p_i(x)) \subseteq C \big\}$$
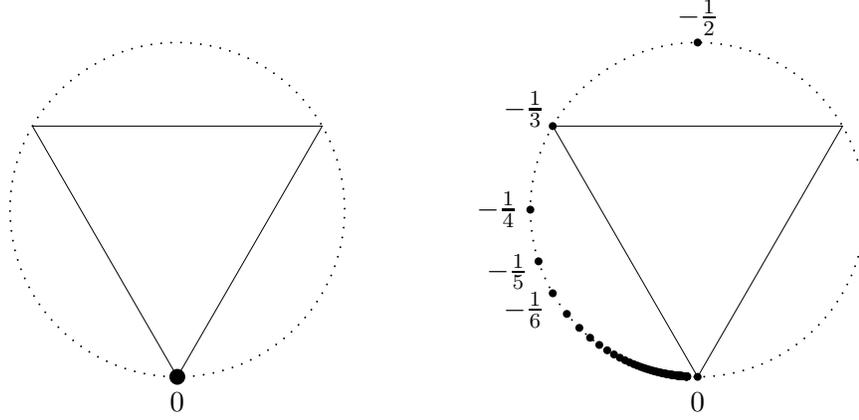
FIGURE 7. Examples of two-level measures K on the circle such that $M_K \neq \lambda_{\mathbb{S}}$. On the left, $K = \delta_{\delta_0}$ and on the right, we represent $K_n = \delta_{\frac{1}{2}\delta_0 + \frac{1}{2}\delta_{n-1}}$ for several $n \in \mathbb{N}$. Then K_n converges weakly to K. If we add the atom on endpoints to the mass of the circular segment after the endpoint, the tree associated with (C, K_n) has two leaves carrying mass but the one associated with (C, K) is a sinle point carrying mass. Therefore the sequence of trees corresponding to (C, K_n) do not converge to the one associated with (C, K).

which is the set of mid-points of "filled" circular segments for C (see Figure 5).

In the tree, the corresponding leaves will be connected to the rest of the tree through line segments with non-atomic measure. For (T, c, ν) an algebraic measure tree, we call an interval and denote it by (v, w) a subset of $\operatorname{supp}(M_\nu)$ such that $(v, w) \cap \operatorname{at}(M_\nu) = \emptyset$ and $(v, w) \cap \operatorname{br}(T, c) = \emptyset$ a *line segment with non-atomic measure*. We denote by

$$(4.11) \qquad \operatorname{seg}(T, c, \nu)$$

the set of maximal (w.r.t. inclusion) line segments with non-atomic measure. Note that if (T, c) is order complete, every line segment with non-atomic measure is included in some $(v, w) \in \operatorname{seg}(T, c, \nu)$.

The analog of these line segments in the sub-triangulation are "filled" areas, that is, elements of

$$(4.12) \qquad \blacksquare(C) := \left\{ \overline{b} \ : \ b \text{ is a connected component of } C^\circ \right\}.$$

For such a filled area $b \in \blacksquare(C)$, it will be important to be able to know the mass distribution of K along $b \cap \mathbb{S}$, and not only the total mass carried by $b \cap \mathbb{S}$. For this reason, we use that by definition of a sub-triangulation, $b \in \blacksquare(C)$ is the union of non-crossing closed straight line segments with endpoints in $\mathbb{S}$. However, there are several possible ways to decompose b into infinitely many line segments. For the sake of simplicity, we choose to partition b "linearly" in the following sense. Recall that we identify $\mathbb{S}$ with $[0, 1]$ where the endpoints are glued. If $b \cap \mathbb{S}$ is not connected, then it is the union of two circular segments $[x, y]$ and $[y', x']$ (with $x = y$ or $x' = y'$ possibly, see Figure 8, left). If $b \cap \mathbb{S}$ is connected, it is a circular segment $[x, x']$ and we take $y = y'$ to be its mid-point (which belongs to $\square(C)$), so that $b \cap \mathbb{S}$ is still the union of two circular segments (see Figure 8, right). We then define $\|_b(C)$ as the set of all straight line segments connecting a point of $[x, y]$ to a point of $[y', x']$ in a linear way (see Figure 8). For
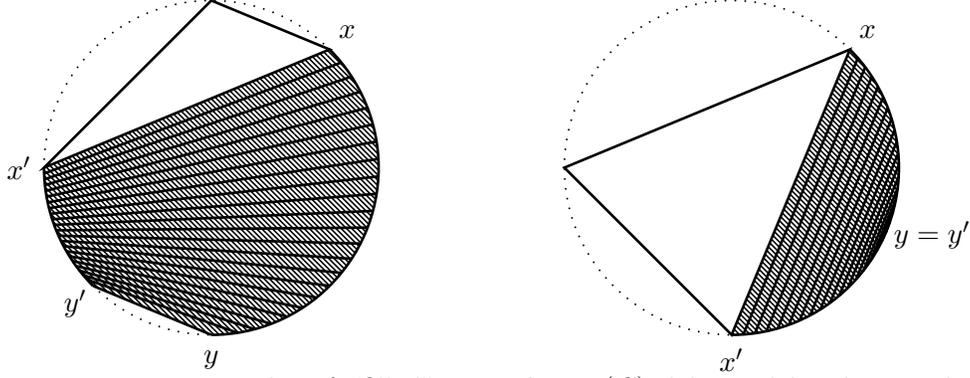
FIGURE 8. Two examples of "filled" areas $b \in \blacksquare(C)$ delimited by the circular segments $[x, y]$ and $[x', y']$. They are partitioned into straight line segments with endpoints in $[x, y]$ and $[x', y']$.

example, if $x \le y \le y' \le x'$,

$$\|_b(C) := \left\{ [(1-t)x + ty, (1-t)x' + ty'] \mid t \in [0, 1] \right\}. \tag{4.13}$$

Define

$$\|(C) := \bigcup_{b \in \blacksquare(C)} \|_b(C). \tag{4.14}$$

Finally, we define for the sub-triangulation the analog of the components in an algebraic tree. For $x \in \Delta(C) \cup \nabla(C)$, and $y \subseteq \mathbb{D}$ connected and disjoint from ∂x, let

$$\text{comp}_x(y) := \text{the connected component of } \mathbb{D} \setminus \partial x \text{ which contains } y. \tag{4.15}$$

Similarly, for $x \in \|(C)$ and $y \subseteq \mathbb{D}$ connected and disjoint from x, let

$$\text{comp}_x(y) := \text{the connected component of } \mathbb{D} \setminus x \text{ which contains } y. \tag{4.16}$$

We also call them *components* (of the sub-triangulation) (see Figure 9). Define also $\text{comp}_x(x) := x$ for $x \in \|(C) \cup \square(C)$.

The following lemma gives the tree associated to a triangulation C. Contrary to [LW21, Lemma 4.7], we add the set $\|(C)$ to the "skeleton" of the tree as it is important in the two-level case to know how the mass is distributed along a line segment of the tree. For (T, c) an algebraic tree and $x \in T$, we first extend the notation (3.4) for components to

$$\mathcal{S}_x(x) := \{x\}. \tag{4.17}$$

Lemma 4.6 (induced branch-point map)**.** *For $C \in \mathcal{T}$, let $V_C := \Delta(C) \cup \nabla(C) \cup \square(C) \cup \|(C)$. If $V_C \ne \emptyset$, then there exists a unique branch-point map $c_V : V_C^3 \to V_C$, such that (V_C, c_V) is an algebraic tree with*

$$\mathcal{S}_x^{(V_C, c_V)}(y) = \{v \in V_C : \text{comp}_x(y) = \text{comp}_x(v)\} \tag{4.18}$$

for $x, y \in V_C$. In particular, $\deg(x) = 3$ for all $x \in \Delta(C)$, $\deg(x) = 1$ for all $x \in \nabla(C) \cup \square(C)$ and $\deg(x) = 2$ for all $x \in \|(C) \setminus \square(C)$.
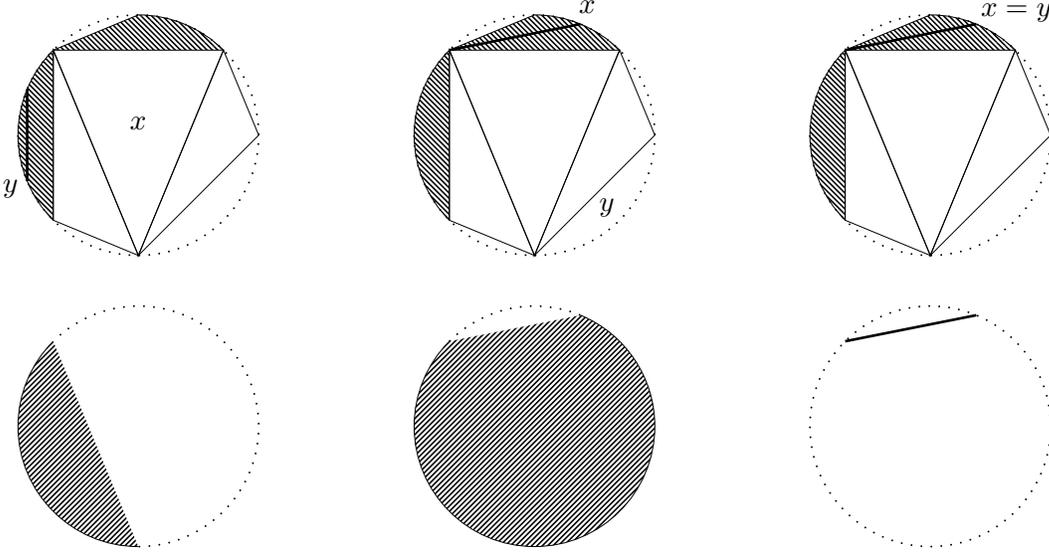
FIGURE 9. Three examples of components $\mathrm{comp}_x(y)$ in a sub-triangulation of the circle C. The top line represents the triangulation, x and y. The second line represents the corresponding components $\mathrm{comp}_x(y)$. On the left, $x \in \Delta(C)$ and $y \in \|(C)$. In the middle, $x \in \|(C)$ and $y \in \nabla(C)$. In these two cases, $\mathrm{comp}_x(y)$ is a filled cap region of the disc with the boundary straight line excluded and the rest of the boundary arc included. On the right, $x = y \in \|(C)$ and $\mathrm{comp}_x(x) = x$.

Proof. Condition (Tri2)' of Proposition 4.2 gives that for pairwise distinct $x, y, z \in \Delta(C) \cup \nabla(C)$, there is a unique triangle $c_{xyz} \in \Delta(C)$ such that x, y, z are subsets of pairwise different connected components of $\mathbb{D} \setminus \partial c_{xyz}$. This can be extended to triples of points in V_C in an obvious way. It is then easy to see that this defines a branch point map on V_C. $\square$

The following theorem is the analog of [LW21, Theorem 2]. It states that each pair (C, K), where C is a sub-triangulation of the circle and K a two-level measure on $\mathbb{S}$ with $M_K = \lambda_{\mathbb{S}}$, can be associated with a binary a2m tree such that $\Delta(C)$ corresponds to the set of branch points and $\nabla(C)$ corresponds to the set of atoms of the intensity measure of ν in the tree. Furthermore, $\mathrm{comp}_v(w)$ corresponds to the component $\mathcal{S}_v(w)$ and its random ν-mass is given by the random K-mass carried by $\mathrm{comp}_v(w) \cap \mathbb{S}$.

Theorem 1 (Coding map).   (i) *For all* $\Gamma = (C, K) \in \mathfrak{D}$, *there is a unique (up to equivalence) a2m tree* $\chi_\Gamma = (T_\Gamma, c_\Gamma, \nu_\Gamma) \in \mathbb{T}_2^{(2)}$ *such that:*
   (CM1) $V_C \subseteq T_\Gamma$, $\mathrm{br}(T_\Gamma, c_\Gamma) = \Delta(C)$, *and* c_Γ *is an extension of* c_V, *where* (V_C, c_V) *is defined in Lemma 4.6.*
   (CM2) *For all* $x, y \in V_C$ *and* $\sigma \geq 0$,

$$(4.19) \qquad \int \nu_\Gamma(\mathrm{d}\mu) e^{-\sigma \mu(\mathcal{S}_x(y))} = \int K(\mathrm{d}\kappa) e^{-\sigma \kappa(\mathrm{comp}_x(y) \cap \mathbb{S})}.$$

   (CM3) $\mathrm{at}(M_{\nu_\Gamma}) = \nabla(C)$.

