The Annals of Applied Probability
2021, Vol. 31, No. 5, 2046–2090
https://doi.org/10.1214/20-AAP1641

EXCHANGEABLE COALESCENTS, ULTRAMETRIC SPACES, NESTED INTERVAL-PARTITIONS: A UNIFYING APPROACH

BY FÉLIX FOUTEL-RODIER[*], AMAURY LAMBERT[†] AND EMMANUEL SCHERTZER[‡]

Laboratoire de Probabilités, Statistiques & Modélisation, Sorbonne Université, and Center for Interdisciplinary Research in Biology, Collège de France, [*]*felix.foutel-rodier@college-de-france.fr;* [†]*amaury.lambert@sorbonne-universite.fr;* [‡]*emmanuel.schertzer@sorbonne-universite.fr*

Kingman's (1978) representation theorem (*J. Lond. Math. Soc. (2)* **18** (1978) 374–380) states that any exchangeable partition of $\mathbb{N}$ can be represented as a paintbox based on a random mass-partition. Similarly, any exchangeable composition (i.e., ordered partition of $\mathbb{N}$) can be represented as a paintbox based on an interval-partition (Gnedin (1997) *Ann. Probab.* **25** (1997) 1437–1450).

Our first main result is that any exchangeable coalescent process (not necessarily Markovian) can be represented as a paintbox based on a random nondecreasing process valued in interval-partitions, called nested interval-partition, generalizing the notion of comb metric space introduced in Lambert and Uribe Bravo (2017) (*p-Adic Numbers Ultrametric Anal. Appl.* **9** (2017) 22–38) to represent compact ultrametric spaces.

As a special case, we show that any Λ-coalescent can be obtained from a paintbox based on a unique random nested interval partition called Λ-comb, which is Markovian with explicit transitions. This nested interval-partition directly relates to the flow of bridges of Bertoin and Le Gall (2003) (*Probab. Theory Related Fields* **126** (2003) 261–288). We also display a particularly simple description of the so-called evolving coalescent (Pfaffelhuber and Wakolbinger (2006) *Stochastic Process. Appl.* **116** (2006) 1836–1859) by a comb-valued Markov process.

Next, we prove that any ultrametric measure space U, under mild measure-theoretic assumptions on U, is the leaf set of a tree composed of a separable subtree called the backbone, on which are grafted additional subtrees, which act as star-trees from the standpoint of sampling. Displaying this so-called weak isometry requires us to extend the Gromov-weak topology of Greven, Pfaffelhuber and Winter (2009) (*Probab. Theory Related Fields* **145** (2009) 285–322), that was initially designed for separable metric spaces, to nonseparable ultrametric spaces. It allows us to show that for any such ultrametric space U, there is a nested interval-partition which is (1) indistinguishable from U in the Gromov-weak topology; (2) weakly isometric to U if U has a complete backbone; (3) isometric to U if U is complete and separable.

1. Introduction.

1.1. *Ultrametric spaces and exchangeable coalescents.* In this paper we extend earlier work from Lambert and Uribe Bravo (2017) on the comb representation of ultrametric spaces. An ultrametric space is a metric space (U, d) such that the metric d fulfills the additional assumption

$$\forall x, y, z \in U, \quad d(x, y) \le \max\bigl(d(x, z), d(z, y)\bigr).$$

Received October 2018; revised October 2020.
MSC2020 subject classifications. Primary 60G09; secondary 60J35, 60C05, 54E70.
Key words and phrases. Combs, compositions, nested compositions, Λ-coalescents, flow of bridges, metric measure spaces, Gromov-weak topology.

In applications, ultrametric spaces are used to model the genealogy of entities co-existing at the same time. The distance between two points x and y of an ultrametric space is interpreted as the time to the most recent common ancestor (MRCA) of x and y. For instance, in population genetics ultrametric spaces model the genealogy of homologous genes in a population. Another example can be found in phylogenetics where ultrametric spaces are used to model the evolutionary relationships between species.

In population genetics and more generally in biology we do not have access to the entire population (that is to the entire ultrametric space) but only to a sample from the population. To model the procedure of sampling we equip the ultrametric space with a probability measure μ (also referred to as the sampling measure), yielding the notion of ultrametric measure spaces.

DEFINITION 1.1. A quadruple $(U, d, \mathscr{U}, \mu)$ is called an ultrametric measure space (UMS) if the following hold:

(i) The distance d is an ultrametric on U which is $\mathscr{U} \otimes \mathscr{U}$ measurable.
(ii) The measure μ is a probability measure defined on $\mathscr{U}$.
(iii) The σ-field $\mathscr{U}$ fulfills $\mathscr{U} \subseteq \mathscr{B}(U)$, where $\mathscr{B}(U)$ is the Borel σ-field of (U, d), and

$$\forall x \in U, \forall t > 0, \quad \{y \in U : d(x, y) < t\} \in \mathscr{U}.$$

If $\mathscr{U} = \mathscr{B}(U)$, we say that $(U, d, \mathscr{U}, \mu)$ is a Borel UMS.

REMARK 1.2. This definition might be surprising as we would naively expect a UMS to be any ultrametric space with a probability measure on its Borel σ-field. However, the previous naive definition is not satisfying for several reasons, that are exposed in Section 4.1. Notice that if (U, d) is separable, then the only σ-field $\mathscr{U}$ satisfying (iii) is the Borel σ-field, and thus condition (i) always holds. We thus recover the usual definition of an ultrametric measure space.

A sample from a UMS is an i.i.d. sequence $(X_i)_{i \geq 1}$ distributed according to μ. The genealogy of the sample is usually encoded as a partition-valued process, $(\Pi_t)_{t \geq 0}$ called a *coalescent*. For any time $t \geq 0$, the blocks of the partition Π_t are given by the following relation:

$$i \sim_{\Pi_t} j \quad \Longleftrightarrow \quad d(X_i, X_j) \leq t. \tag{1}$$

The process $(\Pi_t)_{t \geq 0}$ has two major features. First a well-known characteristic of ultrametric spaces is that for a given t the balls of radius t form a partition of the space that gets coarser as t increases. This implies that given $s \leq t$, the partition Π_t is coarser than Π_s. Second, if σ denotes a finite permutation of $\mathbb{N}$ and $\sigma(\Pi_t)$ is the partition of $\mathbb{N}$ whose blocks are the images by σ of the blocks of Π_t, we have

$$(\Pi_t)_{t \geq 0} \overset{(\mathrm{d})}{=} \left(\sigma(\Pi_t)\right)_{t \geq 0}.$$

We call any càdlàg partition valued process that fulfills these two conditions an *exchangeable coalescent* (note that the process $(\Pi_t)_{t \geq 0}$ is not necessarily Markovian).

1.2. *Combs in the compact case.*

Combs and ultrametric spaces. In this section, we address similar questions in the much simpler framework of comb metric spaces which have been introduced recently by Lambert

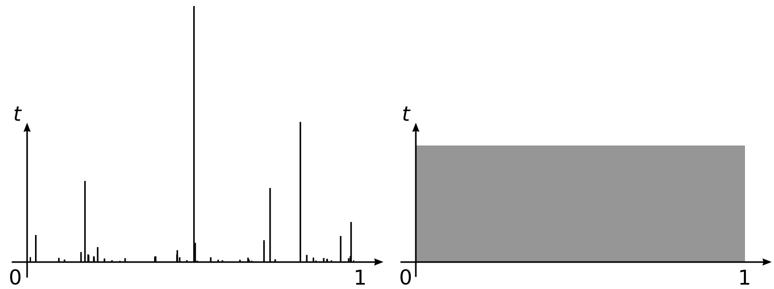

FIG. 1. *Representation of two nested interval-partitions. A point (x,t) is plotted in dark if $x \notin I_t$. Left panel: A realization of the Kingman comb, a tooth of size y at location x represents that $f(x) = y$. Right panel: The star-tree comb, an example of a nested interval-partition that cannot be represented as an original comb.*

and Uribe Bravo (2017) to represent *compact* ultrametric spaces. A comb is a function

$$f : [0,1] \to \mathbb{R}_+$$

such that for any $\varepsilon > 0$ the set $\{f \geq \varepsilon\}$ is finite (see Figure 1 left panel). To any comb is associated a comb metric d_f on $[0,1]$ defined as

$$\forall x, y \in [0,1], \quad d_f(x,y) = \mathbb{1}_{\{x \neq y\}} \sup_{[x\wedge y, x \vee y]} f.$$

In general d_f is only a pseudo-metric on $[0,1]$ and it is easy to verify that it is actually ultrametric. One of the main results in Lambert and Uribe Bravo (2017) shows that any compact ultrametric space is isometric to a properly completed and quotiented comb metric space (see Theorem 3.1 in Lambert and Uribe Bravo (2017)).

Exchangeable coalescents. We also will be interested in the relation between combs and exchangeable coalescents. Any comb metric space $([0,1], d_f)$ can be naturally endowed with the Lebesgue measure on $[0,1]$. Sampling from a comb can be seen as a direct extension of Kingman's paintbox procedure. More precisely, given a comb f, we can generate an exchangeable coalescent $(\Pi_t)_{t\geq 0}$ by throwing i.i.d. uniform random variables $(X_i)_{i\geq 1}$ on $[0,1]$ and declaring that

$$i \sim_{\Pi_t} j \quad \Longleftrightarrow \quad \sup_{[X_i \wedge X_j, X_i \vee X_j]} f \leq t.$$

For the sake of illustration, we recall the comb representation of the Kingman coalescent stated in Kingman (1982). The Kingman comb is constructed out of an i.i.d. sequence $(e_i)_{i\geq 1}$ of exponential variables with parameter 1, and of an independent i.i.d. sequence $(U_i)_{i\geq 1}$ of uniform variables on $[0,1]$. We define the sequence $(T_i)_{i\geq 2}$ as

$$T_i = \sum_{j \geq i} \frac{2}{j(j-1)} e_j.$$

The Kingman comb f_K is defined as

$$f_K = \sum_{i\geq 2} T_i \mathbb{1}_{U_i}.$$

See Figure 1 left panel for an illustration of a realization of the Kingman comb. The paintbox based on f_K is a version of the Kingman coalescent (see Section 4.1.3 of Bertoin (2006)).

More generally, the assumption that $\{f \geq \varepsilon\}$ is finite implies that the coalescent $(\Pi_t)_{t\geq 0}$ obtained from a paintbox based on f has only finitely many blocks for any $t > 0$. This property is usually referred to as "coming down from infinity." It has been shown in Lambert (2017) that any coalescent which comes down from infinity can be represented as a paintbox based on a comb, see Proposition 3.2.

1.3. *General combs.* One of the objectives of this work is to extend Theorem 3.1 of Lambert and Uribe Bravo (2017) and Proposition 3.2 of Lambert (2017) to any ultrametric space (not only compact) and to any exchangeable coalescent (i.e., beyond the "coming down from infinity" property). From a technical point of view, we note that this extension is conceptually harder, and requires the technology of exchangeable nested compositions which were absent in Lambert and Uribe Bravo (2017). This point will be discussed further in Section 2.1.

In order to deal with noncompact metric spaces, we need to generalize the definition of a comb by relaxing the condition on the finiteness of $\{f \geq \varepsilon\}$. We will encode combs as functions taking values in the open subsets of $(0, 1)$. Any open subset I of $(0, 1)$ can be decomposed into an at-most countable union of disjoint intervals denoted by $(I_i)_{i\geq 1}$. For this reason we will call an open subset of $(0, 1)$ an *interval-partition* and each of the intervals I_i is an *interval component* of I. The space of interval-partitions is conveniently topologized with the Hausdorff distance on the complement, d_H, defined as

$$d_H(I, \tilde{I}) = \sup\{d(x, [0, 1] \setminus \tilde{I}), x \notin I\} \vee \sup\{d(x, [0, 1] \setminus I), x \notin \tilde{I}\}.$$

We propose to generalize the notion of comb to the notion of *nested interval-partition*.

DEFINITION 1.3. A nested interval-partition is a càdlàg function $(I_t)_{t\geq 0}$ taking values in the open subsets of $(0, 1)$ verifying

$$\forall s \leq t, \quad I_s \subseteq I_t.$$

Sometimes nested interval-partitions will be called generalized combs or even simply combs.

Let us briefly see how this definition extends the initial comb of Lambert and Uribe Bravo (2017). Starting from a comb function f, we can build a nested interval-partition $(I_t)_{t\geq 0}$ as follows:

$$\forall t > 0, \quad I_t = \{f < t\} \setminus \{0, 1\}$$

and

$$I_0 = \operatorname{int}(\{f = 0\}),$$

where $\operatorname{int}(A)$ denotes the interior of the set A.

Conversely if $(I_t)_{t\geq 0}$ is a nested interval-partition we can define a comb function $f_I : [0, 1] \to \mathbb{R}_+$ as

$$f_I(x) = \inf\{t \geq 0 : x \in I_t\}.$$

In general f_I does not fulfill that $\{f_I \geq t\}$ is finite. A necessary and sufficient condition for this to hold is that for any $t > 0$, I_t has finitely many interval components, and the summation of their lengths is 1. If the latter condition is fulfilled, we say that I_t is proper or equivalently that it has no dust.

A nested interval-partition naturally encodes a (pseudo-)ultrametric d_I on $[0, 1]$ defined as

$$\begin{aligned} d_I(x, y) &= \inf\{t \geq 0 : x \text{ and } y \text{ belong to the same interval of } I_t\} \\ &= \sup_{[x,y]} f_I \end{aligned}$$

for $x < y$. We call the ultrametric space $([0, 1], d_I)$ the *comb metric space* associated to $(I_t)_{t\geq 0}$. In order to turn $([0, 1], d_I)$ into a UMS, we need to define an appropriate σ-field and a sampling measure. The interval $[0, 1]$ is naturally endowed with the usual Borel σ-field $\mathscr{B}([0, 1])$ and the Lebesgue measure. However, the usual Borel σ-field does not fulfill the requirements of Definition 1.1 in general because two points that belong to the same interval component of I_0 are indistinguishable in the metric d_I. This can be addressed by considering a slightly smaller σ-field as follows.

Let $(I_i^0)_{i\geq 1}$ be the interval components of I_0. We define a σ-field $\mathscr{I}$ on $[0,1]$ as

$$\mathscr{I} = \Big\{ A \cup \bigcup_{i\in M} I_i^0 : A \in \mathscr{B}\big([0,1]\setminus I_0\big) \text{ and } M \subseteq \mathbb{N} \Big\},$$

where $\mathscr{B}([0,1]\setminus I_0)$ denotes the usual Borel σ-field on $[0,1]\setminus I_0$. It is clear that $\mathscr{I} \subseteq \mathscr{B}([0,1])$. We call a *comb metric measure space* associated to $(I_t)_{t\geq 0}$ the quadruple $([0,1], d_I, \mathscr{I}, \mathrm{Leb})$, where Leb is the restriction of the Lebesgue measure to $\mathscr{I}$. The following lemma shows that the Lebesgue measure on $\mathscr{I}$ satisfies the requirements of Definition 1.1, and that a comb metric measure space is a UMS.

LEMMA 1.4. *Any comb metric measure space* $([0,1], d_I, \mathscr{I}, \mathrm{Leb})$ *is a UMS.*

PROOF. Let us first prove that (iii) holds. For $x \in [0,1]$ and $t \geq 0$, let $I_t(x)$ denote the interval component of I_t to which x belongs if $x \in I_t$, or let $I_t(x) = \{x\}$ else. Then for $t > 0$ we have

$$\big\{y \in [0,1] : d_I(x,y) < t\big\} = \bigcup_{s<t} I_s(x) \in \mathscr{I}.$$

It remains to show that $\mathscr{I} \subseteq \mathscr{B}_I([0,1])$, where $\mathscr{B}_I([0,1])$ denotes the σ-field induced by d_I. It is sufficient to prove that for all $x, y \notin I_0$, we have $(x,y) \in \mathscr{B}_I([0,1])$. Let $z \in (x,y)$ and suppose that $z \in I_t$ for all $t > 0$. Then $I_{t_z}(z) \subseteq (x,y)$ for a small enough t_z, and thus

$$\big\{z' \in [0,1] : d_I\big(z,z'\big) < t_z\big\} \subseteq (x,y).$$

Otherwise if $z \notin I_{t_z}(z)$ for some t_z, then $\{z' \in [0,1] : d_I(z,z') < t_z\} = \{z\}$. We can now write

$$(x,y) = \bigcup_{z\in(x,y)} \big\{z' \in [0,1] : d_I\big(z,z'\big) < t_z\big\} \in \mathscr{B}_I\big([0,1]\big)$$

which proves that point (iii) of the definition is fulfilled.

Let $I_{t-} = \bigcup_{s<t} I_s$, then

$$\big\{(x,y) \in [0,1]^2 : d(x,y) < t\big\} = \Delta_0 \cup \bigcup_{x\in I_{t-}} I_t(x) \times I_t(x),$$

where $\Delta_0 = \{(x,y) \in ([0,1]\setminus I_0)^2 : x = y\}$. As there are only countably many interval components of I_t, the union on the right-hand side is countable, and this set belongs to the product $\mathscr{I} \otimes \mathscr{I}$. This proves that point (i) holds and that the comb metric measure space is a UMS. $\square$

For later purpose, let us denote by U_I the completion of the quotient space of $\{f_I = 0\}$ by the relation $x \sim y$ iff $d_I(x,y) = 0$. (This completion can be realized explicitly by adding countably many "left" and "right" faces to the comb, see Section 4.5.)

Finally, as in the compact case, an exchangeable coalescent $(\Pi_t)_{t\geq 0}$ can be obtained from a nested interval-partition $(I_t)_{t\geq 0}$ out of an i.i.d. uniform sequence $(X_i)_{i\geq 1}$ by defining

(2) $$i \sim_{\Pi_t} j \quad \Longleftrightarrow \quad X_i \text{ and } X_j \text{ belong to the same interval component of } I_t.$$

Notice that this definition is a multidimensional extension of the original Kingman paintbox procedure; see, for example, the beginning of Section 2.3.2 of Bertoin (2006).

REMARK 1.5. The coalescent obtained through this sampling procedure is not càdlàg in general. As a coalescent is a nondecreasing process, we can (and will) always suppose that we work with a càdlàg modification of the coalescent.

REMARK 1.6. We have defined two natural ways of sampling a coalescent from a nested interval-partition. First, one can realize the extended paintbox procedure described in equation (2). Second, one can consider the comb metric measure space associated to the nested interval-partition and sample the coalescent according to equation (1). It is not hard to see that the coalescent obtained through (1) is the càdlàg version of the one obtained through (2).

We will now demonstrate that nested interval-partitions form a large enough framework to answer our two initial problems: representing any exchangeable coalescent as a paintbox on a comb and representing general ultrametric measure spaces.

1.4. *Comb representation of exchangeable coalescents.*

General comb representation. We start by showing that one can always find a comb representation of any coalescent. First notice that this representation cannot be unique. For example taking the reflection of a comb about the vertical line in the middle of the segment [0, 1] yields a new comb but does not change the associated coalescent. In many applications we will not be interested in this order but only in the genealogical structure of the comb. For this reason we introduce the following relation.

DEFINITION 1.7. Two generalized combs are paintbox-equivalent if their associated coalescents are identical in law. Being paintbox-equivalent is an equivalence relation, we denote by $\mathfrak{I}$ the quotient space.

Given $I \in \mathfrak{I}$ we denote by ρ_I the distribution on the space of coalescents of the paintbox based on any representative of $\mathfrak{I}$. We provide the following version of Kingman's representation theorem (e.g., see Bertoin ((2006), Theorem 2.1)) for exchangeable coalescents.

THEOREM 1.8. *Let $(\Pi_t)_{t\geq 0}$ be an exchangeable coalescent. There exists a unique distribution ν on $\mathfrak{I}$ such that*

$$\mathbb{P}\big((\Pi_t)_{t\geq 0} \in \cdot\big) = \int_{\mathfrak{I}} \rho_I(\cdot)\nu(\mathrm{d}I).$$

REMARK 1.9. It is interesting to relate this result to the original theorem from Kingman. A *mass-partition* is a sequence $\beta = (\beta_i)_{i\geq 1}$ such that

$$\beta_1 \geq \beta_2 \geq \cdots \geq 0, \qquad \sum_{i\geq 1} \beta_i \leq 1.$$

Kingman's representation theorem states that any exchangeable partition can be obtained through a paintbox based on a random mass-partition, and that this correspondence is bijective. A mass-partition can be seen as the ranked sequence of the lengths of the interval components of an interval-partition. Now notice that two interval-partitions are paintbox-equivalent, that is, induce the same exchangeable partition, iff they have the same associated mass-partition. In this one-dimensional setting, any paintbox-equivalence class of interval-partitions can be identified with a random mass-partition. In a similar way, it would be natural to try to identify the elements of $\mathfrak{I}$ with mass-partition valued processes, also called *mass-coalescents*. However, one can easily find two different equivalence classes of $\mathfrak{I}$ that have the same associated mass-coalescent, see Figure 2.

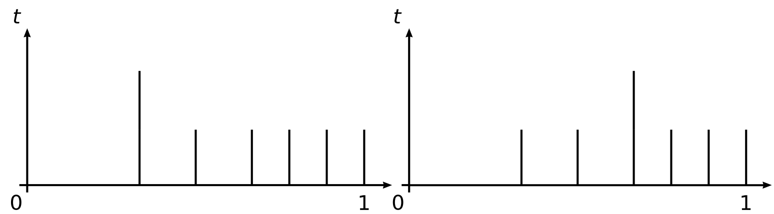

FIG. 2. *An example of two nested interval-partitions that have the same mass-coalescent but different coalescents. For both processes, the initial mass-partition is* $(\frac{1}{3}, \frac{1}{6}, \frac{1}{6}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, 0, \ldots)$, *then* $(\frac{2}{3}, \frac{1}{3}, 0, \ldots)$ *and finally* $(1, 0, \ldots)$. *However, for the process on the left-hand side the first blocks to merge are those of mass* $1/6$ *and* $1/9$, *whereas for the right-hand process, the blocks of mass* $1/6$ *first merge with the block of size* $1/3$.

REMARK 1.10. A result very similar to Theorem 1.8 has been obtained in Forman, Haulk and Pitman ((2018), Theorem 4), in the context of hierarchies. Roughly speaking, an exchangeable hierarchy is obtained from an exchangeable coalescent by "forgetting about time." In this sense, an exchangeable coalescent carries more information, and this part of our work can be seen as an extension of Forman, Haulk and Pitman (2018). However, the forthcoming Section 3 and Section 4 heavily rely on the knowledge of the coalescence times, and could not have been achieved in the framework of hierarchies. We have dedicated Appendix A to the explanation of the links between the present work and Forman, Haulk and Pitman (2018).

Λ-*Coalescents*. Most of the efforts made in the study of exchangeable coalescents have been devoted to the special case of Λ-coalescents (Pitman (1999), Sagitov (1999)). These coalescents are parametrized by a finite measure Λ on $[0, 1]$, and their restriction to $[n] := \{1, \ldots, n\}$ is a Markov chain whose transitions are the following. The process undergoes a transition from a partition π with b blocks to a partition obtained by merging k blocks of π at rate $\lambda_{b,k}$ given by

$$\lambda_{b,k} = \int_{[0,1]} x^{k-2}(1-x)^{b-k} \Lambda(\mathrm{d}x).$$

The next proposition states that we can always find a Markovian comb representation of a Λ-coalescent. Moreover in Section 3 we provide an explicit description of its transition.

PROPOSITION 1.11. *Let* $(\Pi_t)_{t \geq 0}$ *be a* Λ-*coalescent. There exists* $(I_t)_{t \geq 0}$ *a Markov nested interval-partition such that the coalescent obtained from the paintbox based on* $(I_t)_{t \geq 0}$ *is distributed as* $(\Pi_t)_{t \geq 0}$.

REMARK 1.12 (Combs and the flow of bridges). The flow of bridges introduced by Bertoin and Le Gall (2003) represents the dynamics of a population whose genealogy is given by a Λ-coalescent. We will show that we can build a nested interval-partition from the flow of bridges and that it has the same distribution as the Markov nested interval-partition of Proposition 1.11, see Section 3.

REMARK 1.13. There exists a natural extension of the Λ-coalescents called the coalescents with simultaneous multiple collisions or Ξ-coalescents (Schweinsberg (2000)). All our results carry over to Ξ-coalescents, however, for the sake of clarity we will focus on the case of Λ-coalescents.

A coalescent process models the genealogy of a population living at a fixed observation time. Many works have been concerned with the dynamical genealogy obtained by varying the observation time of the population. For example, in Pfaffelhuber and Wakolbinger (2006), Pfaffelhuber, Wakolbinger and Weisshaupt (2011) the authors study some statistics of the dynamical genealogy, namely the time to the MRCA and the total length of the genealogy. In Greven, Pfaffelhuber and Winter (2013) the genealogy is encoded as a metric space (a real tree, see Evans (2008)) and the authors introduce the tree-valued Fleming–Viot process, a process bearing the entire information on the dynamical genealogy. This encoding requires to work with metric space-valued stochastic processes, and with the rather technical Gromov-weak topology for metric spaces.

We address such questions in the framework of combs in Section 3.3. We show that we can naturally encode a dynamical genealogy as a comb-valued process called the *evolving comb*. This process is a Markov process, whose semigroup can be explicitly described. In the particular case of coalescents that come down from infinity, the semigroup of the evolving comb takes a particularly simple form in terms of sampling from an independent comb.

1.5. *Comb representation of ultrametric spaces*. The second main aim of this paper is to provide a comb representation of ultrametric measure spaces in the same vein as Theorem 3.1 of Lambert and Uribe Bravo (2017). We will only state our results informally and refer to Section 4 for the precise statements.

We first introduce the *Gromov-weak topology* on the space of UMS and show that any UMS is indistinguishable from a comb metric space in this topology. To do so, we realize a straightforward extension of the work developed in Greven, Pfaffelhuber and Winter (2009), Gromov (1999) which is focused on separable metric measure spaces. In short, starting from a UMS we can obtain a coalescent by sampling from it as described in Section 1.1. This coalescent can be seen as a random ultrametric on $\mathbb{N}$ called the distance matrix of the UMS, see Section 4.2. We say that a sequence of UMS converges to a limiting UMS in the Gromov-weak sense if the corresponding distance matrices converge weakly (see Section 4.2 for a more precise definition). We are now ready to state our representation result, which is a direct application of Theorem 1.8.

THEOREM 1.14. *For any UMS $(U, d, \mathscr{U}, \mu)$ there exists a comb metric measure space that is indistinguishable in the Gromov-weak topology from $(U, d, \mathscr{U}, \mu)$.*

PROOF. As we have identified any UMS with the distribution of its distance matrix, two UMS are indistinguishable iff their distance matrices have the same distribution, or, equivalently, iff their coalescents have the same distribution. Theorem 1.8 shows that we can always find a nested interval-partition $(I_t)_{t\geq 0}$ such that the coalescent obtained from a paintbox based on $(I_t)_{t\geq 0}$ is distributed as the coalescent obtained by sampling from $(U, d, \mathscr{U}, \mu)$. As noticed in Remark 1.6, the coalescent obtained by sampling in the comb metric measure space $([0, 1], d_I, \mathscr{I}, \mathrm{Leb})$ has the same distribution as the coalescent obtained from the paintbox based on $(I_t)_{t\geq 0}$, and thus this comb metric measure space is indistinguishable from $(U, d, \mathscr{U}, \mu)$. □

The comb representation given by Theorem 1.14 is rather weak, since it only ensures that we can find a comb that has the same sampling structure as a given UMS. We would like to be more precise and obtain an isometry result as in the compact case. This is not possible in general, and we have to consider separately the separable case and the nonseparable case.