    (ii) *The coding map $\tau : \mathfrak{D} \to \mathbb{T}_2^{(2)}$, $\Gamma \mapsto \chi_\Gamma$ is surjective.*

    (iii) *Let $\mathcal{T}$ be equipped with the Hausdorff metric topology, $\mathcal{M}_1(\mathcal{M}_1(\mathbb{S}))$ with the weak topology, $\mathcal{T} \times \mathcal{M}_1(\mathcal{M}_1(\mathbb{S}))$ with the product topology and $\mathbb{T}_2^{(2)}$ with the two-level bpdd-Gromov-weak topology. Then the coding map τ is continuous.*

We first state an extension result from [LW21] for algebraic measure trees that will be a key ingredient in the proof. Recall that for (T,c) an algebraic tree and $x \in T$, we have $\mathcal{S}_x(x) := \{x\}$ so that T is the disjoint union of the $\deg(x) + 1$ sets in

$$(4.20) \qquad\qquad \mathcal{C}_x := \{\mathcal{S}_x(y) \ : \ y \in T\}.$$

For $y \in T$ and $V \subseteq T$, we call a function $f \colon V \to \mathbb{R}$ *order-left continuous* on V with respect to the partial order $\leq_y$ (defined by (3.3)) if for all $x, x_n \in V$ such that $x_n \uparrow_y x$, we have $\lim_{n\to\infty} f(x_n) = f(x)$, where $x_n \uparrow_y x$ means that $x_1 \leq_y x_2 \leq_y \cdots$ and $x = \sup_{n\in\mathbb{N}} x_n$ with respect to $\leq_y$.

From [LW21, Proposition 2.20], we know that if an algebraic tree (T,c) is order separable, there exists a subset V of T that is *order dense*, i.e., such that for all $x, y \in T$ with $x \neq y$,

$$(4.21) \qquad\qquad D \cap [x,y) \neq \emptyset.$$

The following result [LW21, Proposition 3.12] allows to extend the function that associates to subtrees their masses to a measure on the algebraic tree.

Proposition 4.7 (Extension to a measure). *Let (T,c) be an order separable algebraic continuum tree, and $V \subseteq T$ order dense. Then a set-function $\mu_0 : \mathcal{C}_V := \bigcup_{x\in V} \mathcal{C}_x \to [0,1]$ has a unique extension to a probability measure on $\mathcal{B}(T,c)$ if it satisfies*

    (1) *For all $x \in V$, $\sum_{A\in\mathcal{C}_x} \mu_0(A) = 1$.*

    (2) *For all $x, y \in V$ with $x \neq y$,*

$$(4.22) \qquad\qquad \mu_0(\mathcal{S}_x(y)) + \mu_0(\mathcal{S}_y(x)) \geq 1.$$

    (3) *For every $y \in V$, the function $\psi_y : x \mapsto \mu_0(\mathcal{S}_x(y))$ is order left-continuous on V with respect to $\leq_y$.*

Let us now sketch the proof of Theorem 1 which is proved below. (i) We start by extending the tree (V_c, c_V) to an algebraic continuum tree and apply Proposition 4.7 "almost surely" to construct the two-level measure ν. (ii) We give the proof of the surjectivity of the coding map τ in three steps. For $\chi = (T, c, \nu) \in \mathbb{T}_2^{(2)}$,

    (1) we first construct the sub-triangulation C associated with the algebraic measure tree (T, c, M_ν) as in the proof of [LW21, Theorem 2]. That is, for each branch point we remove an open triangle from the set $\mathbb{D}$ and for each atom of M_ν, we remove a circular segment.

    (2) we then build K by weak approximation. For this, we rely on a result from [Mei19] that the two-level measure ν can be reconstructed from an infinite set of points randomly sampled. The two-level measure K is then built in a similar fashion by using a correspondance between points in the sample and subsets of the circle.

    (3) finally, we prove that the constructed pair (C, K) is in $\mathfrak{D}$ by partitioning the circle line into different types of intervals and showing that the restriction of M_K to each interval is the Lebesgue measure.

(iii) The proof of the continuity of τ is defered to the next section, Proposition 5.8.

Proof of Theorem 1. (i) Let $\Gamma = (C, K) \in \mathfrak{D}$. For $\kappa \in \text{supp}(K)$, we define the set-function $\mu_0^\kappa : \bigcup_{x \in V_C} \mathcal{C}_x \to [0, 1]$ by

$$(4.23) \qquad \mu_0^\kappa(\mathcal{S}_x(y)) := \kappa(\text{comp}_x(y) \cap \mathbb{S}), \quad x, y \in V_C.$$

We want to be in a position to apply Proposition 4.7 to each μ_0^κ. Since $M_K = \lambda_\mathbb{S}$, for K-almost every κ, for all $x \in V_C$, $\kappa(\partial x \cap \mathbb{S}) = 0$. Thus, for K-almost every κ and all $x \in V_C$, $\sum_{A \in \mathcal{C}_x} \mu_0^\kappa(A) = \kappa(\mathbb{S}) = 1$ and $\mu_0^\kappa(\mathcal{S}_x(y)) + \mu_0^\kappa(\mathcal{S}_y(x)) \geq 1$ for $y \neq x$. Furthermore, the function $x \mapsto \mu_0^\kappa(\mathcal{S}_x(y))$ is order left-continuous on V_C w.r.t. $\leq_y$ for all $y \in V_C$ for K-almost every κ. This can be shown by using (4.23) and the continuity from above of the measure κ, together with the condition $M_K = \lambda_\mathbb{S}$.

The tree (V_C, c_V) is order separable. Indeed, the set $\Delta(C) \cup \nabla(C)$ is countable and one can construct a countable order dense set of (V_C, c_V) by adding the straight line segments in $\|(C)$ that have at least one rational endpoint in $\mathbb{S} \simeq [0, 1]$. However, (V_C, c_V) is not necessarily order complete. Since we defined $\blacksquare(C)$ to be the set of all closures of connected components of C°, the endpoints of line segments are included in (V_C, c_V). But we still need to add the leaves which are limits of an increasing sequence (w.r.t $\leq_y$ for some $y \in V_C$) of branch points. Thus we define $\overline{V}_C$ as V_C together with an uncountable set of leaves given by the limit points of these increasing sequences, and we can easily extend c_V to $\overline{c}_V$ so that the tree $(\overline{V}_C, \overline{c}_V)$ is order complete and order separable.

Furthermore, $(\overline{V}_C, \overline{c}_V)$ need not to be an algebraic continuum tree because $\text{edge}(\overline{V}_C, \overline{c}_V)$ might be non-empty. Therefore, to apply Proposition 4.7 we also extend the tree $(\overline{V}_C, \overline{c}_V)$ to make it an order separable algebraic continuum tree. Let $\{v, w\}$ be an edge of $(\overline{V}_C, \overline{c}_V)$. Define $\widetilde{V} := \overline{V}_C \uplus (\{(v, w)\} \times (0, 1))$. We can extend $\overline{c}_V$ to a branch point map $\widetilde{c}$ on $\widetilde{V}$ in a canonical way such that $[v, w] = \{v, w\} \cup (\{(v, w)\} \times (0, 1))$ (see Figure 10). $\overline{V}_C$ is not anymore order dense in $(\widetilde{V}, \widetilde{c})$ so we also define, for $\kappa \in \text{supp}(K)$, μ_0^κ on any $\mathcal{C}_x$ with $x \in \{(v, w)\} \times (0, 1)$. For $x \in \{(v, w)\} \times (0, 1)$, let $\mu_0^\kappa(\{x\}) = 0$ and for $y \neq x$,

$$(4.24) \qquad \mu_0^\kappa(\mathcal{S}_x(y)) := \begin{cases} \mu_0^\kappa(\mathcal{S}_x(v)) & \text{if } v \in \mathcal{S}_x(y), \\ \mu_0^\kappa(\mathcal{S}_x(w)) & \text{if } w \in \mathcal{S}_x(y). \end{cases}$$

Then for K-almost every κ, μ_0^κ still satisfies the conditions (1)-(3) of Proposition 4.7 on $(\widetilde{V}, \widetilde{c})$. We repeat this extension for each edge in $\text{edge}(\overline{V}_C, \overline{c}_V)$ which is countable as the set $\Delta(C) \cup \nabla(C)$ is itself countable. We denote by (T, c) the extended tree. By construction, it is still order complete and order separable, and $\text{edge}(T, c) = \emptyset$. Since μ_0^κ was extended to $\bigcup_{x \in T} \mathcal{C}_x$ and T is obviously dense in itself, we can now apply Proposition 4.7. That is, for K-almost every κ, μ_0^κ has a unique extension to a probability measure μ^κ on $\mathcal{B}(T, c)$.

Define $\nu \in \mathcal{M}_1(\mathcal{M}_1(T, c))$ as the pushforward of K under the map $\kappa \mapsto \mu^\kappa$. The triplet (T, c, ν) will be the triplet $(T_\Gamma, c_\Gamma, \nu_{G}amma)$ announced in the point (i) of Theorem 1, but we do not write the subscripts Γ for the sake of notation. Let us verify the a2m tree (T, c, ν) satisfies conditions (CM1)-(CM3). It is easy to see that (CM1) holds, and by (4.23), for all $x, y \in V_C$, and $\sigma > 0$,

$$(4.25) \qquad \int \nu(\mathrm{d}\mu) e^{-\sigma \mu(\mathcal{S}_x(y))} = \int K(\mathrm{d}\kappa) e^{-\sigma \mu^\kappa(\mathcal{S}_x(y))} = \int K(\mathrm{d}\kappa) e^{-\sigma \kappa(\text{comp}_x(y) \cap \mathbb{S})},$$
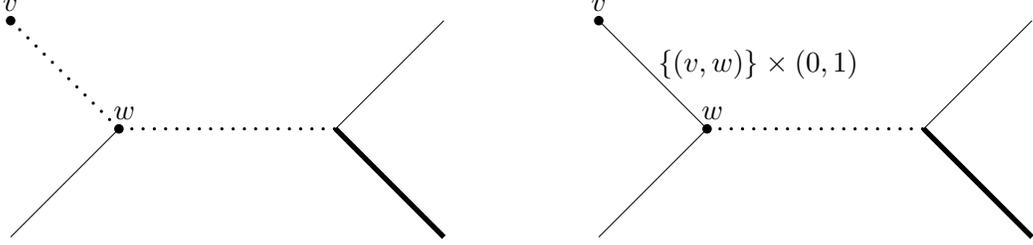
FIGURE 10. The tree is not a continuum tree because $[v, w] = \{v, w\}$ is an edge. We extend the tree in a natural way so that $[v, w]$ is no longer an edge.

which is (CM2). Moreover, points in $T \setminus V_C$ carry no atoms, so $\mathrm{at}(M_\nu) \subseteq V_C$. With (CM2), for each $x \in V_C$,

$$(4.26) \qquad M_\nu\{x\} = M_K(\mathrm{comp}_x(x) \cap \mathbb{S}) = \lambda(\mathrm{comp}_x(x) \cap \mathbb{S}),$$

which is zero if $x \in \Delta(C) \cup \square(C)$ and strictly positive if $x \in \nabla(C)$, which yields (CM3). To see that it is unique (up to equivalence), notice that in the construction above, the extensions of edges of $(\overline{V}_C, \overline{c}_V)$ carry no mass and that the extension of μ_0^κ to μ^κ is unique for K-almost every κ by Proposition 4.7.

(ii) We give the proof of the surjectivity in three steps. Let $\chi = (T, c, \nu) \in \mathbb{T}_2^{(2)}$. We construct a sub-triangulation C and a two-level measure K on $\mathbb{S}$ such that $\tau(C, K) = \chi$. We can suppose w.l.o.g. that (T, c) is order complete and that for all $v \in \mathrm{br}(T, c)$, $c_*(M_\nu^{\otimes 3})\{v\} > 0$. To make the construction of K easier, we also assume that for all $(v, w) \in \mathrm{seg}(T, c, \nu)$, M_ν restricted to (v, w) is the Lebesgue measure (see Example 3.11).

Step 1: construction of C. Fix $\rho \in \mathrm{lf}(T, c)$ and recall that ρ induces a partial order relation $\leq_\rho$. We can extend this partial order to a total order $\leq$ in the following way. For $v \in \mathrm{br}(T, c)$, denote by $S_0(v)$, $S_1(v)$ and $S_2(v)$ the three components of $T \setminus \{v\}$ such that $S_0(v) = \mathcal{S}_v(\rho)$ and with $S_1(v)$ and $S_2(v)$ chosen by picking an order for the two remaining components. Define now for all v, w,

$$(4.27) \qquad v \leq w \quad \Leftrightarrow \quad \big[v \leq_\rho w \big] \text{ or } \big[v \in S_1(c(v, w, \rho)) \text{ and } w \in S_2(c(v, w, \rho)) \big].$$

Using this total order on the tree, the construction of C below can be understood as follows. Starting from the root ρ which corresponds to 0 in the identification $\mathbb{S} \simeq [0, 1]$, we read through T according to $\leq$. For each branch point v, we draw a triangle whose vertices on $\mathbb{S}$ are given by the M_ν-mass of points smaller than v and the M_ν-masses of the two components above v. For each leaf carrying an atom w of M_ν, we add a straight line segment to C according to the M_ν-mass of points smaller than w and the M_ν-mass of w (see Figure 11).