The separable case. In the separable case, the coalescent contains all the information about the UMS. More precisely, the Gromov reconstruction theorem ensures that two complete separable UMS that are indistinguishable in the Gromov-weak topology have the supports of their measures in isometry; see, for example, Gromov ((1999), Section $3.\frac{1}{2}.5$) or Greven, Pfaffelhuber and Winter ((2009), Proposition 2.6). The following refinement of Theorem 1.14 in the separable case is a direct consequence of the Gromov reconstruction theorem and of Theorem 1.14; see Section 4.6 for a proof.

COROLLARY 1.15. *Let* $(U, d, \mathscr{U}, \mu)$ *be a complete separable UMS. There exists a comb metric measure space* $(U_I, d_I, \mathscr{I}, \text{Leb})$ *such that the support of* μ *is isometric to* (U_I, d_I), *and such that the isometry maps* μ *to* Leb.

Additionally, any separable ultrametric space (U, d) can be endowed with a probability measure whose support is the whole space U, see Lemma 4.18. This result combined with Corollary 1.15 yields the following representation result for complete separable ultrametric spaces, which is the direct extension of Theorem 3.1 of Lambert and Uribe Bravo (2017) to the separable case.

PROPOSITION 1.16. *Let* (U, d) *be a complete separable ultrametric space. We can find a nested interval-partition such that* (U_I, d_I) *is isometric to* (U, d).

A proof of this proposition is provided in Section 4.6. Notice that the proof of the previous proposition is very different from the original proof of Lambert and Uribe Bravo (2017) which is no longer valid for noncompact UMS.

The general case. In general, two UMS that are associated to the same coalescent are not isometric. This essentially comes from the fact that a coalescent only bears the information about a sequence of "typical" points of the UMS, and that a nonseparable UMS may contain more information than the topology generated by these "typical" points. The main idea of our approach relies on a new decomposition that we now expose.

A UMS $(U, d, \mathscr{U}, \mu)$ can be seen as the leaves of a tree. We show that we can decompose this tree into two parts. The first part is a separable tree that we call the backbone. Second, one can then recover the tree from the backbone by grafting some "simple" subtrees on the backbone. By "simple," we mean that each of those subtrees has the sampling properties of a star-tree, in the sense that all points sampled in the same subtree are at the same distance to each other. See Figure 3 for an illustration of this decomposition, and Definition 4.8 for a precise definition of the backbone. An object very similar to the backbone is studied in Gufler (2018) but the construction of the backbone from a general UMS is not considered there.

Our result states that if two UMS have complete backbones and are associated to the same coalescent, then the backbones are in isometry in a way that preserves the star-trees attached to it. We say that the two UMS are in *weak isometry*, see Definition 4.11. We provide the following version of the Gromov reconstruction theorem in the case of general UMS.

PROPOSITION 1.17. *Let* $(U, d, \mathscr{U}, \mu)$ *and* $(U', d', \mathscr{U}', \mu')$ *be two UMS with complete backbones. These UMS are indistinguishable in the Gromov-weak topology iff* $(U, d, \mathscr{U}, \mu)$ *and* $(U', d', \mathscr{U}', \mu')$ *are in weak isometry.*

An equivalent reformulation of the previous proposition is stated in Section 4.4, see Proposition 4.12, and proved at the end of Section 4.4. As a consequence of Proposition 1.17 and Theorem 1.14, we have the following version of Theorem 3.1 of Lambert and Uribe Bravo (2017) in the general case. See Section 4.5 for a proof.

COROLLARY 1.18. *Let* $(U, d, \mathscr{U}, \mu)$ *be a UMS with a complete backbone. There exists a nested interval-partition* $(I_t)_{t\geq 0}$ *such that, up to the addition of a countable number of points, the comb metric measure space* $([0, 1], d_I, \mathscr{I}, \text{Leb})$ *is weakly isometric to* $(U, d, \mathscr{U}, \mu)$.

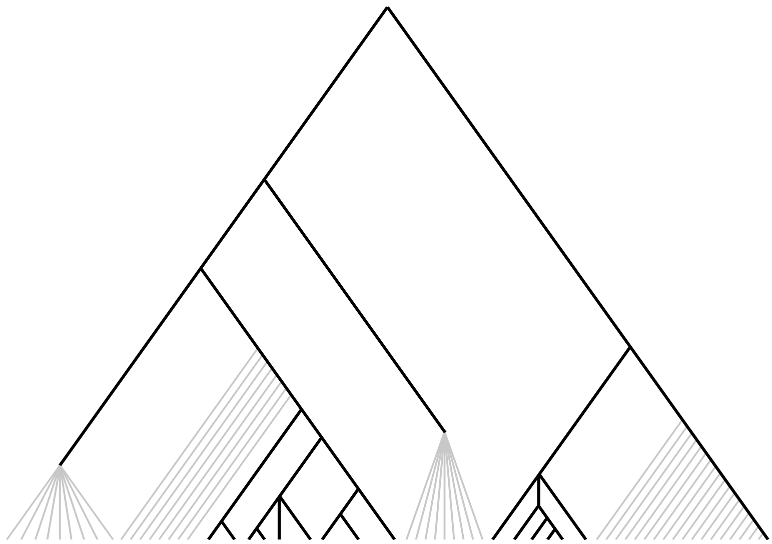

FIG. 3. *Illustration of the backbone decomposition. The dark thick lines represent the backbone. An element of the tree is represented in grey if its descendance has zero mass.*

1.6. *Outline*. The rest of the paper is divided into three parts. In Section 2 we introduce the notion of composition and nested composition which will be our main tool to study combs. Section 2.1 introduces the existing material on random compositions. In Section 2.2 we define exchangeable nested compositions and prove the representation theorem linking combs and nested compositions. The proof of Theorem 1.8 is given in Section 2.3. In Section 3 we restrict our attention to the case of Λ-coalescents. We define there the notion of a Λ-comb and study a family of nested compositions emerging from the Λ-coalescents. The proof of Proposition 1.11 is given in Section 3.2. The evolving comb is introduced and studied in Section 3.3. Finally in Section 4 we envision combs as ultrametric spaces. A precise outline of this section is given at the beginning of Section 4.

2. Combs and nested compositions. The objective of this section is to prove Theorem 1.8 on the comb representation of exchangeable coalescents. As was already mentioned in the Introduction, the correspondence between combs and exchangeable coalescents cannot be bijective. Roughly speaking, this comes from the fact that a nested interval-partition inherits an order from $[0, 1]$, and that changing this order does not modify the associated coalescent. However, we will show in Section 2.2 that there is a bijective correspondence between nested interval-partitions and exchangeable nested compositions, the ordered version of exchangeable coalescents. Exchangeable nested compositions will be our main tool to study combs.

We start this section by recalling existing results and material on exchangeable compositions developed in Donnelly and Joyce (1991), Gnedin (1997) and then show how to extend them to nested compositions.

2.1. *Exchangeable compositions*. In combinatorics, a composition of $[n]$ (resp. $\mathbb{N}$) is a partition of $[n]$ (resp. $\mathbb{N}$) with a total order on the blocks. We write $\mathcal{C} = (\pi, \leq)$ for a composition of $\mathbb{N}$ where π is the partition and $\leq$ the order on the blocks. The blocks of the partition π can always be labeled in increasing order of their least element, that is, the blocks of π are denoted by $(A_1, A_2, \ldots)$ and are such that for any $i, j \geq 1$,

$$i \leq j \iff \min(A_i) \leq \min(A_j).$$

Let σ be a finite permutation of $\mathbb{N}$, we denote by $\sigma(\mathcal{C})$ the composition whose blocks are $(\sigma(A_1), \sigma(A_2), \dots)$ and such that the order of the blocks is

$$\sigma(A_i) \le \sigma(A_j) \quad \iff \quad A_i \le A_j.$$

For example, for $n = 5$, consider $\mathcal{C}^n$ the composition

$$\mathcal{C}^n = \{2, 3\} \le \{5\} \le \{1, 4\}.$$

With our labeling convention, we have $A_1 = \{1, 4\}$, $A_2 = \{2, 3\}$ and $A_3 = \{5\}$ (A_1 needs not be the first block of $\mathcal{C}$ for the order $\le$). If $\sigma = (2, 1, 3, 5, 4)$, the composition $\sigma(\mathcal{C}^n)$ is given by

$$\sigma(\mathcal{C}^n) = \{1, 3\} \le \{4\} \le \{2, 5\}.$$

A random composition $\mathcal{C}$ of $\mathbb{N}$ is called *exchangeable* if for any finite permutation σ,

$$\mathcal{C} \overset{(d)}{=} \sigma(\mathcal{C}).$$

Gnedin (1997) provides a procedure to build an exchangeable composition of $\mathbb{N}$ from any interval-partition I called the *ordered paintbox*. Let $(V_i)_{i\ge1}$ be an i.i.d. sequence of uniform [0, 1] variables. Let $\mathcal{C}$ be the composition of $\mathbb{N}$ whose blocks are given by the relation

$$i \sim j \quad \iff \quad V_i \text{ and } V_j \text{ belong to the same interval component of } I$$

and the order of the blocks is

$$A \le A' \quad \iff \quad V_i \le V_j, \quad \forall i \in A, \forall j \in A'.$$

The main result of Gnedin (1997) shows that any exchangeable composition of $\mathbb{N}$ can be obtained as an ordered paintbox based on a random interval-partition (see Theorem 11 in Gnedin (1997)). We now give a proof of this result that differs from the original proof of Gnedin (1997). We make use of de Finetti's theorem in a similar way as Aldous' proof of Kingman's theorem; see, for example, the proof of Theorem 2.1 in Bertoin (2006). The original proof of Gnedin (1997) relies on a reversed martingale argument combined with the method of moments.

THEOREM 2.1 (Gnedin). *Let $\mathcal{C}$ be an exchangeable composition of $\mathbb{N}$. There exists on the same probability space a random interval-partition I and an independent i.i.d. sequence $(V_i)_{i\ge1}$ of uniform* [0, 1] *variables such that the ordered paintbox based on I by the sequence $(V_i)_{i\ge1}$ is a.s. $\mathcal{C}$.*

Before showing the theorem we need a technical lemma. Any composition $\mathcal{C} = (\pi, \le)$ can be encoded as a total preorder $\preceq$ on $\mathbb{N}$ defined as

$$i \preceq j \quad \iff \quad B_i \le B_j,$$

where B_i (resp. B_j) is the block containing i (resp. j). The blocks of π can be recovered from $\preceq$ by the following relation:

$$i \sim j \quad \iff \quad i \preceq j \quad \text{and} \quad j \preceq i$$

and the order $\le$ by

$$B \le B' \quad \iff \quad i \preceq j, \quad \forall i \in B, \forall j \in B'.$$

LEMMA 2.2. *Let $\mathcal{C}$ be an exchangeable composition of $\mathbb{N}$. We can find an exchangeable sequence of* [0, 1]*-valued random variables $(\xi_i)_{i\ge1}$ such that*

$$i \preceq j \quad \iff \quad \xi_i \le \xi_j.$$

PROOF. Let D_i be the set of integers lower than i

$$D_i = \{k : k \preceq i\}.$$

It is immediate that the partition $(D_i \setminus \{i\}, \mathbb{N} \setminus \{i\} \setminus D_i)$ is an exchangeable partition of $\mathbb{N} \setminus \{i\}$. Thus Kingman's representation theorem (see, e.g., Theorem 2.1 in Bertoin (2006)) ensures that the limit

$$\xi_i = \lim_{n\to\infty} \frac{1}{n} \operatorname{Card}\bigl(D_i \cap [n]\bigr)$$

exists a.s. Fix a finite permutation σ whose support lies in $[n]$, that is, such that $\sigma(i) = i$ for $i \geq n$. For $m \geq n$, the distribution of $(\operatorname{Card}(D_i \cap [m]))_{i\geq 1}$ is invariant by the action of σ. Taking the limit, the distribution of the sequence $(\xi_i)_{i\geq 1}$ is also invariant by the action of σ, and thus it is an exchangeable sequence.

We need to show that

$$i \preceq j \quad \Longleftrightarrow \quad \xi_i \leq \xi_j.$$

The only difficulty here is to show that $\xi_i \leq \xi_j$ implies $i \preceq j$. Suppose that $i \npreceq j$, we need to show that

$$\xi_i - \xi_j = \lim_{n\to\infty} \frac{1}{n} \operatorname{Card}\bigl((D_i \setminus D_j) \cap [n]\bigr) > 0.$$

The partition $(D_j \setminus \{i, j\}, D_i \setminus \{i, j\} \setminus D_j, \mathbb{N} \setminus \{i, j\} \setminus D_i)$ is an exchangeable partition of $\mathbb{N} \setminus \{i, j\}$. Another interesting consequence of Kingman's theorem is that in any exchangeable partition, the blocks are either singletons or have positive asymptotic frequencies. According to this, it is sufficient to show that a.s. $D_i \setminus D_j$ has at least two elements that are not i. Consider B_i (resp. B_j) the block to which i (resp. j) belongs. The set $D_i \setminus D_j$ is the reunion of all the blocks B such that $B_j < B \leq B_i$. Thus $D_i \setminus D_j$ is a singleton iff $B_i = \{i\}$ and there exists at most one singleton block B such that $B_j < B < B_i$. Let $n \geq 1$ and consider the block sizes and order of $\mathcal{C}^n$ as fixed. Exchangeability shows that the labels inside the blocks are chosen uniformly among all the possibilities. In particular this shows that the probability that $(D_i \setminus D_j) \cap [n]$ is a singleton goes to 0 as n goes to infinity. □

Now Theorem 2.1 is essentially a corollary of the previous lemma and of de Finetti's theorem.

PROOF OF THEOREM 2.1. Let $(\xi_i)_{i\geq 1}$ be as above. Applying de Finetti's theorem we know that there exists a random measure μ such that conditionally on it the sequence $(\xi_i)_{i\geq 1}$ is i.i.d. distributed as μ. Consider the distribution function F_μ of μ, and its generalized inverse

$$F_\mu^{-1}(x) = \inf\{r : F_\mu(r) > x\}.$$

The interval-partition associated with μ, I_μ, is defined as the set of flats of F_μ^{-1}:

$$I_\mu = \{x \in [0, 1] : \exists y < x < z, F_\mu^{-1}(y) = F_\mu^{-1}(z)\}.$$

The measure μ has the property that if X is distributed as μ, then μ-a.s. $F_\mu(X) = X$. Conditioning on μ, this can be seen from the definition of the sequence $(\xi_i)_{i\geq 1}$ and the law of large numbers:

$$F_\mu(\xi_1) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \mathbb{1}_{\{\xi_j \leq \xi_1\}} = \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \mathbb{1}_{\{j \preceq 1\}} = \xi_1 \quad \mu\text{-a.s.}$$

In the terminology of Gnedin (1997) this shows that the measure μ is *uniformized*. A uniformized measure has an atomic and a diffuse part. The support of the diffuse part is $[0, 1] \setminus I_\mu$ and coincides with the Lebesgue measure. The atomic part is supported by the right endpoints of the interval components of I_μ. If $J = (\ell, r)$ is an interval component of I_μ, the measure μ has an atom of mass $r - \ell$ located at r.

Let $(J_k)_{k\geq 1}$ be the interval decomposition of I_μ, and write $J_k = (\ell_k, r_k)$. Let $(X_i)_{i\geq 1}$ be an independent i.i.d. sequence of uniform variables, we define

$$V_i = \begin{cases} \xi_i & \text{if } \xi_i \notin I_\mu, \\ (r_k - \ell_k)X_i + \ell_k & \text{if } \xi_i = r_k. \end{cases}$$

In words, the variables from the sequence $(\xi_i)_{i\geq 1}$ which are equal to the atom r_k are uniformly dispersed over the interval J_k. The previous remarks on the structure of uniformized measures show that conditionally on μ, the sequence $(V_i)_{i\geq 1}$ is i.i.d. uniform on $[0, 1]$. The conditional distribution does not depend on μ, thus the sequence $(V_i)_{i\geq 1}$ is independent of μ and of I_μ.

We only need to show that the ordered paintbox based on I_μ using the sequence $(V_i)_{i\geq 1}$ is $\mathcal{C}$ a.s. This is plain from the design of the sequence. □

We end this section with a technical result already present in Gnedin (1997) (see Proposition 9) which we will require. Let $\mathcal{C}$ be an exchangeable composition of $\mathbb{N}$ and $\mathcal{C}^n$ its restriction to $[n]$. Let us denote by n_i the size of the ith block of $\mathcal{C}^n$. The *empirical interval-partition* associated to $\mathcal{C}_n$ is given by

$$I^n = \left(0, \frac{n_1}{n}\right) \cup \left(\frac{n_1}{n}, \frac{n_1 + n_2}{n}\right) \cup \cdots \cup \left(\frac{n_1 + \cdots + n_{k-1}}{n}, 1\right).$$

Here is a more pictorial way of constructing I^n. Divide $[0, 1]$ in intervals of size $1/n$ and label them from 1 to n in such a way that $i \preceq j$ iff the block with label i is before the block with label j. Then I^n is obtained by merging the intervals whose labels are in the same block of the composition. The next result states that the interval-partition representing $\mathcal{C}$ in Theorem 2.1 can be obtained as the limit of the empirical interval-partitions.

PROPOSITION 2.3. *If $\mathcal{C}$ is an exchangeable composition of $\mathbb{N}$, I the interval-partition obtained from Theorem 2.1 and $(I^n)_{n\geq 1}$ the sequence of empirical interval-partitions associated to $\mathcal{C}$, we have*

$$\lim_{n\to\infty} d_H\big(I^n, I \setminus \{0, 1\}\big) = 0 \quad \textit{a.s.}$$

PROOF. Let μ, $(\xi_i)_{i\geq 1}$ and I_μ be as in the proof of Theorem 2.1. De Finetti's theorem ensures that

$$\lim_{n\to\infty} \mu_n := \frac{1}{n} \sum_{i=1}^{n} \delta_{\xi_i} = \mu \quad \text{a.s.}$$

in the sense of weak convergence of probability measures. The interval-partition I_{μ_n} coincides with the empirical interval-partition I^n and as was already noticed in Gnedin (1997), the weak convergence of μ_n to μ implies the convergence of I_{μ_n} to I in the Hausdorff topology. □

REMARK 2.4. This also shows that the representation obtained through Theorem 2.1 is unique in distribution. The interval-partition I is a.s. recovered from I^n whose distribution is fully determined by $\mathcal{C}$.

2.2. *Exchangeable nested compositions.* Gnedin's theorem sets up a correspondence between random interval-partitions and exchangeable compositions. We want to find a similar correspondence between nested interval-partitions and exchangeable nested compositions, the ordered version of exchangeable coalescents. A nested composition of $[n]$ (resp. $\mathbb{N}$) is a càdlàg process $(\mathcal{C}_t)_{t\geq 0}$ taking values in the compositions of $[n]$ (resp. $\mathbb{N}$) such that, as t increases, only adjacent blocks of the composition merge. More precisely, if $(\mathcal{C}_t)_{t\geq 0}$ is a nested composition, for any $s \leq t$, the blocks of $\mathcal{C}_t$ are obtained by merging blocks of $\mathcal{C}_s$, and if $A \leq B$ are two blocks of $\mathcal{C}_s$ that merge, they also merge with any block C such that $A \leq C \leq B$.

Naturally we say that $(\mathcal{C}_t)_{t\geq 0}$ is an exchangeable nested composition of $\mathbb{N}$ if for any finite permutation σ we have

$$(\mathcal{C}_t)_{t\geq 0} \overset{\text{(d)}}{=} \big(\sigma(\mathcal{C}_t)\big)_{t\geq 0}.$$

We can extend the ordered paintbox construction to nested compositions. Let $(I_t)_{t\geq 0}$ be a nested interval-partition, and $(V_i)_{i\geq 1}$ an independent i.i.d. uniform sequence. Let $\mathcal{C}_t$ be the composition obtained from the ordered paintbox based on I_t by $(V_i)_{i\geq 1}$. Then it is immediate that $(\mathcal{C}_t)_{t\geq 0}$ is an exchangeable nested composition. Notice that this is only true because we have used the same sequence $(V_i)_{i\geq 1}$ for all times t.

REMARK 2.5. Similar to Remark 1.5, the nested composition obtained from an ordered paintbox is not càdlàg in general. Again it admits a unique càdlàg modification and we shall always consider this modification.

We have the following direct reformulation of Theorem 2.1 in the framework of nested compositions.

THEOREM 2.6. *Let $(\mathcal{C}_t)_{t\geq 0}$ be an exchangeable nested composition of $\mathbb{N}$. We can find on the same probability space a nested interval-partition $(I_t)_{t\geq 0}$ and an independent i.i.d. sequence $(V_i)_{i\geq 1}$ of uniform variables such that a.s. the ordered paintbox based on $(I_t)_{t\geq 0}$ with $(V_i)_{i\geq 1}$ is $(\mathcal{C}_t)_{t\geq 0}$. This nested interval-partition is unique in distribution.*

PROOF. *Existence.* For any $t \geq 0$, $\mathcal{C}_t$ is an exchangeable composition of $\mathbb{N}$. We can apply Theorem 2.1 distinctly for $t \in \mathbb{Q}_+$ to find on the same probability space a collection of interval-partitions $(I_t)_{t\in\mathbb{Q}_+}$ such that for any $t \in \mathbb{Q}_+$ the ordered paintbox based on I_t is $\mathcal{C}_t$. Let I_t^n be the empirical interval-partition associated to $\mathcal{C}_t \cap [n]$. The fact that $(\mathcal{C}_t)_{t\geq 0}$ is a nested composition ensures that $(I_t^n)_{t\in\mathbb{Q}_+}$ is a nested interval-partition. Taking the limit as n goes to infinity shows that $(I_t)_{t\in\mathbb{Q}_+}$ is also a nested interval-partition. It admits a unique càdlàg extension given by

$$I_s = \operatorname{int}\Big(\bigcap_{\substack{t\geq s\\ t\in\mathbb{Q}_+}} I_t\Big).$$

Let $(V_i)_{i\geq 1}$ be the i.i.d. uniform sequence given by Theorem 2.1 applied at time $t = 0$. To see that $(V_i)_{i\geq 1}$ is independent of $(I_t)_{t\geq 0}$, one can do the exact same steps as in the proof of Theorem 2.1 but using a vectorial version of de Finetti's theorem (see Appendix B).

We now show that for any $t \in \mathbb{Q}_+$, a.s.

$$i \sim_t j \quad \iff \quad V_i \text{ and } V_j \text{ are in the same interval of } I_t, \tag{3}$$

where $\sim_t$ is the relation given by the blocks of $\mathcal{C}_t$.

Let $n \geq 1$ and divide the interval $[0, 1]$ in n intervals of size $1/n$. We label the intervals from 1 to n in the same order as the variables $V_1, \dots, V_n$. Let $t \in \mathbb{Q}_+$, the first step is to notice that the empirical interval-partition I_t^n can be recovered by merging the blocks of size $1/n$ whose labels belong to the same block of $\mathcal{C}_t$. Now, let $V_i^{(n)}$ (resp. $V_j^{(n)}$) be the right-hand extremity of the interval with label i (resp. j). Using twice the law of large numbers shows that $V_i^{(n)}$ and $V_j^{(n)}$ converge to V_i and V_j respectively. Moreover, we know that I_t^n converges a.s. to I_t. If we suppose that $V_i < V_j$ and $i \sim_t j$, then for any $n \geq 1$, $(V_i^{(n)}, V_j^{(n)}) \subset I_t^n$, and taking the limit shows that $(V_i, V_j) \subset I_t$. Conversely if $(V_i, V_j) \subset I_t$, using the convergence, for n large enough we have $(V_i^{(n)}, V_j^{(n)}) \subset I_t^n$ and thus i and j are in the same block of $\mathcal{C}_t$.

That relation (3) holds a.s. for any $t \geq 0$ will follow by right-continuity. However, we have to be careful, in general the nested composition obtained from an ordered paintbox is not càdlàg. By continuity, the relation (3) only holds a.s. for all times t when $(\mathcal{C}_t)_{t \geq 0}$ is continuous. The original nested composition $(\mathcal{C}_t)_{t \geq 0}$ is recovered by considering a càdlàg modification of the nested composition obtained though an ordered paintbox based on $(I_t)_{t \geq 0}$.

Uniqueness. The uniqueness will come from the following convergence result:

$$\lim_{n \to \infty} \sup_{t \geq 0} d_H\big(I_t^n, I_t\big) = 0 \quad \text{a.s.}$$

We start by showing the convergence. Let $\varepsilon > 0$, we can split $[0, 1]$ into a finite number of pairwise disjoint intervals of length smaller than ε denoted by $J_1, \dots, J_p$. Given a combination of such intervals, $J = J_{i_1} \cup \dots \cup J_{i_k}$, let f_J^n denote the fraction of variables $V_1, \dots, V_n$ which belong to J. Then for any $\eta > 0$ using the law of large numbers we can a.s. find a large enough N_J such that

$$\forall n \geq N_J, \quad \big|\mathrm{Leb}(J) - f_J^n\big| < \eta.$$

Let N be large enough such that this condition is fulfilled for all possible combinations of intervals.

We now show that a.s.

$$\forall t \geq 0, \forall n \geq N, \quad d_H\big(I_t^n, I_t\big) \leq \eta + \varepsilon.$$

Let $x \notin I_t$, and $J_x = (\ell_x, r_x)$ be the interval such that $x \in J$ (in case x is the boundary of two intervals, we choose the left interval). First suppose that $\ell_x = 0$ or $r_x = 1$. By construction $0, 1 \notin I_t^n$, thus $d(x, 0) < \varepsilon$ or $d(x, 1) < \varepsilon$. In the other case, the variables $(V_i)_{i \geq 1}$ which are in $[0, \ell_x]$ and those in $[r_x, 1]$ are not in the same interval component of I_t, and by construction of the paintbox, their labels are not in the same block of $\mathcal{C}_t$. For $n \geq 1$, let f_1^n (resp. f_2^n) denote the frequency of the variables $(V_i)_{i \leq n}$ belonging to $[0, \ell_x]$ (resp. $[0, r_x]$). The previous remark shows that there is a point $y \in [f_1^n, f_2^n]$ which does not belong to I_t^n. For $n \geq N$ we know that $y \in [\ell_x - \eta, r_x + \eta]$ and thus $d(x, y) \leq \eta + \varepsilon$. This shows

$$\forall t \geq 0, \forall n \geq N, \quad \sup_{x \notin I_t} d\big(x, [0, 1] \setminus I_t^n\big) \leq \eta + \varepsilon.$$

Similarly consider $x_n \notin I_t^n$. If $x_n \in \{0, 1\}$, clearly $d(x_n, [0, 1] \setminus I_t) = 0$. In the other case the point x_n is the separation between two intervals of I_t^n. These two intervals can be seen as an agglomeration of blocks of size $1/n$ whose labels belong to the same block of I_t. Let i (resp. j) be the label of the right-most (resp. left-most) block of size $1/n$ of the left interval (resp. right interval) separated by x_n. The rules of the paintbox construction imply that V_i and V_j are not in the same interval of I_t, thus there exists $V_i \leq y_n \leq V_j$ such that $y_n \notin I_t$. The value of x_n is exactly the frequency of variables $V_1, \dots, V_n$ which belong to $[0, y_n]$. Let $J_{y_n} = (\ell_{y_n}, r_{y_n})$ be the interval to which y_n belongs, and f_1^n, f_2^n be as above the frequency of the n first variables in $[0, \ell_{y_n}]$ and $[0, r_{y_n}]$. As $\ell_{y_n} \le y_n$, we know that $f_1^n \le x_n$, and similarly $x_n \le f_2^n$. Thus for $n \ge N$, $x_n \in [\ell_{y_n} - \eta, r_{y_n} + \eta]$ and $d(x_n, y_n) \le \eta + \varepsilon$. This shows

$$\forall t \ge 0, \forall n \ge N, \quad \sup_{x \notin I_t^n} d(x, [0,1] \setminus I_t) \le \eta + \varepsilon.$$

Thus, a.s. $(I_t^n)_{t \ge 0}$ converges uniformly to $(I_t)_{t \ge 0}$.