We first introduce some more notations. For $a \in [0, 1]$ and $b, c > 0$ with $a + b + c \leq 1$, let $\Delta(a, b, c) \subseteq \mathbb{D}$ be the open triangle with vertices $a, a+b, a+b+c \in \mathbb{S} \simeq [0, 1]$, and let $\ell(a, b) \subseteq \mathbb{D}$ be the straight line from a to $a+b$, and $L(a, b)$ the connected component of $\mathbb{D} \setminus \ell(a, b)$ containing $a + \frac{b}{2} \in \mathbb{S}$.
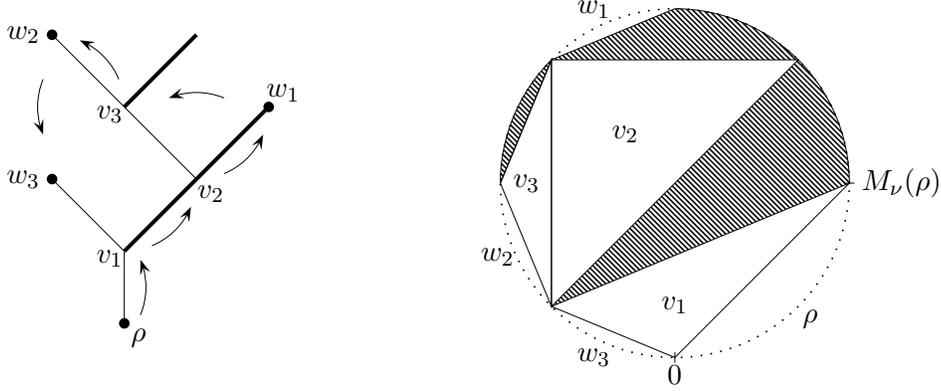
FIGURE 11. An algebraic tree (T, c) and the corresponding sub-triangulation C as constructed in Step 1 of the proof of Theorem 1(ii). v_1, v_2, v_3 are branch points and ρ, w_1, w_2, w_3 are leaves carrying an M_ν-atom. More precisely, $M_\nu(\rho) = \frac{1}{4}$ and $M_\nu(w_1) = M_\nu(w_2) = M_\nu(w_3) = \frac{1}{8}$. We extend the partial order $\leq_\rho$ to a total order such that $\rho \leq v_1 \leq v_2 \leq w_1 \leq v_3 \leq w_2 \leq w_3$.

Now for all $v \in \mathrm{br}(T, c) \cup \mathrm{at}(M_\nu)$, denote by

$$(4.28) \qquad \alpha(v) := M_\nu\left(\{u \in T : u < v\}\right)$$

the total M_ν-mass of points smaller than v (with respect to $\leq$ defined in (4.27)). In the sub-triangulation C defined below, $\alpha(v)$ gives the first vertex of the triangle or circular segment corresponding to v. We define C by

$$(4.29) \qquad \mathbb{D} \setminus C := \biguplus_{v \in \mathrm{br}(T,c)} \Delta\big(\alpha(v), M_\nu(S_1(v)), M_\nu(S_2(v))\big) \uplus \biguplus_{w \in \mathrm{at}(M_\nu)} L\big(\alpha(w), M_\nu\{w\}\big).$$

Let us show that C is a sub-triangulation of the circle. By definition of C, $\mathrm{conv}(C) \setminus C$ is the disjoint union of open triangles, i.e. condition (Tri1) is satisfied. Furthermore, the extreme points of $\mathrm{conv}(C)$ are contained in $\mathbb{S}$, and for $x, y, z \in \Delta(C) \cup \nabla(C)$ distinct, there are corresponding $u, v, w \in T$, and a triangle $c_{xyz} \in \Delta(C)$ corresponding to $c(u, v, w)$, which satisfies the requirements of (Tri2)'. Thus, by Proposition 4.2, $C \in \mathcal{T}$. Note that C is the sub-triangulation associated with the algebraic measure tree (T, c, M_ν) in the proof of Theorem 4.8 in [LW21].

To make the notations coincide with the ones from Lemma 4.6, we will use the following correspondences in the rest of the proof. That is, a branch point $v \in \mathrm{br}(T, c)$ will correspond to the triangle $\overline{v} = \Delta\big(\alpha(v), M_\nu(S_1(v)), M_\nu(S_2(v))\big)$ in $\Delta(C)$ and an atom $w \in \mathrm{at}(M_\nu)$ to the circular segment $\overline{w} = L\big(\alpha(w), M_\nu\{w\}\big)$ in $\nabla(C)$. A leaf with zero mass $u \in \mathrm{lf}(T, c)$, but connected to the rest of the tree by a line segment carrying mass, will correspond to the midpoint $\overline{u} \in \square(C)$ of the corresponding "filled" circular segment. Finally, we associate, in a "linear" way, a point $z \in T$ belonging to a line segment with non-atomic mass in $\mathrm{seg}(T, c, \nu)$ to a straight line segment $\overline{z} \in \|(C)$. That is, if the segment $(v, w) \in \mathrm{seg}(T, c, \nu)$ corresponds to the circular segments $]x, y[$ and $]x', y'[$, the point $y = (1 - t)v + tw$ with $t \in (0, 1)$ is associated with the straight line segment $\overline{y} := [(1 - t)x + ty, (1 - t)x' + ty']$ (see (4.13)). Denote by

$$(4.30) \qquad V_T \subseteq T$$

the set of all points x in the tree that have a corresponding subset $\overline{x} \in V_C$.

With these notations, V_T is a subtree of (T, c) and by construction, we have that for all $x, y \in V_T$,

$$(4.31) \qquad\qquad M_\nu(\mathcal{S}_x(y)) = \lambda_\mathbb{S}(\mathrm{comp}_{\overline{x}}(\overline{y}) \cap \mathbb{S}).$$

Let $b \in \blacksquare(C)$ such that $b \cap \square(C) = \emptyset$, that is, b corresponds to a segment line of the tree carrying mass and that is not connected to a leaf without mass. In this case, $b \cap \mathbb{S}$ is the union of two disjoint intervals (see Figure 8, left). It is important for the next step of the proof to note that from the construction of C, one (and only one) of the two intervals is a singleton. Thus, the random K-mass we will assign to b needs to be carried by the interval with non-empty interior to ensure that $M_K = \lambda_\mathbb{S}$.

Step 2: construction of K. Let $(x_{ij})_{(i,j)}$ be a random infinite matrix with distribution

$$(4.32) \qquad\qquad \int \nu^{\otimes\infty}(\mathrm{d}\underline{\mu}) \int \bigotimes_{i \geq 1} \mu_i^{\otimes\infty}(\cdot).$$

We assumed the tree (T, c) to be order separable and order complete. Thus the component topology on (T, c) is Polish (see Remark 3.5) and we know from [Mei19, Proposition 3.7] that, almost surely,

(i) for every $i \in \mathbb{N}$, the weak limit $\mu_i := \underset{n\to\infty}{\text{w-lim}}\frac{1}{n}\sum_{j=1}^n \delta_{x_{ij}}$ exists and has law ν,

(ii) the two-level measure $\frac{1}{m}\sum_{i=1}^m \delta_{\mu_i}$ converges weakly to ν, and

(iii) for every $j \in \mathbb{N}$, the sequence $(x_{ij})_i$ is dense in $\mathrm{supp}(M_\nu)$.

Fix $i, n \in \mathbb{N}$. We write $\mu_i^n := \frac{1}{n}\sum_{j=1}^n \delta_{x_{ij}}$ and we first define κ_i^n a probability measure on the circle line corresponding to μ_i^n. For each x_{ij}, we put weight $\frac{1}{n}$ on a subset of the circle depending on the position of x_{ij} in (T, c). We will make use of the correspondence between points of $V_T \subseteq T$ and subsets of $\mathbb{D}$ discussed at the end of Step 1. Since $(T, c, \nu) \in \mathbb{T}_2^{(2)}$, $\mathrm{at}(\mu) \subseteq \mathrm{lf}(T, c)$ for all $\mu \in \mathrm{supp}(\nu)$ so we have the following possible cases (see Figure 12) that correspond to a partition of T in five sets $A_1, \cdots A_5$:

(1) Set A_1: if $x_{ij} \in \mathrm{at}(M_\nu)$, $\overline{x_{ij}}$ is the circular segment $L(\alpha(x_{ij}), M_\nu\{x_{ij}\})$ and we will put the rescaled Lebesgue measure on the arc $\overline{x_{ij}} \cap \mathbb{S}$.

(2) Suppose now that x_{ij} belongs to a line segment with non-atomic mass, or equivalently that $\deg(x_{ij}) = 2$. Since (T, c) is order complete, there exist $v, w \in T$ such that $(v, w) \in \mathrm{seg}(T, c, \nu)$ and $x_{ij} \in (v, w)$.

    (a) Set A_2: if (v, w) is adjacent to a leaf, i.e. $\overline{x_{ij}}$ is included in a filled circular segment, then we put weight $\frac{1}{2n}$ on the endpoints $\widetilde{x_{ij1}}$ and $\widetilde{x_{ij2}}$ of the line segment $\overline{x_{ij}}$.

    (b) Set A_3: if not, then $\overline{x_{ij}} \subset b$ where $b \in \blacksquare(C)$ is such that $b \cap \mathbb{S}$ is the union of an interval and a singleton. In this case, we put weight $\frac{1}{n}$ on the endpoint $\widetilde{x_{ij}}$ of the line segment $\overline{x_{ij}}$ that is not the singleton.

(3) Finally, suppose that $x_{ij} \in \mathrm{lf}(T, c) \setminus \mathrm{at}(M_\nu)$.

    (a) Set A_4: if x_{ij} is adjacent to a line segment with non-atomic mass, then $\overline{x_{ij}} \in \square(C)$ and we put weight $\frac{1}{n}$ on $\widetilde{x_{ij}} = \overline{x_{ij}}$.

    (b) Set A_5: if not, $x_{ij} \notin V_T$, i.e. $\overline{x_{ij}}$ is not defined. However, it naturally corresponds to a specific point of the circle. To see this, let $y \in \mathrm{br}(T, c)$. Then there exists $(z_n)_{n\in\mathbb{N}}$ a sequence of distinct branch points in T such that $z_n \leq_y z_{n+1} \leq_y x_{ij}$ for all $n \in \mathbb{N}$.
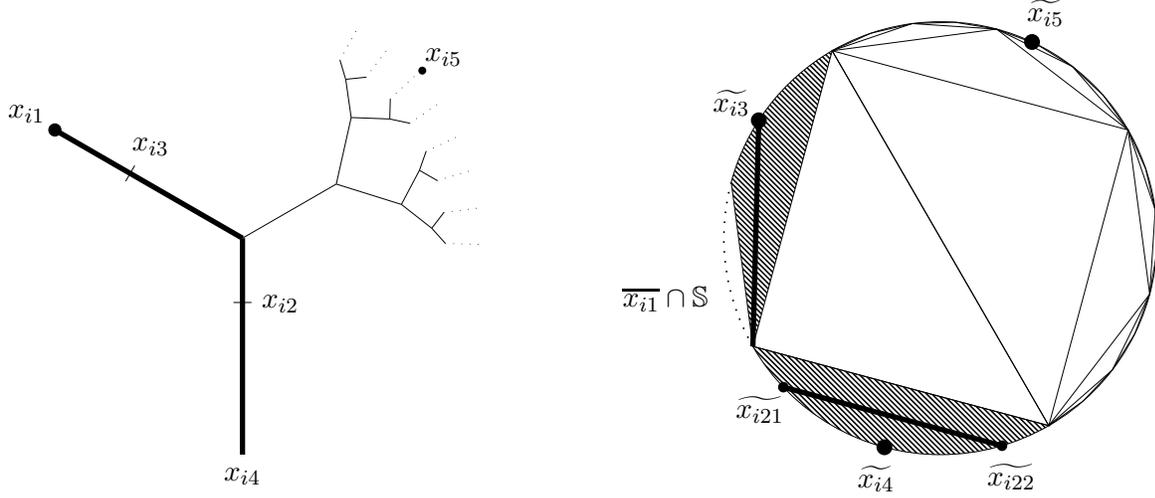
FIGURE 12. For each atom $x_{i1}, ..., x_{i5}$ of μ_i^n, κ_i^n assigns weight $\frac{1}{n}$ on a subset of the circle, which depends on the position of the atom in the tree. For the M_ν-atom leaf x_{i1}, κ_i^n assigns the rescaled Lebesgue measure on the corresponding circular segment $\overline{x_{i1}} \cap \mathbb{S}$. For x_{i2}, κ_i^n splits the mass $\frac{1}{n}$ between the endpoints $\widetilde{x_{i21}}$ and $\widetilde{x_{i22}}$ of the segment $\overline{x_{i2}}$. For the three other ones, $\widetilde{x_k}$ carries weight $\frac{1}{n}$.

For all $n \in \mathbb{N}$, let $a_n, b_n \in \mathbb{S} \simeq [0,1]$ such that $\mathrm{comp}_{\overline{z_n}}(\overline{z_{n+1}}) \cap \mathbb{S} = [a_n, b_n]$. We have $[a_{n+1}, b_{n+1}] \subseteq [a_n, b_n]$. By (4.31), $\lambda_{\mathbb{S}}([a_n, b_n]) = M_\nu(\mathcal{S}_{z_n}(z_{n+1})) = M_\nu(\mathcal{S}_{z_n}(x_{ij}))$. But since $x_{ij} \notin \mathrm{at}(M_\nu)$ and x_{ij} is not adjacent to a line segment with non-atomic mass, $\lim_{n \to \infty} M_\nu(\mathcal{S}_{z_n}(x_{ij})) = 0$. Therefore, there exists $\widetilde{x_{ij}} \in \mathbb{S}$ unique such that $\widetilde{x_{ij}} = \bigcap_{n \in \mathbb{N}}[a_n, b_n]$. We put weight $\frac{1}{n}$ on $\widetilde{x_{ij}}$.

Define the random probability measure κ_i^n on $\mathbb{S}$ by

$$(4.33) \qquad \kappa_i^n := \sum_{x_{ij} \in A_1} \frac{\lambda_{\mathbb{S}}(\cdot \cap \overline{x_{ij}})}{n\lambda_{\mathbb{S}}(\mathbb{S} \cap \overline{x_{ij}})} + \sum_{x_{ij} \in A_2} \frac{\delta_{\widetilde{x_{ij1}}}(\cdot) + \delta_{\widetilde{x_{ij2}}}(\cdot)}{2n} + \sum_{x_{ij} \in A_3 \cup A_4 \cup A_5} \frac{\delta_{\widetilde{x_{ij}}}(\cdot)}{n}.$$

Now for each $i, n \in \mathbb{N}$, κ_i^n is a random variable in $\mathcal{M}_1(\mathbb{S})$ which is compact when equipped with the weak topology. Therefore, the sequence $(\kappa_i^n)_n$ is tight and has a subsequence $(\kappa_i^{\varphi_i(n)})_n$ that converges to some random measure κ_i, for each $i \in \mathbb{N}$. Before taking the weak limit of the empirical distribution of the sequence $(\kappa_i)_i$, note first that a.s. μ_i and κ_i agree on (union of) corresponding components on the tree T and on the circle line $\mathbb{S}$. Indeed, by construction, a.s. for all $x, y \in V_T$,

$$(4.34) \qquad \kappa_i^n(\mathrm{comp}_{\overline{x}}(\overline{y}) \cap \mathbb{S}) = \mu_i^n(\mathcal{S}_x(y)).$$

Furthermore, consider $x, y, x', y' \in V_C$. Then $\mathrm{comp}_x(y)$ and $\mathrm{comp}_{x'}(y')$ are two circular segments delimited by two straight line segments that are disjoint or equal. Thus $\mathrm{comp}_x(y) \cup \mathrm{comp}_{x'}(y')$ is either disjoint, the whole circle $\mathbb{S}$ or one component if one is included in the other (see Figure 13).
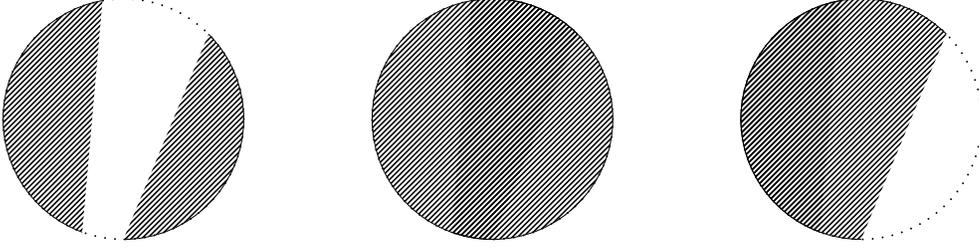
FIGURE 13. Thus union of two components is either disjoint (*left*), the whole circle $\mathbb{S}$ (*middle*) or one component if one is included in the other (*right*).

Therefore, we have a.s. for all $x_k, y_k \in V_T$,

$$(4.35) \qquad \kappa_i^n \left(\bigcup_{k \in \mathbb{N}} \mathrm{comp}_{\overline{x_k}}(\overline{y_k}) \cap \mathbb{S} \right) = \mu_i^n \left(\bigcup_{k \in \mathbb{N}} \mathcal{S}_{x_k}(y_k) \right).$$

Then we can generalize it by saying that for n large enough, a.s. for all $x_k, y_k \in V_C$,

$$(4.36) \qquad (\kappa_1^n, ..., \kappa_m^n) \left(\bigcup_{k \in \mathbb{N}} \mathrm{comp}_{\overline{x_k}}(\overline{y_k}) \cap \mathbb{S} \right) = (\mu_1^n, ..., \mu_m^n) \left(\bigcup_{k \in \mathbb{N}} \mathcal{S}_{x_k}(y_k) \right).$$

Taking the limit of the subsequences $(\kappa_i^{\varphi_i(n)})_n$ when n goes to infinity for $1 \le i \le m$,

$$(4.37) \qquad (\kappa_1, ..., \kappa_m) \left(\bigcup_{k \in \mathbb{N}} \mathrm{comp}_{\overline{x_k}}(\overline{y_k}) \cap \mathbb{S} \right) \quad \text{and} \quad (\mu_1, ..., \mu_m) \left(\bigcup_{k \in \mathbb{N}} \mathcal{S}_{x_k}(y_k) \right).$$

have same distribution for all $x_k, y_k \in V_T$.

We now denote by K_m the empirical distribution of the sequence $(\kappa_i)_i$, that is,

$$(4.38) \qquad K_m := \frac{1}{m} \sum_{i=1}^m \delta_{\kappa_i},$$

which is a random variable in $\mathcal{M}_1(\mathcal{M}_1(\mathbb{S}))$. Since $\mathcal{M}_1(\mathcal{M}_1(\mathbb{S}))$ is also compact, the sequence $(K_m)_m$ is tight, so that there exists a subsequence $(K_{\varphi(m)})_m$ that converges to some random two-level measure K. For all $m \in \mathbb{N}$, $x_k, y_k \in V_T$, and $\sigma \ge 0$,

$$(4.39) \qquad \frac{1}{m} \sum_{i=1}^m e^{-\sigma \kappa_i \left(\bigcup_{k \in \mathbb{N}} \mathrm{comp}_{\overline{x_k}}(\overline{y_k}) \cap \mathbb{S} \right)} \quad \text{and} \quad \frac{1}{m} \sum_{i=1}^m e^{-\sigma \mu_i \left(\bigcup_{k \in \mathbb{N}} \mathcal{S}_{x_k}(y_k) \right)}$$

have same distribution. Recall that the two-level measure $\frac{1}{m} \sum_{i=1}^m \delta_{\mu_i}$ converges weakly to ν. Thus, taking the limit of the subsequence $(K_{\varphi(m)})_m$, for all $x_k, y_k \in V_T$, and $\sigma \ge 0$,

$$(4.40) \qquad \int K(\mathrm{d}\kappa) e^{-\sigma \kappa \left(\bigcup_{k \in \mathbb{N}} \mathrm{comp}_{\overline{x_k}}(\overline{y_k}) \cap \mathbb{S} \right)} \quad \text{and} \quad \int \nu(\mathrm{d}\mu) e^{-\sigma \mu \left(\bigcup_{k \in \mathbb{N}} \mathcal{S}_{x_k}(y_k) \right)}$$

have same distribution. But the term on the right above is deterministic, so we have that a.s. for all $x, y \in V_T$, and $\sigma \geq 0$,

$$(4.41) \qquad \int K(\mathrm{d}\kappa) e^{-\sigma\kappa\left(\bigcup_{k\in\mathbb{N}} \mathrm{comp}_{\overline{x_k}}(\overline{y_k}) \cap \mathbb{S}\right)} = \int \nu(\mathrm{d}\mu) e^{-\sigma\mu\left(\bigcup_{k\in\mathbb{N}} \mathcal{S}_{x_k}(y_k)\right)},$$

which yields in particular (CM2).

Step 3: $M_K = \lambda_\mathbb{S}$ **a.s.** By taking the derivative of (4.41) at $\sigma = 0$ and with (4.31), a.s. for all $x_k, y_k \in V_C$,

$$(4.42) \qquad M_K\left(\bigcup_{k\in\mathbb{N}} \mathrm{comp}_{\overline{x_k}}(\overline{y_k}) \cap \mathbb{S}\right) = M_\nu\left(\bigcup_{k\in\mathbb{N}} \mathcal{S}_{x_k}(y_k)\right) = \lambda_\mathbb{S}\left(\bigcup_{k\in\mathbb{N}} \mathrm{comp}_{\overline{x_k}}(\overline{y_k}) \cap \mathbb{S}\right).$$

In other words, $M_K = \lambda_\mathbb{S}$ on the σ-algebra generated by $\{\mathrm{comp}_{\overline{x}}(\overline{y}) \cap \mathbb{S} : x, y \in V_T\}$. To prove the equality on the Borel σ-algebra of $\mathbb{S}$, we partition the circle line $\mathbb{S}$ in the following way. Since $\Delta(C)$ is countable, there exist countably many disjoint intervals of non-empty interior I_p, $p \in \mathbb{N}$, such that

$$(4.43) \qquad \mathbb{S} = \biguplus_{p\in\mathbb{N}} I_p,$$

and for each $p \in \mathbb{N}$, I_p satisfies one of the following (see Figure 12 for a sub-triangulation with the different types of boundaries):

    (a) I_p corresponds to an atom leaf of (T, c), that is, there exists $w \in \mathrm{at}(M_\nu)$ such that $(I_p)^\circ = \overline{w} \cap \mathbb{S}$,

    (b) I_p corresponds to a line segment in $\mathrm{seg}(T, c, \nu)$, i.e. there exists $b \in \blacksquare(C)$ such that $(I_p)^\circ = (b \cap \mathbb{S})^\circ$,

    (c) $I_p \subseteq C$ and for all distinct $x, y \in I_p$, there exists $c_{xy} \in \Delta(C)$ such that x, y belong to different connected components of $\mathbb{D} \setminus \partial c_{xy}$.

We show that a.s. $M_K = \lambda_\mathbb{S}$ by showing that a.s. for any $I := I_p$, $M_{K|I} = \lambda_I$.

    (a) Suppose $I^\circ = \overline{w} \cap \mathbb{S}$ for $w \in \mathrm{at}(M_\nu)$. For all $n \leq i$, we defined κ_i^n such that

$$(4.44) \qquad \kappa_{i|I}^n = C_I^{i,n} \lambda_I$$

        where $C_I^{i,n}$ is a real-valued random-variable. Taking the weak limit of $(\kappa_i^{\varphi_i(n)})_n$ when n goes to infinity,

$$(4.45) \qquad \kappa_{i|I} = C_I^i \lambda_I$$

        a.s. for some real-valued random-variable C_I^i. Therefore, a.s. for all $m \in \mathbb{N}$,

$$(4.46) \qquad M_{K_m|I} = \left(\frac{1}{m} \sum_{i=1}^{m} C_I^i\right) \lambda_I.$$

        Taking the weak limit of $(K_{\varphi(m)})_m$ when m goes to infinity,

$$(4.47) \qquad M_{K|I} = C_I \lambda_I$$

        a.s. for some real-valued random-variable C_I. Noticing that I belongs to the σ-algebra generated by $\{\mathrm{comp}_{\overline{x}}(\overline{y}) \cap \mathbb{S} : x, y \in V_C\}$, we have $M_K(I) = \lambda_I(I)$ so that $C_I = 1$ a.s.

    (b) Suppose that $I^\circ = (b \cap \mathbb{S})^\circ$ for some $b \in \blacksquare(C)$. We will use here the assumption that for the corresponding $(v, w) \in \mathrm{seg}(T, c, \nu)$, M_ν restricted to (v, w) is the Lebesgue measure.

- If $b \cap \mathbb{S}$ is the union of two disjoint intervals (one of which is a singleton), then the map that associates a point in (v, w) to the endpoint of the corresponding segment in b is linear. Therefore, with the assumption on ν, $M_{K|I} = \lambda_I$.

- If not, let $x, x' \in \mathbb{S}$ such that $b \cap \mathbb{S} = [x, x']$, $y = y' = \frac{x+x'}{2}$ (see Figure 8). In this case, (v, w) is adjacent to a leaf and we can assume w.l.o.g. that $w \in \mathrm{lf}(T, c)$. The point $u = (1 - t)v + tw$, $t \in (0, 1)$ corresponds to the segment with endpoints $\widetilde{u_1} = (1-t)x + ty$ and $\widetilde{u_2} = (1-t)x' + ty'$ (see Figure 12). Therefore, by construction and assumption on ν, $M_{K|(x,y)}$ is proportional to $\lambda_{(x,y)}$ and $M_K(x, y) = \frac{M_K(x,x')}{2} = \frac{\lambda_{\mathbb{S}}(x,x')}{2} = \lambda_{\mathbb{S}}(x, y)$ because $[x, x']$ is the intersection of a component with $\mathbb{S}$. Thus $M_{K|(x,y)} = \lambda_{(x,y)}$ and since it also holds for (y', x'), $M_{K|I} = \lambda_I$.