To get uniqueness, it is sufficient to notice that the distribution of the sequence $((I_t^n)_{t \ge 0}; n \ge 1)$ is determined uniquely by that of $(\mathcal{C}_t)_{t \ge 0}$. As we can recover a.s. $(I_t)_{t \ge 0}$ from $((I_t^n)_{t \ge 0}; n \ge 1)$, the distribution of $(I_t)_{t \ge 0}$ is also determined by that of $(\mathcal{C}_t)_{t \ge 0}$. □

REMARK 2.7. This also proves Proposition 2.3 in a more detailed way.

2.3. *Uniform nested compositions*, *proof of Theorem* 1.8. We recall that $\mathfrak{I}$ stands for the quotient space of combs for the paintbox-equivalence relation. To be entirely rigorous we need to define a suitable σ-field on $\mathfrak{I}$. By definition of $\mathfrak{I}$ a paintbox based on any of the representatives of a class yields the same distribution on the space of coalescents. We can identify each class with this distribution and endow $\mathfrak{I}$ with the weak convergence topology of probability measures on the space of coalescents. We consider the associated Borel σ-field. This approach bears similarity with the Gromov-weak topology introduced in Greven, Pfaffelhuber and Winter (2009), more on this can be found in Section 4.

The first step to find a comb representation of a given exchangeable coalescent $(\Pi_t)_{t \ge 0}$ is to order the blocks of $(\Pi_t)_{t \ge 0}$ to obtain a nested composition. We will do that using the notion of uniform nested composition that we now introduce.

DEFINITION 2.8. Let $(\mathcal{C}_t)_{t \ge 0}$ be an exchangeable nested composition of $\mathbb{N}$ and $(\Pi_t)_{t \ge 0}$ be the associated coalescent. We say that $(\mathcal{C}_t)_{t \ge 0}$ is uniform if for any $n \ge 1$, conditionally on $(\Pi_t^n)_{t \ge 0}$, the order of the blocks of $(\mathcal{C}_t^n)_{t \ge 0}$ is uniform among all the possible orderings, that is, all the orderings such that $(\mathcal{C}_t^n)_{t \ge 0}$ is a nested composition.

The following lemma shows that any exchangeable coalescent can be turned into a uniform exchangeable nested composition.

LEMMA 2.9. *Let $(\Pi_t)_{t \ge 0}$ be an exchangeable coalescent. There exists a uniform exchangeable nested composition $(\mathcal{C}_t)_{t \ge 0}$ whose associated coalescent is $(\Pi_t)_{t \ge 0}$.*

PROOF. We proceed by induction. For $n = 1$ there is a unique trivial possible order on the blocks. Suppose that we have built for n an order on the blocks of $(\Pi_t^n)_{t \ge 0}$ such that only adjacent blocks can merge, we call such an order an order *consistent with the genealogy*. Then there are finitely many orders on the blocks of $(\Pi_t^{n+1})_{t \ge 0}$ that extend the previous order and are consistent with the genealogy. More precisely, if $n + 1$ is in a block of Π_0^{n+1} the extension is unique. If $n + 1$ is a singleton of Π_0^{n+1}, suppose that $\{n + 1\}$ coalesce at some point and that k blocks are involved in this coalescence event. Then there are k consistent extensions: $\{n + 1\}$ can be placed between any of the $k - 1$ other blocks, or at the left-most (resp. right-most) position. If $\{n + 1\}$ does not coalesce, the singleton can be placed at any position between blocks that do not coalesce. We pick one of these orders independently and uniformly.

By induction, we have built on the same probability space as $(\Pi_t)_{t \ge 0}$ a nested composition of $\mathbb{N}$ whose blocks merge according to $(\Pi_t)_{t \ge 0}$. It is easily checked from the construction that $(\mathcal{C}_t)_{t \ge 0}$ is a uniform nested composition. It remains to show that it is exchangeable. Fix $0 \le t_1 < \cdots < t_p$, and let $c_1, \ldots, c_p$ be compositions of $[n]$, whose block partitions are

$\pi_1, \ldots, \pi_n$ respectively. Fix some trajectory $\Pi^n := (\Pi^n_t)_{t\geq 0}$ of the coalescent. Let us denote by $O(\Pi^n)$ the number of orderings of the blocks of Π^n_0 yielding a nested composition, and let $O(c_1, \ldots, c_p; \Pi^n)$ be the number of such orderings verifying that $\mathcal{C}^n_{t_i} = c_i$, for $i \in \{1, \ldots, p\}$. Then for any permutation σ of $[n]$, the following direct calculation:

$$
\begin{aligned}
&\mathbb{P}(\mathcal{C}^n_{t_1} = c_1, \ldots, \mathcal{C}^n_{t_p} = c_p) \\
&\quad = \mathbb{E}\left[\frac{O(c_1, \ldots, c_p; \Pi^n)}{O(\Pi^n)} \mathbb{1}_{\{\Pi^n_{t_1} = \pi_1, \ldots, \Pi^n_{t_p} = \pi_p\}}\right] \\
&\quad = \mathbb{E}\left[\frac{O(c_1, \ldots, c_p; \sigma(\Pi^n))}{O(\sigma(\Pi^n))} \mathbb{1}_{\{\sigma(\Pi^n_{t_1}) = \pi_1, \ldots, \sigma(\Pi^n_{t_p}) = \pi_p\}}\right] \\
&\quad = \mathbb{E}\left[\frac{O(\sigma^{-1}(c_1), \ldots, \sigma^{-1}(c_p); \Pi^n)}{O(\Pi^n)} \mathbb{1}_{\{\Pi^n_{t_1} = \sigma^{-1}(\pi_1), \ldots, \Pi^n_{t_p} = \sigma^{-1}(\pi_p)\}}\right] \\
&\quad = \mathbb{P}(\mathcal{C}^n_{t_1} = \sigma^{-1}(c_1), \ldots, \mathcal{C}^n_{t_p} = \sigma^{-1}(c_p))
\end{aligned}
$$

proves that the nested composition is exchangeable. □

PROOF OF THEOREM 1.8. Let $(\Pi_t)_{t\geq 0}$ be an exchangeable coalescent. Let $(\mathcal{C}_t)_{t\geq 0}$ be the uniform nested compositions obtained through Lemma 2.9. Invoking Theorem 2.6 shows that there exists a comb representation $(I_t)_{t\geq 0}$ of $(\Pi_t)_{t\geq 0}$. The uniqueness is immediate from the definition of the quotient. □

3. Comb representation of Λ-coalescents. In this section, we restrict our attention to the well-studied case of Λ-coalescents. A process $(\Pi_t)_{t\geq 0}$ is a Λ-coalescent if for any $n \geq 1$, its restriction $(\Pi^n_t)_{t\geq 0}$ to $[n]$ is a Markov process such that starting from a partition with b blocks, any k blocks coalesce at rate

$$
\lambda_{b,k} = \int_{[0,1]} x^{k-2}(1-x)^{b-k} \Lambda(\mathrm{d}x)
$$

for a finite measure Λ on [0, 1].

The broad aim of this section is to find a Markovian comb representation of a given Λ-coalescent, and to provide its transitions. Recall from the last section the path followed to obtain a comb associated to an exchangeable coalescent. The first step is to order the blocks of the coalescent to get a nested composition, and then to use Theorem 2.6 to define a comb. Here we will follow this path in the special case of Λ-coalescents where we can have an explicit description of both the nested composition and the comb.

Let us first define the nested composition associated to a Λ-coalescent. Consider the modified transition rates

$$
\tilde{\lambda}_{b,k} = \frac{1}{b-k+1} \binom{b}{k} \lambda_{b,k}.
$$

Let $n \geq 1$, we define a Markov chain $(\mathcal{C}^n_t)_{t\geq 0}$ taking values in the space of composition of $[n]$ as follows. Starting from c, a composition of $[n]$ with b blocks, any k adjacent blocks merge at rate $\tilde{\lambda}_{b,k}$. These transition rates have a natural combinatorial interpretation. Consider $(\Pi^n_t)_{t\geq 0}$ the restriction to $[n]$ of a Λ-coalescent. Starting from a partition with b blocks, there are $\binom{b}{k}$ ways of merging k distinct blocks. Thus the total transition rate from b to $b-k+1$ blocks is $\binom{b}{k}\lambda_{b,k}$. Given that k blocks merge, the blocks that merge are chosen uniformly among the $\binom{b}{k}$ possible choices. Starting from a composition with b blocks, there are only $b-k+1$ ways to merge k adjacent blocks. Thus, the total transition rate of $(\mathcal{C}^n_t)_{t\geq 0}$ from b to $b-k+1$

blocks is the same as $(\Pi_t^n)_{t\geq 0}$, but instead of choosing uniformly k blocks among the $\binom{b}{k}$ possibilities, we choose k *adjacent* blocks among the $b-k+1$ possibilities.

We now extend this sequence of nested compositions to a nested composition of $\mathbb{N}$. To fully determine the distribution of $(\mathcal{C}_t^n)_{t\geq 0}$ we have to specify an initial distribution. We will always assume in this section that the process $(\mathcal{C}_t^n)_{t\geq 0}$ starts from the composition of $[n]$ composed of only singletons ordered uniformly. Using the Markov projection theorem (see, e.g., Kemeny and Snell ((1976), Section 6.3)), it is not hard to see that the sequence of processes $((\mathcal{C}_t^n)_{t\geq 0}; n\geq 1)$ is sampling consistent, that is, that the restriction of $(\mathcal{C}_t^{n+1})_{t\geq 0}$ to $[n]$ is distributed as $(\mathcal{C}_t^n)_{t\geq 0}$. Using the Kolmogorov extension theorem we can find $(\mathcal{C}_t)_{t\geq 0}$ an exchangeable nested composition of $\mathbb{N}$ whose projections to $[n]$ is distributed as $(\mathcal{C}_t^n)_{t\geq 0}$ for all $n\geq 1$. The process $(\mathcal{C}_t)_{t\geq 0}$ is a nested composition whose blocks merge according to a Λ-coalescent.

LEMMA 3.1. *Let $(\Pi_t)_{t\geq 0}$ be the coalescent associated to $(\mathcal{C}_t)_{t\geq 0}$. Then $(\Pi_t)_{t\geq 0}$ is a Λ-coalescent. Moreover for any $t\geq 0$, conditionally on Π_t^n, the composition $\mathcal{C}_t^n$ is obtained by ordering uniformly the blocks of Π_t^n.*

PROOF. Let $(\mathcal{C}_t^n)_{t\geq 0}$ and $(\Pi_t^n)_{t\geq 0}$ be the restriction to $[n]$ of $(\mathcal{C}_t)_{t\geq 0}$ and $(\Pi_t)_{t\geq 0}$ respectively. Let $\hat{Q}_n$ be the generator of $(\mathcal{C}_t^n)_{t\geq 0}$ and Q_n be the generator of a Λ-coalescent on $[n]$. The result will follow by using a Markov projection theorem from Rogers and Pitman (1981), see their Theorem 2. To apply this result, we need to find a probability kernel L_n from the space of partitions of $[n]$ to the space of compositions of $[n]$ such that for any function f from the space of compositions of $[n]$ to $\mathbb{R}$,

$$\forall\pi,\quad \hat{Q}_n L_n f(\pi) = L_n Q_n f(\pi)$$

and such that the initial distribution of $(\mathcal{C}_t^n)_{t\geq 0}$ is the push-forward by L_n of the initial distribution of $(\Pi_t^n)_{t\geq 0}$.

Let f be such a function. For π a partition of $[n]$, let $\mathcal{C}_\pi$ be the random composition of $[n]$ obtained by ordering the blocks of π uniformly. We set

$$\forall\pi,\quad L_n f(\pi) = \mathbb{E}[f(\mathcal{C}_\pi)].$$

Our choice of initial distribution for $(\mathcal{C}_t)_{t\geq 0}$ ensures that the second condition holds. A straightforward generator calculation shows that the above equality is fulfilled and that the desired result holds. See Appendix C for the details of the calculation. □

Using Theorem 2.6, the nested composition $(\mathcal{C}_t)_{t\geq 0}$ defines a unique nested interval-partition $(I_t)_{t\geq 0}$ that we call the Λ-*comb*. In the remainder of the section we want to show that the Λ-comb is a Markov process and give its transitions. We will express the transitions in terms of composition of bridges that we now introduce.

We say that a function $B\colon [0,1]\to[0,1]$ is a bridge if it is of the form

$$B(x) = x\left(1-\sum_{i\geq 1}\beta_i\right) + \sum_{i\geq 1}\beta_i \mathbb{1}_{\{x\leq V_i\}}$$

for a random mass-partition β and an independent i.i.d. sequence $(V_i)_{i\geq 1}$ of uniform $[0,1]$-valued variables. To any bridge we associate an interval-partition defined as

$$I(B) = \operatorname{int}\big([0,1]\setminus B([0,1])\big),$$

where $B([0,1])$ is the range of B. We can ask if the converse holds. The correct notion to answer this question is that of uniform order.

DEFINITION 3.2. Let I be a random interval-partition and $\mathcal{C}$ be the composition of $\mathbb{N}$ obtained through an ordered paintbox based on I. We say that I has a uniform order if for any $n \geq 1$, the order of the blocks of $\mathcal{C} \cap [n]$ is uniform.

The following lemma shows that having a uniform order is a necessary and sufficient condition for an interval-partition to be represented by a bridge. See Section 3.1 for a proof.