(c) Finally, we assume that $I \subseteq C$ and for all distinct $x, y \in I$, there exists $c_{xy} \in \Delta(C)$ such that x, y belong to different connected components of $\mathbb{D} \setminus \partial c_{xy}$. Note that in this case, $[x, y] \cap c_{xy} \neq \emptyset$. In other words, the set of endpoints of (boundaries of) triangles in $\Delta(C)$ is dense in I. Let $J = [a, b] \subseteq I$. There exist two sequences $(a_n)_n$ and $(b_n)_n$ in I such that for all n, $a_n \in \partial_{\mathbb{D}} v$ and $b_n \in \partial_{\mathbb{D}} w$ for some $v, w \in \Delta(C)$ and $a_n \uparrow a$ and $b_n \downarrow b$. Since $a_n \in \partial_{\mathbb{D}} v$ and $b_n \in \partial_{\mathbb{D}} w$, $[a_n, b_n]$ is the union of countably many disjoint components of C intersected with $\mathbb{S}$, so that $M_K([a_n, b_n]) = \lambda_{\mathbb{S}}([a_n, b_n])$. Therefore, $M_K(J) = \lim_{n \to \infty} M_K([a_n, b_n]) = \lambda_{\mathbb{S}}(J)$. Since it holds for all interval $J \subseteq I$, $M_{K|I} = \lambda_I$.

We have shown that a.s. the random two-level measure K together with the sub-triangulation C is such that $\tau(C, K) = (T, c, \nu)$ and $M_K = \lambda_{\mathbb{S}}$. Therefore, by taking a realization of K such that it holds, we have shown surjectivity of τ. $\qquad \square$

5. Topologies on the subspace of binary algebraic two-level measure trees

Recall that in Section 2.2 we have equiped $\mathbb{T}^{(2)}$ with the Gromov-weak topology using a metric representation that comes from the branch point distribution. Here we introduce on the space of binary algebraic two-level measure trees another notion of convergence called *two-level sample shape convergence*. It exploits the idea of the Gromov-weak topology to sample finite sub-spaces and then require these to converge in distribution. Whereas one samples metric sub-spaces in the Gromov-weak topology, we now consider subtrees of a2m trees as combinatorial objects, which will make it useful to show convergence of tree-valued Markov chains in the future. We show that this topology is equivalent to the two-level bpdd-Gromov-weak topology on $\mathbb{T}_2^{(2)}$, and that both topologies are compact.

5.1. Two-level sample shape convergence.
Obviously the two-level sample shape convergence is similar to the notion of sample shape convergence defined on the space of algebraic measure trees (see [LW21], that we follow closely for this generalization to the two-level case). A sequence of trees converges to a limit tree if all random tree shapes spanned by finite samples converge weakly to the corresponding limit shapes. On $\mathbb{T}_2^{(2)}$, there is a two-level sampling procedure: we first sample a finite number of measures on the tree from the two-level measure and then sample a finite number of points according to each sampled measure. Therefore, the notion of tree shape that we define keeps track of this two-level sampling by using double indices.
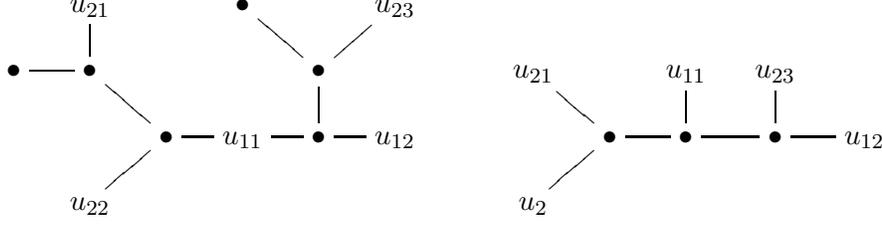
FIGURE 14. A tree T and the shape $\mathfrak{s}_{(T,c)}(u_{11}, u_{12}, u_{21}, u_{22}, u_{23})$.

For $m \in \mathbb{N}$ and $\underline{n} \in \mathbb{N}^m$, we denote by

$$\text{(5.1)} \qquad \mathbb{N}_{m,\underline{n}} = \{(i,j) : 1 \leq i \leq m, 1 \leq j \leq n_i\}$$

the set of indices. The sampled subtrees from an a2m tree will be cladograms, which are binary finite trees with labelled leaves. Since the two-level measure on the tree may have atoms on leaves, a given leaf may be sampled several times. We therefore need to allow multi-labels in cladograms.

Definition 5.1 $((m,\underline{n})$-cladogram). *For $m \in \mathbb{N}$ and $\underline{n} \in \mathbb{N}^m$, an $(m, \underline{n})$-labelled cladogram is a binary, finite algebraic tree (C, c) consisting only of leaves and branch points together with a surjective labelling map $\zeta \colon \mathbb{N}_{m,\underline{n}} \to \mathrm{lf}(C)$. An $(m, \underline{n})$-cladogram (C, c, ζ) is an $(m, \underline{n})$-labelled cladogram such that ζ is also injective.*

We call two $(m, \underline{n})$-labelled cladograms (C_1, c_1, ζ_1) and (C_2, c_2, ζ_2) isomorphic if there exists a tree isomorphism ϕ from (C_1, c_1) onto (C_2, c_2) such that $\zeta_2 = \phi \circ \zeta_1$. We then write

$$\text{(5.2)} \qquad \overline{\mathfrak{C}}_{m,\underline{n}} := \{\textit{isomorphism classes of } (m, \underline{n})\textit{-labelled cladograms}\}$$

and

$$\text{(5.3)} \qquad \mathfrak{C}_{m,\underline{n}} := \{(C, c, \zeta) \in \overline{\mathfrak{C}}_{m,\underline{n}} : \zeta \textit{ injective}\}.$$

The shape function encodes as cladograms the shape of the subtree spanned by finite samples of points (see Figure 14).

Definition 5.2 (Shape function). *Let (T, c) be a binary algebraic tree, $m \in \mathbb{N}$, $\underline{n} \in \mathbb{N}^m$, and $u_{ij} \in T \setminus \mathrm{br}(T)$ for $(i, j) \in \mathbb{N}_{m,\underline{n}}$. Then there exists a unique (up to isomorphism) (m, n)-labelled cladogram*

$$\text{(5.4)} \qquad \mathfrak{s}_{(T,c)}(\underline{\underline{u}}) = (C, c_C, \zeta)$$

with $\mathrm{lf}(C) = \{u_{ij}\}_{(i,j) \in \mathbb{N}_{m,\underline{n}}}$ and $\zeta(i, j) = u_{ij}$, such that the identity on $\mathrm{lf}(C)$ extends to a tree homomorphism π from C onto $c\left((\{u_{ij}\}_{(i,j) \in \mathbb{N}_{m,\underline{n}}})^3 \right)$, i.e. for all $(i_1, j_1), (i_2, j_2), (i_3, j_3) \in \mathbb{N}_{m,\underline{n}}$,

$$\text{(5.5)} \qquad \pi(c_C(u_{i_1 j_1}, u_{i_2 j_2}, u_{i_3 j_3})) = c(u_{i_1 j_1}, u_{i_2 j_2}, u_{i_3 j_3}).$$

We will refer to $\mathfrak{s}_{(T,c)}(\underline{\underline{u}}) \in \overline{\mathfrak{C}}_m$ as the shape of $\underline{\underline{u}}$ in (T, c).

We introduce a notion of convergence on $\mathbb{T}_2^{(2)}$ based on the weak convergence of random tree shapes spanned by finite samples. This topology is defined as the topology induced by the two-level shape polynomials, which are test functions evaluating the tree shape distributions.

Definition 5.3 (Two-level shape polynomials). *A two-level shape polynomial is a function* $\Phi^{m,\underline{n},\varphi} \colon \mathbb{T}_2^{(2)} \to \mathbb{R}$ *of the form*

$$(5.6) \qquad \Phi^{m,\underline{n},\varphi}(\chi) := \int_{(\mathcal{M}_1(T))^m} \nu^{\otimes m}(\mathrm{d}\underline{\mu}) \int_{T^{|\underline{n}|}} \bigotimes_{i=1}^m \mu_i(\mathrm{d}\underline{u_i}) \varphi\left(\mathfrak{s}_{(T,c)}(\underline{u})\right),$$

where $\chi = (T, c, \nu)$, $m \in \mathbb{N}$, $\underline{n} \in \mathbb{N}^m$ *and* $\varphi \colon \mathfrak{C}_{m,\underline{n}} \to \mathbb{R}$. *We write* $\Pi_{\mathfrak{s}}^{(2)}$ *for the set of all two-level shape polynomials.*

Definition 5.4 (Two-level sample shape topology). *The two-level sample shape topology on* $\mathbb{T}_2^{(2)}$ *is defined as the initial topology induced by* $\Pi_{\mathfrak{s}}^{(2)}$.

Remark 5.5 (Tree shape distribution). Fix $m \in \mathbb{N}$ and $\underline{n} \in \mathbb{N}^m$. For all $\chi = (T, c, \nu)$, we define the $(m, \underline{n})$-*tree shape distribution* $\mathfrak{S}_{m,\underline{n}}(\chi)$ as the probability measure on $\mathfrak{C}_{m,\underline{n}}$ such that for all $\varphi \colon \mathfrak{C}_{m,\underline{n}} \to \mathbb{R}$,

$$(5.7) \qquad \int_{\mathfrak{C}_{m,\underline{n}}} \mathrm{d}\mathfrak{S}_{m,\underline{n}}(\chi)\, \varphi = \int_{\mathcal{M}_1(T)} \nu^{\otimes m}(\mathrm{d}\underline{\mu}) \int_{T^{|\underline{n}|}} \bigotimes_{i=1}^m \mu_i(\mathrm{d}\underline{u_i}) \varphi\left(\mathfrak{s}_{(T,c)}(\underline{u})\right).$$

Then the two-level sample shape topology is induced by the set of functions

$$(5.8) \qquad \mathcal{F}_{\mathfrak{s}} := \{\mathfrak{S}_{m,\underline{n}} : m \in \mathbb{N}, \underline{n} \in \mathbb{N}\}.$$

The following result gives that on $\mathbb{T}_2^{(2)}$, two-level sample shape convergence implies two-level bpdd-Gromov-weak convergence.

Proposition 5.6. *On* $\mathbb{T}_2^{(2)}$, *the two-level sample shape topology is stronger than the two-level bpdd-Gromov-weak topology.*

Proof. By definition, the two-level bpdd-Gromov-weak topology is induced by the set $\Pi_\iota^{(2)}$ of polynomials of the form

$$(5.9) \qquad \Phi(\chi) := \int \nu^{\otimes m}(\mathrm{d}\underline{\mu}) \int \bigotimes_{i=1}^m \mu_i(\mathrm{d}\underline{u_i}) \varphi\left((r_\lambda(u_{ij}, u_{i'j'})_{(i,j),(i',j')})\right),$$

where $\chi = (T, c, \nu)$, $\lambda := c_* M_\nu^{\otimes 3}$, $m \in \mathbb{N}$, $\underline{n} \in \mathbb{N}^m$ and $\varphi \in \mathcal{C}_b(\mathbb{R}^{|\underline{n}|^2})$ (see Definition 2.9 and Definition 3.15). Now, the set of $\phi \in \mathcal{C}_b(\mathbb{R}^{|\underline{n}|^2})$ that are Lipschitz continuous is convergence determining for probability measure on $\mathbb{R}^{|\underline{n}|^2}$. Hence, the subset of $\Psi \in \Pi_\iota^{(2)}$ with

$$(5.10) \qquad \Psi(T, c, \nu) = \int \nu^{\otimes m}(\mathrm{d}\underline{\mu}) \int \bigotimes_{k=1}^m \mu_k^{\otimes n_k}(\mathrm{d}\underline{u_k}) \phi\left((r_\lambda(u_{ij}, u_{i'j'})_{(i,j),(i',j')})\right)$$

for some $m \in \mathbb{N}$, $\underline{n} \in \mathbb{N}^m$ and a Lipschitz continuous function $\phi \in \mathcal{C}_b(\mathbb{R}^{|\underline{n}|^2})$ also induces the two-level bpdd-Gromov-weak topology. Therefore, it is enough to show that such a Ψ is continuous on $\mathbb{T}_2^{(2)}$ with respect to the sample shape topology. To do this, we show that the restriction of Ψ to $\mathbb{T}_2^{(2)}$ can be uniformly approximated by polynomials in $\Pi_{\mathfrak{s}}^{(2)}$.