LEMMA 3.3. *Let I be a random interval-partition. There exists a bridge B such that $I(B) = I$ iff I has a uniform order. If I has a uniform order, the bridge B such that $I(B) = I$ is unique in distribution.*

Notice that for any $t \geq 0$, the Λ-comb I_t at time t has a uniform order. We will denote by B^{I_t} the bridge associated to I_t through Lemma 3.3. We are now in position to provide the transitions of the Λ-comb.

PROPOSITION 3.4. *Let $(I_t)_{t\geq 0}$ be the Λ-comb. The process $(I_t)_{t\geq 0}$ is Markovian, and for any $s, t \geq 0$, conditionally on I_t,*

$$I_{t+s} \overset{(d)}{=} I(B^{I_t} \circ B'_s), \tag{4}$$

where B'_s is an independent bridge distributed as B^{I_s}.

REMARK 3.5. In the coming down from infinity case we have a simpler description of the semigroup of the Λ-comb. Suppose that $(I_t)_{t\geq 0}$ starts from an interval-partition I_0 with b blocks and no dust. Then any k adjacent blocks of I_0 merge at rate $\tilde{\lambda}_{b,k}$.

The above proposition shows that the Λ-comb can be represented in terms of composition of independent bridges. As a direct corollary, we provide an alternative construction of the Λ-comb based on the flow of bridges of Bertoin and Le Gall (2003). A flow of bridges is a collection $(B_{s,t})_{s\leq t}$ of bridges which fulfills the following three conditions:

(i) For any $s < r < t$, $B_{s,t} = B_{s,r} \circ B_{r,t}$ (cocycle property).

(ii) For any $t_1 < \cdots < t_p$, the bridges $(B_{t_1,t_2}, \dots, B_{t_{p-1},t_p})$ are independent, and B_{t_1,t_2} is distributed as B_{0,t_2-t_1} (stationarity and independence of the increments).

(iii) The bridge $B_{0,t}$ converges to the identity map Id as $t \downarrow 0$ in probability in Skorohod topology.

It can be seen from the cocycle property that the interval-partition-valued process $(I(B_{0,t}))_{t\geq 0}$ is a nested interval-partition. Bertoin and Le Gall (2003) have defined a sampling procedure to obtain a coalescent from a flow of bridges. In our context, sampling from the flow of bridges according to this procedure is the same as doing a paintbox based on $(I(B_{0,t}))_{t\geq 0}$. An important result from Bertoin and Le Gall (2003) states that given a Λ-coalescent $(\Pi_t)_{t\geq 0}$, there exists a unique flow of bridges whose associated coalescent is distributed as $(\Pi_t)_{t\geq 0}$ (see Theorem 1 in Bertoin and Le Gall (2003)). We call it the Λ-flow of bridges. As a corollary of this correspondence and of Proposition 3.4, we are able to show that the comb associated to the Λ-flow of bridges is the Λ-comb introduced above from the transition rates.

COROLLARY 3.6. *Let Λ be a finite measure on* $[0, 1]$, *and let $(I_t)_{t\geq 0}$ be the Λ-comb and $(B_{s,t})_{s\leq t}$ be the Λ-flow of bridges. Then*

$$(I_t)_{t\geq 0} \overset{(d)}{=} \big(I(B_{0,t})\big)_{t\geq 0}.$$

PROOF. Let $p \geq 1$ and $0 \leq t_1 < \cdots < t_p$. Using the Markov property of $(I_t)_{t\geq 0}$ and the expression of the transitions (4) we know that

$$(I_{t_1}, \ldots, I_{t_p}) \stackrel{(d)}{=} (I_{t_1}, I(B^{I_{t_1}} \circ B'_1), \ldots, I(B^{I_{t_1}} \circ B'_1 \circ \cdots \circ B'_{p-1})),$$

where $(B'_1, \ldots, B'_{p-1})$ are independent bridges and for $1 \leq k \leq p-1$, B'_k is distributed as $B^{I_{t_{k+1}-t_k}}$.

Let $(B_{s,t})_{s\leq t}$ be the Λ-flow of bridges. Then from the cocycle property

$$(I(B_{0,t_1}), \ldots, I(B_{0,t_p})) = (I(B_{0,t_1}), \ldots, I(B_{0,t_1} \circ B_{t_1,t_2} \circ \cdots \circ B_{t_{p-1},t_p})).$$

Moreover as the flow of bridges has independent and stationary increments, $(B_{t_1,t_2}, \ldots, B_{t_{p-1},t_p})$ are independent bridges with the same distribution as above. $\square$

3.1. *Proof of Lemma* 3.3. We will need the following continuity result.

LEMMA 3.7. *The map $I \colon B \mapsto I(B)$ that maps a bridge to its associated interval-partition is continuous when the space of interval-partitions is endowed with the Hausdorff topology and the space of bridges with the Skorohod topology.*

PROOF. Let B^n be a sequence of bridges that converge to B in the Skorohod topology. We know that we can find a sequence of continuous bijections λ_n from $[0,1]$ to $[0,1]$ such that

$$\lim_{n\to\infty} \|\lambda_n - \mathrm{Id}\|_\infty = 0$$

and

$$\lim_{n\to\infty} \|B - B^n \circ \lambda_n\|_\infty = 0.$$

Let $I = I(B)$ and $I_n = I(B^n)$. As the interval-partitions are obtained from bridges, we can re-write the Hausdorff distance as

$$d_H(I, I_n) = \sup_{x\in[0,1]} \inf_{y\in[0,1]} |B^n(x) - B(y)| \vee \sup_{x\in[0,1]} \inf_{y\in[0,1]} |B^n(y) - B(x)|.$$

We have

$$\sup_{x\in[0,1]} \inf_{y\in[0,1]} |B(x) - B^n(y)| \leq \sup_{x\in[0,1]} |B(x) - B^n(\lambda_n(x))|$$

and

$$\sup_{x\in[0,1]} \inf_{y\in[0,1]} |B(y) - B^n(x)| \leq \sup_{x\in[0,1]} |B(\lambda_n^{-1}(x)) - B^n(x)|$$

and thus

$$\lim_{n\to\infty} d_H(I, I^n) = 0,$$

which ends the proof. $\square$

PROOF OF LEMMA 3.3. First suppose that I is of the form $I(B)$ for some bridge B. Consider B^{-1} the generalized inverse of B. Let $(V_i)_{i\geq 1}$ be i.i.d. uniform variables and $\mathcal{C}$ be the composition obtained through an ordered paintbox using these variables. By construction of the ordered paintbox and as B^{-1} is nondecreasing, the order of the blocks of $\mathcal{C}$ is given by the order of the variables $(B^{-1}(V_i))_{i\geq 1}$. Conditionally on the bridge these variables are i.i.d. and thus their order is uniform.

Now let I be an interval-partition with a uniform order and $\mathcal{C}$ be the composition obtained by an ordered paintbox. We will first consider the case where I has finitely many interval components and no dust. The fact that the order of the blocks of the composition $\mathcal{C}$ is uniform shows that the order of the interval components of I is uniform (each block of $\mathcal{C}$ corresponds to an interval of I). Let K be the number of blocks of I, and let $V_1^* < \cdots < V_K^*$ be the order statistics of independent uniform variables. Suppose that β_1 is the length of the left-most interval of I, β_2 that of the second left-most, etc. then

$$\forall u \in [0,1], \quad B(u) = \sum_{i=1}^{K} \beta_i \mathbb{1}_{\{V_i^* \le u\}}$$

is a bridge such that $I(B) = I$. Indeed, since the order of the intervals is uniform, there is a uniform permutation σ of $[K]$ independent of $V_1^*, \dots, V_K^*$, such that $(\beta_{\sigma(i)})$ is ranked in nonincreasing order. This shows that

$$B(u) = \sum_{i=1}^{K} \beta_{\sigma(i)} \mathbb{1}_{\{V_{\sigma(i)}^* \le u\}}$$

indeed defines a bridge. This also shows the uniqueness in distribution of B.

Let us turn to the general case. Let $n \ge 1$ and consider I^n the empirical interval-partition associated to $\mathcal{C} \cap [n]$. By assumption the interval-partition I^n has a uniform order, thus using the above argument we can find a unique bridge B^n such that $I(B^n) = I^n$. We know that I^n converges a.s. to I. Let β_n (resp. β) be the mass-partition associated to I^n (resp. I). As the function that maps an interval-partition to its mass-partition is continuous, we have that β_n converges a.s. to β (see, e.g., Bertoin ((2006), Proposition 2.2)). We can now make use of another continuity result, namely Lemma 1 from Bertoin and Le Gall (2003), to show that the sequence of bridges $(B^n)_{n\ge1}$ converges in distribution to a bridge B obtained from the mass-partition β. Using Lemma 3.7, we know that $I(B^n)$ converges in distribution to $I(B)$. By uniqueness of the limit, we get that

$$I \overset{(\mathrm{d})}{=} I(B),$$

and that B is unique. $\square$

3.2. *Proof of Proposition* 3.4. We will first prove Proposition 3.4 for empirical interval-partitions and then take the limit. We start by proving the following lemma, which is the direct reformulation of Proposition 3.4 for empirical interval-partitions.

LEMMA 3.8. *Let $\mathcal{C}_0^n$ be an exchangeable composition of $[n]$ with a uniform order on its blocks, and let $(\mathcal{C}_t^n)_{t\ge0}$ be the Markov process started from $\mathcal{C}_0^n$ with transitions $(\tilde{\lambda}_{b,k}; 2 \le k \le b < \infty)$. If $(I_t^n)_{t\ge0}$ denotes the empirical nested interval-partition associated to $(\mathcal{C}_t^n)_{t\ge0}$, then conditionally on $\mathcal{C}_0^n$,*

$$I_t^n \overset{(\mathrm{d})}{=} I\big(B_0^n \circ B_t\big),$$

where B_0^n and B_t are independent bridges such that $I(B_0^n) = I_0^n$ and $I(B_t) = I_t$, the Λ-comb at time t.

PROOF. Let us denote by $(A_1, \dots, A_K)$ the blocks of $\mathcal{C}_0^n$ in order of their least element. As $\mathcal{C}_0^n$ has a uniform order on its blocks, according to Lemma 3.3 we can find $(U_1, \dots, U_K)$ such that conditionally on K these are i.i.d. uniform variables on [0, 1] and

$$\forall r \in [0,1], \quad B_0^n(r) = \frac{1}{n} \sum_{i=1}^{K} \mathrm{Card}(A_i) \mathbb{1}_{\{U_i \le r\}}$$

defines a bridges satisfying $I(B_0^n) = I_0^n$. Let B_t be independent and such that $I(B_t) = I_t$. To each interval component of I_0^n corresponds a unique block A_i of $\mathcal{C}_0^n$, and thus a unique jump time U_i of B_0^n. We claim that $I(B_0^n \circ B_t)$ is obtained by merging the intervals of I_0^n whose jump times belong to the same interval component of I_t. To see this, notice that by definition $I(B_0^n \circ B_t)$ is the set of flats of $(B_0^n \circ B_t)^{-1} = B_t^{-1} \circ (B_0^n)^{-1}$. Thus x and y belong to the same flat of $(B_0^n \circ B_t)^{-1}$ iff $(B_0^n)^{-1}(x)$ and $(B_0^n)^{-1}(y)$ belong to the same flat of B_t^{-1}, that is to the same interval component of I_t. The claim is proved by further noting that $(B_0^n)^{-1}(x)$ is the jump time of the interval component of I_0^n to which x belongs.

The previous procedure can be rephrased in terms of an ordered paintbox. The interval-partition $I(B_0^n \circ B_t)$ is obtained by labeling uniformly the K blocks of I_0^n, sampling a composition $\mathcal{C}_t'$ of $[K]$ according to an ordered paintbox based on I_t and merging the intervals of I_0^n whose labels belong to the same block of $\mathcal{C}_t'$. As I_t is the Λ-comb at time t, the composition $\mathcal{C}_t'$ is distributed as $\mathcal{C}_t^K$, the nested composition at time t obtained by merging K initial singleton blocks ordered uniformly according to the rates $(\tilde{\lambda}_{b,k}; 2 \leq k \leq b < \infty)$. Thus $I(B_0^n \circ B_t)$ can be obtained by letting its intervals merge at rate $(\tilde{\lambda}_{b,k}; 2 \leq k \leq b < \infty)$, and is distributed as I_t^n. □

PROOF OF PROPOSITION 3.4. Let $(I_t)_{t\geq 0}$ be the Λ-comb, and $(V_i)_{i\geq 1}$ be an independent sequence of i.i.d. uniform variables on [0, 1]. Denote by $(\mathcal{C}_t^n)_{t\geq 0}$ the nested composition of $[n]$ obtained by an ordered paintbox based on $(I_t)_{t\geq 0}$ using the sampling variables $(V_i)_{i\geq 1}$, and let $(I_t^n)_{t\geq 0}$ be the corresponding empirical nested interval-partition. According to Lemma 3.1 the interval-partition I_t has a uniform order, and thus there exists a bridge B_t such that $I(B_t) = I_t$. Conditionally on B_t, the sequence

$$\forall i \geq 1, \quad \xi_i = B_t^{-1}(V_i)$$

is i.i.d. We denote by μ_t the (random) law of ξ_1 conditionally on B_t, and by μ_t^n its empirical distribution defined as

$$\mu_t^n = \frac{1}{n}\sum_{i=1}^{n} \delta_{\xi_i}.$$

Note that B_t is the distribution function of μ_t. If B_t^n denotes the distribution function of μ_t^n, then B_t^n is a bridge such that $I(B_t^n) = I_t^n$. It follows from Lemma 3.8 that

$$(I_t^n, I_{t+s}^n) \overset{(\mathrm{d})}{=} (I_t^n, I(B_t^n \circ B_s')), \tag{5}$$

where B_s' is an independent bridge distributed as B^{I_s}. The result will follow by taking the limit in (5).

According to the Glivenko–Cantelli theorem (see, for instance, Proposition 4.24 in Kallenberg (2002)), the sequence of bridges $(B_t^n)_{n\geq 1}$ converges almost surely to B_t in the uniform topology. Thus $B_t^n \circ B_s'$ converges a.s. in the uniform topology to $B_t \circ B_s'$, and by Lemma 3.7 and Proposition 2.3 the right-hand side of (5) converges a.s. to $(I_t, I(B_t \circ B_s'))$. According to Proposition 2.3, the left-hand side converges a.s. to (I_t, I_{t+s}) and we have proved that (4) holds.

It remains to show that $(I_t)_{t\geq 0}$ is Markovian. As $(\mathcal{C}_t)_{t\geq 0}$ is obtained from $(I_t)_{t\geq 0}$ through the ordered paintbox procedure, it is sufficient to prove that $(\mathcal{C}_t)_{t\geq 0}$ is Markovian. This follows from standard arguments from measure theory by noting that the filtration of $(\mathcal{C}_t)_{t\geq 0}$ is induced by that of its restrictions to $[n]$, and that all of these restrictions are Markov. □

3.3. *Dynamical combs.* As mentioned in the Introduction, an exchangeable coalescent models the genealogy of a population observed at a given time. By varying the observation time we obtain a dynamical genealogy that has been named the *evolving coalescent*. There has been much interest into studying evolving coalescents. For example, if the coalescent at a fixed time is the Kingman coalescent, the authors of Pfaffelhuber and Wakolbinger (2006), Pfaffelhuber, Wakolbinger and Weisshaupt (2011) have studied statistics of the evolving coalescent using a look-down representation, the authors of Greven, Pfaffelhuber and Winter (2013) studied the dynamics of the entire tree structure using the framework of the Gromov-weak topology. Evolving coalescents such that the coalescent at a fixed time is a more general Λ-coalescent have also been considered; see, for example, Kersting, Schweinsberg and Wakolbinger (2014) for the case of Beta-coalescents and Schweinsberg (2012) for the Bolthausen–Sznitman coalescent.

In this section we show that the previous results on the Markov property of the Λ-comb allow us to define a comb-valued process, the *evolving comb*, such that sampling from the evolving comb at a fixed time yields a Λ-coalescent. The evolving comb contains all the information about the dynamical genealogy but does not require the cumbersome framework of random metric spaces endowed with the Gromov-weak topology as in Greven, Pfaffelhuber and Winter (2013). For the sake of clarity we will only consider the evolving Kingman comb where we have an explicit construction of the genealogy at a fixed time.

We will build the evolving Kingman comb by defining its semigroup. Recall that when the coalescent associated to a nested interval-partition comes down from infinity, the comb can be represented using a comb function, see Section 1.2. Let f be a deterministic comb function and $s > 0$, we want to describe the genealogy of the population at time s given that its genealogy at time 0 is encoded by f. The procedure we follow is illustrated in Figure 4. Recall the Kingman comb construction discussed in the Introduction. Let $(e_i)_{i\geq 1}$ be a sequence of i.i.d. exponential variables, and $(U_i)_{i\geq 1}$ a sequence of i.i.d. uniform $[0, 1]$ variables. For $i \geq 1$, we set

$$T_i = \sum_{k\geq i+1} \frac{2}{k(k-1)} e_k.$$

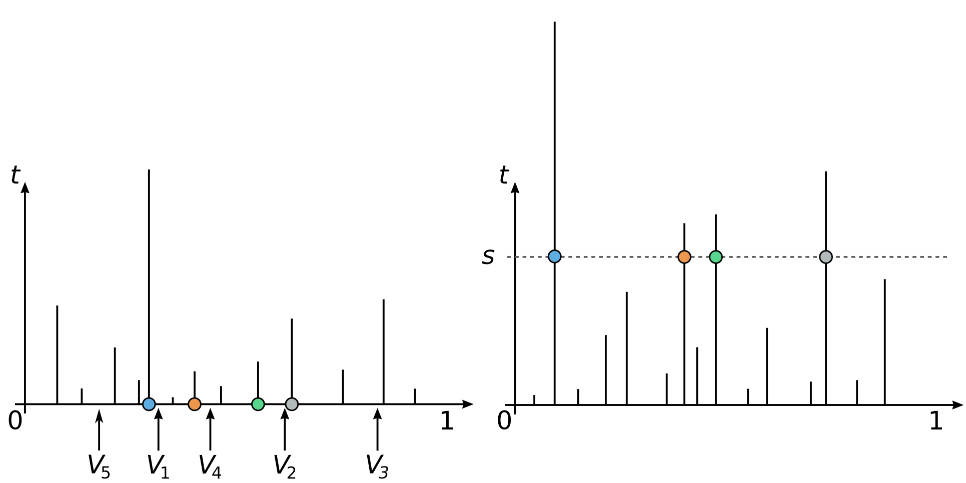

FIG. 4. *Transition of the evolving Kingman comb. The comb at time s, $\hat{f}_K$, is represented on the right, and the initial comb f is on the left. To obtain $\hat{f}_K$, one has first to erase the part of the right comb lying above level s. Here we have erased $N_s = 4$ teeth. Then throw $N_s + 1$ uniform variables $V_1, \ldots, V_{N_s+1}$, this defines N_s intervals between these variables, here (V_5, V_1), (V_1, V_4), (V_4, V_2) and (V_2, V_3). Finally take the largest tooth of f in each of these intervals, represented with a coloured root, and paste it in place of the erased tooth.*

The Kingman comb is given by

$$f_K = \sum_{i \geq 1} T_i \mathbb{1}_{U_i}.$$

It is known from Lambert and Schertzer ((2019), Proposition 3.1), that the above construction generates the comb associated to the flow of bridges, that is, the Λ-comb associated to the Kingman coalescent. There are only finitely many teeth of f_K that are larger than s, that is, such that $T_i \geq s$, say N_s. Let σ be their order, for example, $\sigma(1)$ is the label of the left-most tooth. Consider $V_1^* < \cdots < V_{N_s+1}^*$ the order statistics of $N_s + 1$ independent i.i.d. uniform variables. For $1 \leq k \leq N_s$ let M_k be the greatest tooth of f in the interval (V_k^*, V_{k+1}^*), that is,

$$M_k = \sup_{(V_k^*, V_{k+1}^*)} f.$$

We define new variables $(\hat{T}_i)_{i \geq 1}$ as follows:

$$\forall i > N_s, \quad \hat{T}_i = T_i,$$

and

$$\forall i \leq N_s, \quad \hat{T}_{\sigma(i)} = M_i + s.$$

We define

$$\hat{f}_K = \sum_{i \geq 1} \hat{T}_i \mathbb{1}_{U_i}.$$

Geometrically, the comb $\hat{f}_K$ is obtained through a cutting and pasting procedure illustrated in Figure 4.

The above construction defines an operator given by

$$P_t F(f) = \mathbb{E}\big[F(\hat{f}_K)\big],$$

for all continuous bounded functions F. We will show below that the family of operators $(P_t)_{t \geq 0}$ is a semigroup. Thus we can define a comb-valued Markov process $(\mathcal{I}^r)_{r \geq 0}$ whose transitions are given by the above construction. We call the process $(\mathcal{I}^r)_{r \geq 0}$ the *evolving Kingman comb*.

LEMMA 3.9. *The family of operators $(P_t)_{t \geq 0}$ is a semigroup. Moreover the Kingman comb is a stationary distribution of the evolving Kingman comb.*

PROOF. Let $s, t \geq 0$, let f be a deterministic comb. We call f_t the comb obtained through the above procedure at level t starting from f, and f_{t+s} the one obtained according to the above procedure at level s, but using f_t as starting comb. We need to show that f_{t+s} is distributed as f'_{t+s}, the comb obtained at level $t + s$ starting from f.

It is sufficient to show that the portion of the comb f_{t+s} lying between level 0 and $t + s$ is distributed as a Kingman comb truncated at height $t + s$. To show that, it is more convenient to see combs as nested interval-partitions. The procedure described above can be rephrased in terms of composition. Suppose that f_{t+s} has K truncated teeth at time s, this defines $K + 1$ intervals of $[0, 1]$. For each of these intervals of f_{t+s}, we throw a uniform variable. Two intervals merge at the first moment when their corresponding variables belong to the same subinterval of f_t. This is exactly the description of the ordered paintbox procedure. Thus, using the Markov property of the Kingman comb we know that f_{t+s}, between level 0 and $t + s$, is distributed as the truncation of a Kingman comb. This argument also shows that the Kingman comb is a stationary distribution. □

This construction can be easily extended to the case of Λ-coalescents that come down from infinity, even though we do not have an explicit construction of the comb in this case. In short, to obtain the evolving comb at time s, one needs to sample independently a new comb, erase the portion lying above height s and replace it by teeth sampled from the original comb. In the general case, we have to define the transition of the evolving comb using composition of bridges.

Again, the evolving comb can be built from the flow of bridges. Let $(B_{s,t})_{t\geq 0}$ be a Λ-flow of bridges, for any time r we can build a nested interval-partition by setting

$$(I_t^r)_{t\geq 0} = \big(I(B_{r,r+t})\big)_{t\geq 0}.$$

Then, using a similar argument as in the proof of Corollary 3.6 we could show that the comb-valued process $(\mathcal{I}^r)_{r\geq 0} = ((I_t^{-r})_{t\geq 0})_{r\geq 0}$ is distributed as the evolving comb introduced above. As a remark this provides a càdlàg modification of the evolving comb, and the Feller property of the flow of bridges ensures that the evolving comb is a Feller process.

4. Combs and ultrametric spaces. In this section we envision combs as random UMS. Random metric measure spaces have already been studied in Greven, Pfaffelhuber and Winter (2009), Gromov (1999). A key working hypothesis there is that the metric spaces are *separable*. In terms of combs and coalescents, separability translates into absence of dust (see Section 4.6). While separability is a very natural hypothesis when considering metric measure spaces, restricting our attention to combs without dust seems arbitrary, as dust has not raised any difficulty so far. In this section we provide a straightforward extension of the framework of random metric measure spaces to account for nonseparable UMS.

Let us recall the heuristic of our approach and give a short outline of this section. After a discussion on the assumptions of Definition 1.1 in Section 4.1, we define a topology on the space of UMS in Section 4.2 by saying that a sequence of UMS converges if the associated sequence of distance matrices converges weakly as probability measures. In the separable case, the Gromov reconstruction theorem (see Section $3.\frac{1}{2}.5$ of Gromov (1999)) ensures that spaces that are indistinguishable have the support of their measures in isometry. In general this result does not hold, we want to obtain a similar result for general UMS. In order to do that, we introduce in Section 4.3 the notion of a backbone of a UMS. A UMS can be seen as the leaves of a tree. This tree can be decomposed into (1) a separable part, that we call the backbone and (2) additional subtrees grafted on this backbone. Even though these subtrees can have a complex geometry, from a sampling standpoint they behave as star-trees (recall Figure 3). In Section 4.4, we show that if two UMS are indistinguishable in the Gromov-weak topology, then they are weakly isometric, in the sense that we can find an isometry between their backbones and a measure-preserving correspondence between the star-trees attached to them (see Proposition 4.12 for a rigorous statement). Finally Section 4.5 is dedicated to showing Corollary 1.18, that is, that we can always find a comb metric space weakly isometric to a given UMS with complete backbone, and Section 4.6 is devoted to showing Corollary 1.15 and Proposition 1.16 which are the analogous results in the complete and separable case.

4.1. *Discussion of Definition* 1.1. Recall Definition 1.1 of a UMS from the Introduction. This definition has two differences with the "naive" definition of a UMS (that is, any ultrametric space endowed with a probability measure on its Borel σ-field). First, we impose a measurability condition on the metric d. Second we allow the measure μ to be defined on a σ-field that is smaller than the usual Borel σ-field. In this section, we start with a discussion of the assumptions of Definition 1.1.

Let $\mathcal{P}_{\text{coal}}$ denote the state space of coalescents, endowed with its usual Borel σ-field (see Bertoin ((2006), Lemma 2.6)), and let Π be the map defined as

$$\Pi: \quad \begin{cases} U^{\mathbb{N}} \to \mathcal{P}_{\text{coal}}, \\ (x_i)_{i\geq 1} \mapsto (\Pi_t)_{t\geq 0}, \end{cases}$$

where

$$i \sim_{\Pi_t} j \quad \Longleftrightarrow \quad d(x_i, x_j) \leq t.$$

The following simple lemma proves that the measurability of d is the minimal requirement so that the coalescent obtained by sampling from U is a measurable process.

LEMMA 4.1. *The map Π is measurable when $U^{\mathbb{N}}$ is endowed with the product σ-algebra $\mathscr{U}^{\otimes\mathbb{N}}$ iff the distance d is $\mathscr{U} \otimes \mathscr{U}$ measurable.*

PROOF. Notice that by the definition of Π we have

$$\{d(x_1, x_2) \leq t\} = \{1 \sim_{\Pi_t} 2\}$$

which yields the "only if" part of the proof.

To prove the converse implication, let π be a partition of $[n]$ and define

$$R_{i,j} = \begin{cases} \{d(x_i, x_j) \leq t\} & \text{if } i \sim_\pi j, \\ \{d(x_i, x_j) > t\} & \text{if } i \nsim_\pi j. \end{cases}$$

Then

$$\{\Pi_t|_{[n]} = \pi\} = \bigcap_{i,j\leq n} R_{i,j},$$

which ends the proof. $\square$

We now turn to the second point of the definition. Roughly speaking, the Borel σ-field of a nonseparable ultrametric space tends to be large, and fewer measures can be defined on it. It is natural to ask whether all coalescents (especially coalescents with dust) can be represented as samples from ultrametric measure spaces, endowed with their natural Borel σ-field. (Recall that such ultrametric spaces are called Borel UMS.) It turns out that this question can be linked to a deep measure-theoretic problem known as the Banach–Ulam problem. It can be formulated as follows: can we find a space X and a probability measure μ defined on the power set of X such that $\mu(\{x\}) = 0$ for all $x \in X$? The next proposition connects our question to the Banach–Ulam problem. Note that point (iii) yields a positive answer to the problem.