For $p \in \mathbb{N}$ with $3p \geq m$, we define

$$(5.11) \quad \Phi_p(T, c, \nu) = \int \nu^{\otimes 3p}(\mathrm{d}\underline{\mu}) \int \bigotimes_{k=1}^{m} \mu_k^{\otimes n_k}(\mathrm{d}\underline{u_k}) \bigotimes_{k=m+1}^{3p} \mu_k(\mathrm{d}u_{k1}) \phi\left(\left(r_{\lambda_{p,\underline{u}}}(u_{ij}, u_{i'j'})_{(i,j),(i',j')}\right)\right)$$

where

$$(5.12) \qquad\qquad \lambda_{p,\underline{u}} := \frac{1}{p} \sum_{l=0}^{p} \delta_{c(u_{3l+1,1}, u_{3l+2,1}, u_{3l+3,1})}.$$

Note that whether or not $c(u_{3l+1,1}, u_{3l+2,1}, u_{3l+3,1})$ lies on $[u_{(i,j)}, u_{(i',j')}]$, for some $l \in \{0, ..., p-1\}$ and $(i,j), (i', j') \in \mathbb{N}_{m,\underline{n}}$ only depends on the shape $\mathfrak{s}_{(T,c)}(\underline{u})$ and thus the restriction of Φ_p to $\mathbb{T}_2^{(2)}$ belongs to $\Pi_{\mathfrak{s}}^{(2)}$.

To show that Φ_p approximates Ψ, we uniformly bound the distance of the empirical branch point distribution to the branch point distribution. Recall that for $x, y \in T$,

$$(5.13) \qquad\qquad r_\lambda(x, y) := \lambda([x, y]) - \frac{1}{2}\lambda(\{x\}) - \frac{1}{2}\lambda(\{y\}).$$

Thus, denoting by L the Lipschitz constant of ϕ w.r.t. the l_∞-norm on $\mathbb{R}^{|\underline{n}|^2}$,

$$(5.14)$$

$$||\Psi - \Phi_p||_\infty \leq \sup_{(T,c,\nu)\in\mathbb{T}_2^{(2)}} \int \nu^{\otimes 3p}(\mathrm{d}\underline{\mu}) \int \bigotimes_{k=1}^{m} \mu_k^{\otimes n_k}(\mathrm{d}\underline{u_k}) \bigotimes_{k=m+1}^{3p} \mu_k(\mathrm{d}u_{k1}) L.2 \sup_{I \in \mathcal{I}_T} |\lambda(I) - \lambda_{p,\underline{u}}(I)|$$

$$\leq 2L \sup_{(T,c,\nu)\in\mathbb{T}_2^{(2)}} \int (M_\nu)^{\otimes 3p}(\mathrm{d}\underline{u}) \sup_{I \in \mathcal{I}_T} |\lambda(I) - \lambda_{p,\underline{u}}(I)|,$$

where $\mathcal{I}_T := \{[x, y] : x, y \in T\}$ is the collection of all intervals of the algebraic tree T.

Then [LW21, Lemma A.4] gives the following estimates of the rate of convergence in the approximation of the branch point distribution by empirical distribution:

$$(5.15) \qquad \int (M_\nu)^{\otimes 3p}(\mathrm{d}\underline{u}) \sup_{I \in \mathcal{I}_T} |\lambda(I) - \lambda_{p,\underline{u}}(I)| \leq 96\sqrt{\frac{\dim_{\mathrm{VC}}(\mathcal{I}_T)}{p}} = 96\sqrt{\frac{2}{p}},$$

where $dim_{\mathrm{VC}}(\mathcal{I}_T)$ is the Vapnik-Chervonenkis dimension of $\mathcal{I}_T$ which can easily be shown to be 2 (see [LW21, Example A.2] for more details). This concludes the proof. $\qquad\square$

Corollary 5.7 (Metrizability). *The two-level sample shape topology is metrizable on* $\mathbb{T}_2^{(2)}$.

Proof. By Proposition 5.6, the two-level sample shape topology is stronger than the two-level bpdd-Gromov-weak topology, which is Hausdorff by Proposition 3.16. Hence the two-level sample shape topology is also Hausdorff. Moreover, by Remark 5.5, it is induced by the set $\mathcal{F}_\mathfrak{s}$, which is a countable family of functions with values in metrizable spaces. Thus we can define a pseudo-metric on $\mathbb{T}_2^{(2)}$ that induces the two-level sample shape topology as follows: for $\chi, \chi' \in \mathbb{T}_2^{(2)}$,

$$(5.16) \qquad d_\mathfrak{s}(\chi, \chi') := \sum_{m \in \mathbb{N}} \frac{1}{2^m} \sum_{\underline{n} \in \mathbb{N}^m} \frac{1}{2^{|\underline{n}|}} \min\left(d_{\mathrm{Pr}}(\mathfrak{S}_{m,\underline{n}}(\chi), \mathfrak{S}_{m,\underline{n}}(\chi')), 1\right),$$

where d_{Pr} denotes the Prokhorov distance. Therefore, the two-level sample shape topology is Hausdorff and pseudo-metrizable. In particular, if $d_{\mathfrak{s}}(\chi, \chi') = 0$ then every open set that contains χ also contains χ', so that $d_{\mathfrak{s}}$ is actually a metric. $\qquad\square$

5.2. **Equivalence and compactness of topologies.** In this section, we first show that the coding map τ is continuous when $\mathbb{T}_2^{(2)}$ is equipped with the two-level sample shape topology. This implies that it is a compact topology. Finally, we prove that the two-level sample shape convergence and the two-level bpdd Gromov-weak convergence are equivalent on $\mathbb{T}_2^{(2)}$.

Recall from (4.8) the space $\mathfrak{D}$ of pairs $(C, K) \in \mathcal{T} \times \mathcal{M}_1(\mathcal{M}_1(\mathbb{S}))$ such that $M_K = \lambda_{\mathbb{S}}$.

Proposition 5.8. *Let $\mathcal{T}$ be equipped with the Hausdorff metric topology, $\mathcal{M}_1(\mathcal{M}_1(\mathbb{S}))$ with the weak topology, $\mathcal{T} \times \mathcal{M}_1(\mathcal{M}_1(\mathbb{S}))$ with the product topology and $\mathbb{T}_2^{(2)}$ with the two-level sample shape topology. Then the coding map $\tau : \mathfrak{D} \to \mathbb{T}_2^{(2)}$ is continuous.*

We will use the following lemma in the proof.

Lemma 5.9. *Let (S, d) be separable and let $P, Q \in \mathcal{M}_1(S)$, $\epsilon > 0$ such that $d_{\mathrm{Pr}}(P, Q) \leq \epsilon$ where d_{Pr} denotes the Prokhorov distance. Then we can define two random variables X, resp. Y, distributed according to P, resp. Q, on the same probability space $(\Omega, \mathbb{P})$ such that*

$$(5.17) \qquad\qquad \mathbb{P}\{d(X, Y) \geq \epsilon\} \leq \epsilon.$$

Proof. The lemma directly follows from [EK86, Theorem 3.1.2]:

$$(5.18) \qquad\qquad d_{\mathrm{Pr}}(P, Q) = \inf_{\gamma} \inf \left\{ \epsilon > 0 \,\big|\, \gamma\{(x, y) | d(x, y) \geq \epsilon\} \leq \epsilon \right\},$$

where the first infimum is taken over the set of measures $\gamma \in \mathcal{M}_1(S \times S)$ with marginals P and Q. $\qquad\square$

We can now proceed with the proof of the continuity of the coding map.

Proof of Proposition 5.8. Fix $\Gamma = (C, K) \in \mathfrak{D}$, $m \in \mathbb{N}$ and $\underline{n} \in \mathbb{N}^m$. By Remark 5.5, it is enough to show that $\mathfrak{S}_{m,\underline{n}} \circ \tau : \mathcal{T} \to \mathcal{M}_1(\mathfrak{C}_{m,\underline{n}})$ is continuous at (C, K).

Let $\kappa_1, ..., \kappa_m$ be independent, identically distributed measures on $\mathcal{M}_1(\mathbb{S})$ with distribution K. For all $1 \leq i \leq m$, let $U_{i1}, ..., U_{in_i}$ be independent, identically distributed points on $\mathbb{S}$ with distribution κ_i. Fix $\epsilon > 0$. Since $M_K = \lambda_{\mathbb{S}}$, there exist $N \in \mathbb{N}$ and $v_1, ..., v_N \in \Delta(C) \cap \nabla(C)$ distinct, such that with probability at least $1 - \epsilon$ the following holds:

- if $\{U_{ij} : (i, j) \in \mathbb{N}_{m,\underline{n}}\} \cap \{v\} \neq \emptyset$ for some $v \in \nabla(C)$, then $v \in \{v_1, ..., v_N\}$, and
- if $\{U_{ij} : (i, j) \in \mathbb{N}_{m,\underline{n}}\} \cap \mathrm{comp}_v(w) \neq \emptyset$ for some $v \in \Delta(C)$ and all $w \in \Delta(C) \cup \nabla(C) \cup \square(C)$ with $w \neq v$, then $v \in \{v_1, ..., v_N\}$.

Put $\epsilon' := \epsilon(12N|\underline{n}|)^{-1}$. Then, since $M_K = \lambda_{\mathbb{S}}$, each U_{ij} is distributed according to the Lebesgue measure on $\mathbb{S}$, so that

$$
\begin{aligned}
& \mathbb{P}\big(\{d(U_{ij}, \partial v_k) \geq \epsilon', \forall (i, j) \in \mathbb{N}_{m,\underline{n}}, k = 1, ..., N\}\big) \\
(5.19) \qquad & \geq 1 - \sum_{(i,j) \in \mathbb{N}_{m,\underline{n}}} \mathbb{P}\big(\{d(U_{ij}, \partial v_k) \leq \epsilon', \forall k = 1, ..., N\}\big) \\
& \geq 1 - |\underline{n}|\mathbb{P}\big(\{d(U_{11}, \partial v_k) \leq \epsilon', \forall k = 1, ..., N\}\big) \\
& \geq 1 - \epsilon.
\end{aligned}
$$

Now, there is a $\delta = \delta(\epsilon) > 0$ small enough such that for any $C' \in \mathcal{T}$ satisfying $d_H(C, C') < \delta$ there are distinct $v'_1, ..., v'_N \in \Delta(C') \cup \nabla(C')$ such that $d_H(v_k, v'_k) \leq \frac{\epsilon'}{2}$ for $k = 1, ..., N$. Let $K' \in \mathcal{M}_1(\mathcal{M}_1(\mathbb{S}))$ such that $M_K = \lambda_{\mathbb{S}}$. Suppose that $d_{\mathrm{Pr}}(K, K') \leq \epsilon'' := \min(\frac{\epsilon}{2|\underline{n}|}, \frac{\epsilon'}{2})$. By Lemma 5.9, there exist $\kappa'_1, ..., \kappa'_m$ independent, identically distributed measures on $\mathcal{M}_1(\mathbb{S})$ with distribution K' coupled to $\kappa_1, ..., \kappa_m$ such that for all $i = 1, ..., m$, $\mathbb{P}\{d_{\mathrm{Pr}}(\kappa_i, \kappa'_i) \geq \epsilon''\} \leq \epsilon''$. Applying again Lemma 5.9 for each $1 \leq i \leq m$, there exist $U'_{i1}, ..., U'_{in_i}$ independent, identically distributed points on $\mathbb{S}$ with distribution κ'_i coupled to $U_{i1}, ..., U_{in_i}$ such that for all $j = 1, ..., n_i$,

$$(5.20) \quad \mathbb{P}\{d(U_{ij}, U'_{ij}) \leq \epsilon''\} = \mathbb{P}\{d(U_{ij}, U'_{ij}) \leq \epsilon'' | d_{\mathrm{Pr}}(\kappa_i, \kappa'_i) \leq \epsilon''\}\mathbb{P}\{d_{\mathrm{Pr}}(\kappa_i, \kappa'_i) \leq \epsilon''\} \geq 1 - 2\epsilon''.$$

Therefore,

$$(5.21) \qquad\qquad \mathbb{P}\big(\big\{d(U_{ij}, U'_{ij}) \leq \epsilon'', \forall (i,j) \in \mathbb{N}_{m,\underline{n}}\big\}\big) \geq 1 - 2|\underline{n}|\epsilon''.$$

Thus, since $\epsilon'' \leq \frac{\epsilon'}{2}$, and using (5.19),

$$(5.22) \qquad\qquad \mathbb{P}\left(\left\{d(U'_{ij}, \partial v_k) \geq \frac{\epsilon'}{2}, \forall (i,j) \in \mathbb{N}_{m,\underline{n}}, k = 1, ..., N\right\}\right) \geq 1 - 2\epsilon.$$

Let $\chi = (T, c, \nu) := \tau(C, K)$ and $(V_{ij})_{(i,j)\in\mathbb{N}_{m,\underline{n}}}$ be distributed according to

$$(5.23) \qquad\qquad M^{m,\underline{n}}_\nu(\cdot) := \int \nu^{\otimes m}(\mathrm{d}\underline{\mu}) \int \bigotimes_{i=1}^m \mu_i^{\otimes n_i}(\cdot),$$

coupled to $(U_{ij})_{(i,j)\in\mathbb{N}_{m,\underline{n}}}$ such that $V_{ij} \in \mathcal{S}_v(w)$ if and only if $U_{ij} \in \mathrm{comp}_v(w)$. This is possible due to the properties of the coding map τ stated in Theorem 1. Define χ' and $(V'_{ij})_{(i,j)\in\mathbb{N}_{m,\underline{n}}}$ similarly with (C, K) replaced by (C', K'). Then

$$(5.24) \qquad\qquad \mathbb{P}\big(\big\{\mathfrak{s}_{(T,c)}(\underline{V}) = \mathfrak{s}_{(T',c')}(\underline{V'})\big\}\big) \geq 1 - 2\epsilon.$$

Therefore,

$$(5.25) \qquad\qquad d_{\mathrm{Pr}}\big(\mathfrak{S}_{m,\underline{n}}(\tau(C, K)), \mathfrak{S}_{m,\underline{n}}(\tau(C', K'))\big) \leq 2\epsilon.$$

We proved that $\mathfrak{S}_{m,\underline{n}} \circ \tau$ is continuous at (C, K), with $m, \underline{n}$ and (C, K) arbitrary. Therefore, the coding map τ is continuous. $\square$

We finish this section with our second main result.

Theorem 2 (Equivalence of topologies and compactness). *The two-level sample shape topology and the two-level bpdd-Gromov-weak topology coincide on $\mathbb{T}_2^{(2)}$. Furthermore, $\mathbb{T}_2^{(2)}$ is compact and metrizable in this topology.*

Proof. We start by showing that the two-level sample shape topology on $\mathbb{T}_2^{(2)}$ is compact. The set $\mathcal{T}$ equipped with the Hausdorff metric topology is compact (see [LW21, Lemma 4.2]). Moreover, the circle line $\mathbb{S}$ is compact so that $\mathcal{M}_1(\mathbb{S})$ is compact and $\mathcal{M}_1(\mathcal{M}_1(\mathbb{S}))$ as well. Since the subset of two-level measures K such that $M_K = \lambda_{\mathbb{S}}$ is closed in $\mathcal{M}_1(\mathcal{M}_1(\mathbb{S}))$, $\mathfrak{D}$ is a compact space. By Theorem 1, the coding map is surjective and by Proposition 5.8, it is continuous when $\mathbb{T}_2^{(2)}$ is equipped with the two-level sample shape topology. Therefore, the sample shape topology is a compact topology.

Furthermore, the two-level bpdd-Gromov-weak topology is a Hausdorff topology by Proposition 3.16 and weaker than the two-level sample shape topology by Proposition 5.6. Thus, both topologies coincide on $\mathbb{T}_2^{(2)}$, and we know from Corollary 5.7 that is is metrizable. $\qquad\square$

6. Example: the Kingman algebraic two-level measure tree

One of the first stochastic models describing the genealogy of a large population is the Kingman coalescent. It is a partition-valued process introduced in 1982 [Kin82] to model an evolutionary tree by looking at the genealogy backwards in time. There is not hierarchy in the original Kingman coalescent. Independently and with the same rate, each pair of individuals coalesce, meaning that their partitions merge. A natural extension of this model takes into account that the population consists of several species. Both at the individual and species level, the coalescences follow the rule of the classical Kingman coalescent. Pairs of individuals belonging to the same species can merge as before, but individuals belonging to different species cannot coalesce before their species themselves have coalesced. Such process, named as the nested Kingman coalescent, has been considered by [BDLS18, BRSSJ19, Daw18] or [Mei19] for instance.

We introduce here the Kingman algebraic two-level measure tree, which corresponds to the nested Kingman coalescent measure tree, as defined in [Mei19], without branch length. For this, we rely on a sampling consistency of the nested Kingman coalescent and the compactness of $\mathbb{T}_2^{(2)}$. We first recall a definition of the nested Kingman coalescent for a host-parasite population indexed in $\mathbb{N}^2$, that is, (i, j) is the j-th parasite in the i-th host.

For $I \subseteq \mathbb{N}$, let $\mathcal{E}(I)$ be the set of equivalence relations on I. The equivalence classes of an equivalence relation are called *blocks*. Let

$$(6.1) \qquad \mathcal{P}_{nest}(I) \subset \mathcal{E}(I) \times \mathcal{E}(I)$$

denote the set of all pairs $(\mathfrak{p}_H, \mathfrak{p}_P)$ of *nested* equivalence relations on I, i.e. such that

- if $i_1 = i_2$, then (i_1, j_1) and (i_2, j_2) belong to the same block of $\mathfrak{p}_H$, that is, all parasites of the same host belong to the same block of $\mathfrak{p}_H$,
- if (i_1, j_1) and (i_2, j_2) belong to the same block of $\mathfrak{p}_P$, then they belong to the same block of $\mathfrak{p}_H$. In other words, each block of $\mathfrak{p}_P$ is contained in a single block of $\mathfrak{p}_H$. This property is the reason to call it "nested".

Therefore, $\mathfrak{p}_H$ represents the population of hosts and $\mathfrak{p}_P$ the population of parasites.

We also define the following equivalence relations on $\mathbb{N}^2$:

$$(6.2) \qquad \begin{aligned} P_0 &:= \left\{ \left((i,j),(i,j)\right) \big| (i,j) \in \mathbb{N}^2 \right\} \\ H_0 &:= \left\{ \left((i,j),(i,k)\right) \big| i,j,k \in \mathbb{N} \right\}, \end{aligned}$$

which will be the initial states of the nested Kingman coalescent.

Definition 6.1 (Finite nested Kingman coalescent). *Let I be a finite subset of $\mathbb{N}^2$ and $\gamma_H, \gamma_P > 0$. The* finite nested Kingman coalescent

$$(6.3) \qquad (\mathcal{K}^I(t))_{t \geq 0} = \left(\kappa_H^I(t), \kappa_P^I(t) \right)_{t \geq 0}$$

on I with rates (γ_H, γ_P) is a continuous-time Markov process with values in $\mathcal{P}_{nest}(I)$ such that:
  (1) *The initial state is $\mathcal{K}^I(0) = (H_0 \cap I^2, P_0 \cap I^2)$.*
  (2) *$(\kappa_H^I(t))_{t \geq 0}$ is a Kingman coalescent with rate γ_H, i.e. any pair of blocks in $\kappa_H^I(t)$ merge at rate γ_H.*

(3) $(\kappa_P^I(t))_{t \geq 0}$ *behaves in the following way: any pair of blocks* π_1, π_2 *of* $\kappa_P^I(t)$ *such that* $\pi_1 \cup \pi_2$ *is contained in a single block of* $\kappa_H^I(t)$ *merge at rate* γ_P. *Other blocks of* $\kappa_P^I(t)$ *cannot merge.*

Roughly speaking, hosts merge as in a Kingman coalescent with rate γ_H, and parasites within the same host merge as in a Kingman coalescent with rate γ_P. Since this process has only finitely many states, it is well defined and unique. Furthermore, it satisfies an important property:

Proposition 6.2 (Sampling consistency). *Let* I *and* J *be two finite subsets of* $\mathbb{N}^2$ *such that* $J \subset I$ *and let* $\gamma_H, \gamma_P > 0$. *The restriction of* $\mathcal{K}^I$ *to* J *has same distribution as* $\mathcal{K}^J$, *that is, the restriction of* $\mathcal{K}^I$ *to* J *is a finite nested Kingman coalescent on* J *with rates* (γ_H, γ_P).

Using this property, we can show the existence of the nested Kingman coalescent for an infinite set of hosts and of parasites (see [BDLS18, Section 5] for a construction of more general nested coalescents).

Definition 6.3 (Nested Kingman coalescent). *Let* $\gamma_H, \gamma_P > 0$. *The* nested Kingman coalescent $\mathcal{K}$ *with rates* (γ_H, γ_P) *is a continuous-time Markov process with values in* $\mathcal{P}_{nest}(\mathbb{N}^2)$ *such that for any finite subset* I *of* $\mathbb{N}^2$, *the restriction of* $\mathcal{K}$ *to* I *is a finite nested Kingman coalescent on* I *with rates* (γ_H, γ_P).

In order to define the Kingman algebraic two-level measure tree, we first introduce the nested Kingman *rooted* algebraic tree. For this, let us shortly explain how algebraic trees are extended to rooted algebraic trees.

Definition 6.4 (Rooted and partially ordered algebraic tree). *A partially ordered algebraic tree is a non-empty set* T *together with a symmetric map* $c_\wedge : T \times T \to T$ *such that:*

(M1) *For all* $x \in T$, $c_\wedge(x, x) = x$.

(M2) *For all* $x_1, x_2, x_3 \in T$, $c_\wedge(x_1, c_\wedge(x_2, x_3)) = c_\wedge(c_\wedge(x_1, x_2), x_3)$.

(M3) *For all* $x_1, x_2, x_3 \in T$, $\#\{c_\wedge(x_1, x_2), c_\wedge(x_1, x_3), c_\wedge(x_2, x_3)\} \leq 2$ *and if* $c_\wedge(x_1, x_2) = c_\wedge(x_1, x_3)$, *then*

(6.4) $$c_\wedge(x_1, x_2) = c_\wedge(c_\wedge(x_1, x_2), c_\wedge(x_2, x_3)).$$

We refer to $c_\wedge$ *as the* minimum map.

A rooted algebraic tree $(T, c_\wedge)$ *is a partially ordered algebraic tree for which there exists a point* $\rho \in T$ *with* $c_\wedge(\rho, x) = \rho$ *for all* $x \in T$. *We will refer to (this unique)* ρ *as the* root *of* $(T, c_\wedge)$.

Remark 6.5 ((M1) and (M2) define a partial order). Let $(T, c_\wedge)$ be a partially ordered algebraic tree. In what follows, we write for $x, y \in T$, $x \leq y$ if and only if $x = c_\wedge(x, y)$ respectively, $x < y$ if and only if $x = c_\wedge(x, y) \neq y$. Notice that the first two conditions (M1) and (M2) ensure that $\leq$ defines a partial order relation. Indeed, reflexivity follows from (M1), antisymmetry follows from the fact that $c_\wedge$ is a symmetric map and transitivity follows from (M2), i.e., if $x, y, z \in T$ are such that $x \leq y$ and $y \leq z$, or equivalently, if $x = c_\wedge(x, y)$ and $y = c_\wedge(y, z)$, then $x = c_\wedge(x, y) = c_\wedge(x, c_\wedge(y, z)) = c_\wedge(c_\wedge(x, y), z) = c_\wedge(x, z)$.

When we add the third condition (M3) to the partially ordered set defined by (M1) and (M2), we ensure that there are no loops, so that a rooted algebraic tree defines an (unrooted) algebraic tree. Reciprocally, by distinguishing a point in an algebraic tree, we can define a rooted algebraic tree (see [NW, Section 2.2] for more details).

Proposition 6.6 (Rooted versus unrooted algebraic trees). *Let $T \neq \emptyset$.*

(i) *If $(T, c_\wedge)$ is a partially ordered, algebraic tree and $c : T^3 \to T$ the symmetric map defined as*

(6.5) $$c(x, y, z) := \max\{c_\wedge(x, y), c_\wedge(x, z), c_\wedge(y, z)\}, \quad x, y, z \in T,$$

then (T, c) is an algebraic tree.

(ii) *If (T, c) is an algebraic tree, $\rho \in T$, and $c_\wedge : T^2 \to T$ the symmetric map defined as*

(6.6) $$c_\wedge(x, y) := c(x, y, \rho), \quad x, y \in T,$$

then $(T, c_\wedge, \rho)$ is a rooted algebraic tree.

We now construct the Kingman a2m tree with rates (γ_H, γ_P). Given $\mathcal{K} = (\kappa_H, \kappa_P)$, we define the *nested Kingman rooted algebraic tree* as the random rooted algebraic tree $(T, c_\wedge, \rho)$ with the vertex set

(6.7) $$T := \{\varpi_\rho\} \uplus \bigcup_{t \geq 0} \bigcup_{\varpi \in \kappa_P(t)} \varpi,$$

with $\varpi_\rho = \mathbb{N}^2 \uplus \{\rho\}$ for a point $\rho \notin \mathbb{N}^2$ and the minimal map c_ρ which sends two elements $\varpi, \varpi' \in T$ to the smallest $\tilde{\varpi} \in T$ which contains both ϖ and ϖ', i.e.,

(6.8) $$c_\wedge(\varpi, \varpi') := \bigcap_{\tilde{\varpi} \in T : \varpi, \varpi' \subseteq \tilde{\varpi}} \tilde{\varpi}.$$

Further, we define the *nested Kingman algebraic tree* as the random algebraic space (T, c) obtained from the rooted nested Kingman algebraic tree $(T, c_\wedge, \varpi_\rho)$ as in (6.6).

For $M \in \mathbb{N}$ and $\underline{N} \in \mathbb{N}^M$, define also the two-level measure

(6.9) $$\nu^{M,\underline{N}} := \frac{1}{M} \sum_{i=1}^{M} \delta_{\frac{1}{N_i} \sum_{j=1}^{N_i} \delta_{\{(i,j)\}}},$$

so that $(T, c, \nu^{M,\underline{N}})$ is a random binary a2m tree in $\mathbb{T}_2^{(2)}$. We define the Kingman a2m tree as the weak limit (with respect to the two-level sample shape convergence) of these random trees as $M, \underline{N} \to \infty$, that is, M and $\inf_{i \in M} N_i$ simultaneously go to infinity.

Proposition 6.7 (Kingman algebraic two-level measure tree). *For $M \in \mathbb{N}$ and $\underline{N} \in \mathbb{N}^M$, let $\chi^{M,\underline{N}} = (T, c, \nu^{M,\underline{N}})$. Then their exists a random binary a2m tree $\chi \in \mathbb{T}_2^{(2)}$ such that*

(6.10) $$\chi^{M,\underline{N}} \underset{M, \underline{N} \to \infty}{\Longrightarrow} \chi,$$

where $\Rightarrow$ stands for weak convergence on $\mathbb{T}_2^{(2)}$ equipped with the two-level sample shape convergence.

Proof. By Proposition 2, the space $\mathbb{T}_2^{(2)}$ equipped with the two-level sample shape convergence is compact. Therefore the sequence $\{\chi^{M,\underline{N}} : M \in \mathbb{N}, \underline{N} \in \mathbb{N}^M\}$ is clearly tight.

The *uniqueness* of the limit follows from the sampling consistency of the family of finite nested Kingman coalescents. Fix $M \in \mathbb{N}$ and $\underline{N} \in \mathbb{N}^M$, and $m \in \mathbb{N}$, $\underline{n} \in \mathbb{N}^M$ such that $m \leq M$ and for each $1 \leq i \leq m$, $n_i \leq N_i$. Consider $\{U_{ij} : (i, j) \in \mathbb{N}_{m,\underline{n}}\}$ sampled (without repetition) at random from $\chi^{M,\underline{N}}$. By Proposition 6.2, the restriction of $\mathcal{K}^{\mathbb{N}_{M,\underline{N}}}$ to $\{U_{ij} : (i, j) \in \mathbb{N}_{m,\underline{n}}\}$ is a finite

nested Kingman coalescent on $\{U_{ij} : (i,j) \in \mathbb{N}_{m,\underline{n}}\}$. Therefore, the shape $\mathfrak{s}_{(T,c)}(\underline{U}) \in \mathfrak{C}_{m,\underline{n}}$ of the subtree spanned by the leaves $\{U_{ij} : (i,j) \in \mathbb{N}_{m,\underline{n}}\}$ has the distribution of the shape of $\chi^{m,\underline{n}}$. Since sampling with and without repetition is asymptotically equivalent when $M, \underline{N} \to \infty$, the claim follows by definition of the two-level sample shape convergence. $\square$

We can now conclude the paper with the construction of the Kingman algebraic two-level measure tree, which is obtained as the limit of finite nested Kingman coalescent trees and which is the algebraic tree corresponding to the two-level Kingman coalescent introduced in Meizis [Mei19].

Definition 6.8 (Kingman algebraic two-level measure tree)**.** *The* Kingman algebraic two-level measure tree χ *with rates* (γ_H, γ_P) *is the unique limit in* $\mathbb{T}_2^{(2)}$ *of the sequence* $(\chi^{M,\underline{N}})_{M,\underline{N}}$*, where* $\chi^{M,\underline{N}}$ *is the random algebraic two-level measure tree obtained from the finite nested Kingman coalescent on* $\mathbb{N}_{M,\underline{N}}$ *with rates* (γ_H, γ_P)*.*

REFERENCES

[Ald94a]   David Aldous. Recursive self-similarity for random trees, random triangulations and Brownian excursion. *Ann. Probab.*, 22(2):527–545, 1994. 5

[Ald94b]   David Aldous. Triangulating the circle, at random. *The American Mathematical Monthly*, 101(3):223–233, 1994. 5

[Ald00]   David Aldous. Mixing time for a Markov chain on cladograms. *Combinatorics, Probability and Computing*, 9:191–204, 2000. 3

[ALW17]   Siva Athreya, Wolfgang Löhr, and Anita Winter. Invariance principle for variable speed random walks on trees. *Ann. Probab.*, 45(2):625–667, 2017. 2, 10

[Ban08]   V. Bansaye. Proliferating parasites in dividing cells : Kimmel's branching model revisited. *Annals of Applied Probability*, 2008. 3

[BDLS18]   Airam Blancas Benítez, Jean-Jil Duchamps, Amaury Lambert, and Arno Siri-Jégousse. Trees within Trees: Simple Nested Coalescents. *Electron. J. Probab.*, 23:1–27, 2018. 34, 35

[BJ10]   M.G.B. Blum and M. Jakobsson. Deep Divergences of Human Gene Trees and Models of Human Origins. *Molecular Biology and Evolution*, 28(2):889–898, 10 2010. 3

[BRSSJ19]   Airam Blancas Benítez, Tim Rogers, Jason Schweinsberg, and Arno Siri-Jégousse. The Nested Kingman Coalescent: Speed of Coming Down from Infinity. *Annals of Applied Probability*, 29(3):1808–1836, February 2019. 24 pages. 34

[BT11]   V. Bansaye and V.C. Tran. Branching Feller diffusion for cell division with parasite infection. *ALEA*, 8:95–127, 2011. 3

[Chi01]   I. Chiswell. *Introduction to Λ-trees*. World Scientific, 2001. 2

[Daw18]   D. Dawson. Multilevel mutation-selection systems and set-valued duals. *J. Math. Biol.*, 76:295–378, 2018. 3, 34

[DGW04]   Donald Dawson, Luis Gorostiza, and Anton Wakolbinger. Hierarchical Equilibria of Branching Populations. *Electronic Journal of Probability*, 9(none):316 – 381, 2004. 3

[DHV95]   Donald A. Dawson, Kenneth J. Hochberg, and Vladimir Vinogradov. On path properties of super-2 processes ii. In M.G. Cranston and M.A. Pinsky, editors, *Proceedings of Symposia in Pure Mathematics Series*, volume 57, pages 385–403. AMS, Providence, 1995. 3

[DHV96]   Donald A. Dawson, Kenneth J. Hochberg, and Vladimir Vinogradov. High-density limits of hier-archically structured branching-diffusing populations. *Stochastic Processes and their Applications*, 62(2):191–222, July 1996. 3

[DHW90]   D.A. Dawson, K.J. Hochberg, and Y. Wu. *White Noise Analysis: Mathematics and Applications*, chapter Multilevel branching systems. World Scientific Publ., 1990. 3

[DMT96]   Andreas Dress, Vincent Moulton, and Werner Terhalle. T-theory: An overview. *European Journal of Combinatorics*, 17(2):161–175, 1996. 2

[Dre84]   Andreas W.M. Dress. Trees, tight extensions of metric spaces, and the cohomological dimension of certain groups: A note on combinatorical properties of metric spaces. *Adv. Math.*, 53:321–402, 1984. 2

[DT96]   A.W.M. Dress and W.F. Terhalle. The real tree. *Advances in Mathematics*, 120(2):283–301, 1996. 2

[EK86]   Stewart N. Ethier and Thomas G. Kurtz. *Markov Processes: Characterization and Convergence*. John Wiley, New York, 1986. 32

[Eva08]   Steven N. Evans. *Probability and real trees*, volume 1920 of *Lecture Notes in Mathematics*. Springer, Berlin, 2008. Lectures from the 35th Summer School on Probability Theory held in Saint-Flour, July 6–23, 2005. 2

[For]   Daniel J. Ford. Probabilities on cladograms: introduction to the alpha model. arXiv:math/0511246. 3

[GH00]   Andreas Greven and Kenneth Hochberg. New behavioral patterns for two-level branching systems. *Canadian Mathematical Society Conference Proceedings*, 6:205–215, 01 2000. 3

[GHW95]   Luis G. Gorostiza, Kenneth J. Hochberg, and Anton Wakolbinger. Persistence of a critical super-2 process. *Journal of Applied Probability*, 32(2):534–540, 1995. 3

[GPW09]   Andreas Greven, Peter Pfaffelhuber, and Anita Winter. Convergence in distribution of random metric measure spaces (Λ-coalescent measure trees). *Probab. Theo. Rel. Fields*, 145:285–322, 2009. 3, 6, 7

[Gro99]   Misha Gromov. *Metric structures for Riemannian and non-Riemannian spaces*, volume 152 of *Progress in Mathematics*. Birkhäuser Boston Inc., Boston, MA, 1999. 3, 6

[JBA19]   F. Jay, S. Boitard, and F. Austerlitz. An ABC method for whole-genome sequence data: inferring Paleolithic and Neolithic human expansions. *Molecular Biology and Evolution*, 36:1565–1579, 2019. 3

[Kim97]   M. Kimmel. Quasistationarity in a branching model of division-within-division. In *Classical and modern branching processes (Minneapolis, MN, 1994)*, volume 84 of *IMA Vol. Math. Appl.*, pages 157–164, New-York, 1997. Springer. 3

[Kin82]   J. F. C. Kingman. The coalescent. *Stochastic Process. Appl.*, 13(3):235–248, 1982. 34

[LBP+21] C. Lepers, S. Billiard, M. Porte, S. Méléard, and V.C. Tran. Inference with selection, varying population size and evolving population structure: Application of abc to a forward-backward. *Heredity*, 126:335–350, 2021. 3

[LMW20]   Wolfgang Löhr, Leonid Mytnik, and Anita Winter. The Aldous chain on cladograms in the diffusion limit. *The Annals of Probability*, 48:2565–2590, 09 2020. 3

[Löh13]   Wolfgang Löhr. Equivalence of Gromov-Prohorov- and Gromov's $\square_\lambda$-metric on the space of metric measure spaces. *Electronic Communications in Probability*, 18(none):1 – 10, 2013. 3

[LW21]   Wolgang Löhr and Anita Winter. Spaces of algebraic measure trees and triangulations of the circle. *Bulletin de la Société Mathématique de France*, 149(1):1–63, 2021. 1, 2, 3, 4, 5, 8, 9, 10, 11, 12, 13, 14, 15, 18, 19, 20, 23, 28, 31, 33

[Mei19]   Roland Meizis. Convergence of metric two-level measure spaces. *Stochastic Processes and their Applications*, 130, 10 2019. 3, 4, 5, 7, 8, 11, 13, 20, 24, 34, 37

[MNO92]   John C. Mayer, Jacek Nikiel, and Lex G. Oversteegen. Universal spaces for R-trees. *Transactions of the American Mathematical Society*, 334(1):411–432, 1992. 2

[MO90]   John Mayer and Lex Oversteegen. A topological characterization of R-trees. *Trans. Amer. Math. Soc.*, 320:395–415, 07 1990. 2

[MR13]   Sylvie Méléard and Sylvie Rœlly. Evolutive two-level population process and large population approximations. *Ann. Univ. Buchar. Math. Ser.*, 4(LXII)(1):37–70, 2013. 3

[Nus]   Josué Nussbaumer. Resampling dynamics on metric two-level measure trees. Manuscript in prep. 3

[NW]   Josué Nussbaumer and Anita Winter. The algebraic alpha-ford tree under evolution. arXiv:2006.09316. 3, 35

[Ter97]   W.F. Terhalle. R-trees and symmetric differences of sets. *Europ. J. Combinatorics*, 18:825–833, 1997. 2

[Tit77]   Jacques Tits. A "theorem of Lie-Kolchin" for trees. *Contributions to Algebra: A Collection of Papers Dedicated to Ellis Kolchin*, 12 1977. 2

[VAE+09] P. Verdu, F. Austerlitz, A. Estoup, R. Vitalis, M. Georges, S. Théry, A. Froment, S. Le Bomin, A. Gessain, J.-M. Hombert, L. Van der Veen, L. Quintana-Murci, S. Bahuchet, and E. Heyer. Origins and genetic diversity of Pygmy hunter-gatherers from Western Central Africa. *Current Biology*, 19(4):312–318, 2009. 3

[Wu91]   Yadong Wu. *Dynamic Particle Systems and Multilevel Measure Branching Processes*. PhD thesis, Carleton University, Ottawa, Canada, 1991. 3

[Wu94]   Y. Wu. Asymptotic behavior of the two-level measure branching process. *Annals of Probab.*, 22(2):854–874, 1994. 3

LAMA, Univ Gustave Eiffel, Univ Paris Est Creteil, CNRS, F-77454 Marne-la-Vallée, France.

Fakultät für Mathematik, Universität Duisburg-Essen, Germany.