PROPOSITION 4.2. *The following statements are equivalent.*

(i) *There exists an exchangeable coalescent with dust that can be obtained as a sample from a Borel UMS.*

(ii) *Any exchangeable coalescent can be obtained as a sample from a Borel UMS.*

(iii) *There exists an extension of the Lebesgue measure to all subsets of* $\mathbb{R}$.

This proposition is proved in Appendix F. Proposition 4.2 shows that answering our initial question, that is, representing coalescents with dust as samples from Borel UMS, amounts to finding an extension of the Lebesgue measure to all subsets of $\mathbb{R}$. A treatment of the latter problem requires advanced tools from set theory. Let us recall some basic facts about it. The interested reader is referred to Fremlin (1993) for a complete account on this question and on the Banach–Ulam problem.

A consequence of the various results stated in Fremlin (1993) is that point (iii) of the previous proposition has a greater *consistency strength* than the usual axioms Zermelo–Fraenkel–Choice (ZFC) of set theory. This means that, if ZFC is consistent, further assuming that there exists *no* extension of the Lebesgue measure does not lead to any contradiction. However, even under the assumption that ZFC is consistent, it *cannot* be shown that there is no contradiction in assuming the existence of an extension of the Lebesgue measure.

In other words, assuming that ZFC is consistent, we can safely assume that no extension of the Lebesgue measure exists, and thus that no coalescent with dust can be obtained by sampling from Borel UMS. On the contrary, even assuming that ZFC is consistent, we cannot be sure that further assuming that the Lebesgue measure can be extended (and thus that coalescents with dust are obtained as samples from Borel UMS) will not lead to a contradiction. However, according to the discussion in Remark 1E(e) in Fremlin (1993), it is extremely unlikely that such a contradiction exists.

REMARK 4.3. There is a short direct proof that, if the continuum hypothesis and the axiom of choice both hold, there can be no extension of the Lebesgue measure to all subsets of $\mathbb{R}$; see, for instance, the end of Section 3 of Chapter 2 of Billingsley (1995). As it is well known that the continuum hypothesis is relatively consistent with ZFC, this shows that the converse of point (iii) of Proposition 4.2 is also relatively consistent with ZFC.

The greater consistency strength of (iii) is a consequence of Corollary 2E of Fremlin (1993), which states that (iii) is equiconsistent with the existence of a measurable cardinal. Measurable cardinals are instances of (strongly) inaccessible cardinals, whose existence is well known to have greater consistency strength than ZFC alone, see, for instance Theorem 12.12 in Jech (2003).

Obviously, all these considerations go far beyond the scope of the current work. The approach we propose is to let the sampling measure be defined on a σ-field smaller than the usual Borel σ-field, namely $\mathscr{U}$. The previous discussion shows that this is not a necessary assumption to be able to represent all coalescents as samples from UMS, but that without it we would need to assume the existence of an extension of the Lebesgue measure to all subsets of $\mathbb{R}$. However, we hope that this short digression has led the reader to the conclusion that, as allowing the sampling measure to be defined on $\mathscr{U}$ avoids the aforementioned set-theoretic issues, it is a more natural framework for discussing coalescent theory on nonseparable UMS than having to assume that one of the statements of Proposition 4.2 holds.

Let us finally discuss the last point of Definition 1.1. This point can be reformulated in terms of the ball σ-field which is defined as follows.

DEFINITION 4.4. Let (U, d) be an ultrametric space. The ball σ-field denoted by $\mathscr{U}_\mathrm{b}$ is the σ-field induced by the open balls of (U, d), that is,

$$\mathscr{U}_\mathrm{b} = \sigma\big(\{B(x,t) : x \in U, t > 0\}\big),$$

where

$$\forall x \in U, \forall t > 0, \quad B(x,t) = \{y \in U : d(x,y) < t\}.$$

EXAMPLE 4.5. Consider any set U endowed with the metric

$$\forall x, y \in U, \quad d(x,y) = \mathbb{1}_{\{x \neq y\}}.$$

In this case $\mathscr{U}_\mathrm{b}$ is the countable-cocountable σ-field.

The last point of Definition 1.1 can now be rephrased as $\mathscr{U}_{\mathrm{b}} \subseteq \mathscr{U} \subseteq \mathscr{B}(U)$. It is important to notice that if $\mathscr{B}(U)$ denotes the Borel σ-field of (U,d), then $\mathscr{U}_{\mathrm{b}} \subseteq \mathscr{B}(U)$ always holds. In that sense, our definition of a UMS should be seen as a generalization of the naive definition as more measures can be defined on $\mathscr{U}_{\mathrm{b}}$ than on $\mathscr{B}(U)$. The converse statement, that is, that $\mathscr{B}(U) \subseteq \mathscr{U}_{\mathrm{b}}$, does not hold in general, as Example 4.5 shows. Nevertheless, in the important case where (U,d) is separable, we have that $\mathscr{U}_{\mathrm{b}} = \mathscr{B}(U)$, and the ultrametric d is $\mathscr{B}(U) \otimes \mathscr{B}(U)$-measurable. We thus recover the usual framework of metric measure spaces.

REMARK 4.6. The ball σ-field appears in other contexts where the underlying metric space is not separable, for example when considering the space of càdlàg functions with the uniform topology, as in Billingsley ((1999), Section 6 and Section 15).

4.2. *The Gromov-weak topology.* We now define the Gromov-weak topology on the space of UMS. Let $(U,d,\mathscr{U},\mu)$ be a UMS, and consider $(X_i)_{i\geq 1}$ an i.i.d. sequence distributed as μ. Recall that we define an exchangeable coalescent through the set of relations

$$i \sim_{\Pi_t} j \quad \Longleftrightarrow \quad d(X_i,X_j) \leq t.$$

Alternatively, we can see this coalescent as a random pseudo-ultrametric on $\mathbb{N}$ defined as

$$\forall i,j \geq 1, \quad d_\Pi(i,j) = d(X_i,X_j).$$

Both objects encode the same information, as d_Π can be recovered from $(\Pi_t)_{t\geq 0}$ through the equality

$$\forall i,j \geq 1, \quad d_\Pi(i,j) = \inf\{t \geq 0 : i \sim_{\Pi_t} j\}.$$

The distribution of this pseudo-ultrametric is called the *distance matrix distribution* of the UMS. Note that, as the correspondence between coalescents and distance matrices is bijective and bi-measurable, two UMS have the same distance matrix distribution iff their associated coalescents have the same distributions.

REMARK 4.7. From a topological point of view, a pseudo-ultrametric on $\mathbb{N}$ can be seen as an element of $\mathbb{R}_+^{\mathbb{N}\times\mathbb{N}}$ endowed with its product topology. Note that the correspondence between pseudo-ultrametric on $\mathbb{N}$ and coalescents outlined above is *not* a homeomorphism when the space of coalescents is endowed with the usual Skorohod topology. To see this, consider the (deterministic) coalescent on $\{1,2,3\}$ defined as

$$\forall t \geq 0, \quad \Pi_t^\varepsilon = \begin{cases} \{1\},\{2\},\{3\} & \text{if } t < 1, \\ \{1\},\{2,3\} & \text{if } 1 \leq t < 1+\varepsilon, \\ \{1,2,3\} & \text{if } t \geq 1+\varepsilon. \end{cases}$$

The corresponding ultrametric d_Π^ε on $\{1,2,3\}$ is given by

$$d_\Pi^\varepsilon(1,2) = 1+\varepsilon, \qquad d_\Pi^\varepsilon(1,3) = 1+\varepsilon, \qquad d_\Pi^\varepsilon(2,3) = 1.$$

It is clear that d_Π^ε converges to d_Π where

$$d_\Pi(1,2) = d_\Pi(1,3) = d_\Pi(2,3) = 1,$$

however, $(\Pi_t^\varepsilon)_{t\geq 0}$ does not converge to the coalescent $(\Pi_t)_{t\geq 0}$ corresponding to d_Π, given by

$$\forall t \geq 0, \quad \Pi_t^\varepsilon = \begin{cases} \{1\},\{2\},\{3\} & \text{if } t < 1, \\ \{1,2,3\} & \text{if } t \geq 1. \end{cases}$$

We use distance matrix distributions to define a topology on the space of UMS. Consider a sequence $(U_n, d_n, \mathscr{U}_n, \mu_n)_{n\geq 1}$ of UMS, and denote by $(\nu_n)_{n\geq 1}$ the associated sequence of distance matrix distributions. We say that the sequence $(U_n, d_n, \mathscr{U}, \mu_n)_{n\geq 1}$ converges in the Gromov-weak topology to $(U, d, \mathscr{U}, \mu)$ if $(\nu_n)_{n\geq 1}$ converges weakly to ν, the distance matrix distribution of (U, d, μ), in the space of probability measures on $\mathbb{R}_+^{\mathbb{N}\times\mathbb{N}}$, endowed with the product topology.

4.3. *Backbone*. It is well known that any ultrametric space (U, d) can be seen as the leaves of a tree. This is illustrated in Figure 3. Formally, we work on the space $U \times \mathbb{R}_+$ and consider the pseudo-metric

$$d_T\big((x,s),(y,t)\big) = \max\left(d(x,y) - \frac{s+t}{2}, \frac{|t-s|}{2}\right).$$

Let T be the space $U \times \mathbb{R}_+$ quotiented by the equivalence relation

$$z \sim z' \quad \Longleftrightarrow \quad d_T(z,z') = 0.$$

Then the space (T, d_T) is a real tree (see Evans ((2008), Definition 3.15)) whose leaves can be identified with (U, d).

DEFINITION 4.8 (Backbone of T). Define

$$f: \quad \begin{cases} U \to \mathbb{R}_+, \\ x \mapsto \inf\{t \geq 0 : \mu\big(B(x,t)\big) > 0\}, \end{cases}$$

(note that f is measurable since $\mathscr{U}_b \subseteq \mathscr{U}$) and let

$$\mathcal{S} := \{(x,t) \in T : t \geq f(x)\}.$$

The space $\mathcal{S}$ will be referred to as the backbone of the tree T, and we denote by $d_{\mathcal{S}}$ the distance d_T restricted to $\mathcal{S}$.

Let us now motivate the next result that will be fundamental to our approach. In words, Proposition 4.9 states that even if the underlying UMS is not separable, the backbone is always a separable tree. Second, one can recover the whole tree from the backbone by grafting some "simple" subtrees on the skeleton. By "simple," we mean that each of those subtrees has the sampling properties of a star-tree. Let us be more explicit about this last statement and discuss an example.

Consider the space $[0,1] \times \{0,1\}$ endowed with the ultrametric

$$\forall x,y \in [0,1], \forall a,b \in \{0,1\}, \quad d\big((x,a),(y,b)\big) = \begin{cases} 1 & \text{if } x \neq y, \\ 1/2 & \text{if } x = y \text{ and } a \neq b, \\ 0 & \text{if } (x,a) = (y,b). \end{cases}$$

The space $([0,1] \times \{0,1\}, d)$ is a star-tree where each branch splits in two at height $1/2$ (see Figure 5 left panel), we call it the bifurcating star-tree. We endow this space with the product measure of the Lebesgue measure on $[0,1]$ and the uniform measure on $\{0,1\}$, defined on the usual product Borel σ-field. Consider two independent random variables (X, A) and (Y, B) distributed according to the above measure. We see that these two variables lie at distance $1/2$ iff $X = Y$ and $A \neq B$, which happens with probability 0. Thus, from a sampling point of view, all points of the space lie at distance 1 from one another, that is, the bifurcating star-tree is a star-tree (see Figure 5 right panel).

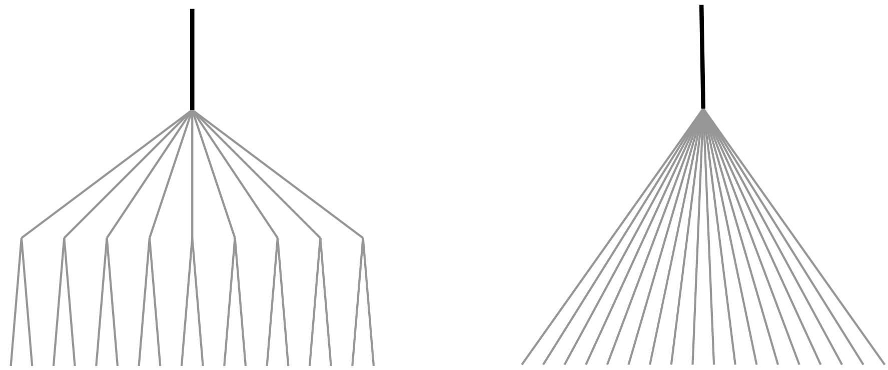

FIG. 5. *Left panel*: *The bifurcating star-tree*. *Right panel*: *The bifurcating star-tree simplified according to the metric* $\tilde{d}$. *In both cases*, *the backbone is illustrated with a bold black line and the subtrees attached to it with thin grey lines*.

This examples illustrates the more general phenomenon that from the measure point of view, the subtrees attached to the backbone behave like star-trees. More formally, consider a UMS $(U, d, \mathscr{U}, \mu)$. We introduce the distance

$$\forall x, y \in U, \quad \tilde{d}(x, y) = \mathbb{1}_{\{x \neq y\}} \inf\{t \geq 0 : d(x, y) \leq t \text{ and } \mu(B(x, t)) > 0\},$$

which replaces each subtree attached to the backbone by a star-tree. The point (iii) of the following proposition shows that the coalescent obtained by sampling from $(U, d, \mathscr{U}, \mu)$ is the same as the coalescent obtained by sampling from $(U, \tilde{d}, \mathscr{U}, \mu)$.

PROPOSITION 4.9.

(i) *The space* $(\mathcal{S}, d_{\mathcal{S}})$ *is a separable real tree*.

(ii) *The map*

$$\psi : \quad \begin{cases} (U, \mathscr{U}) \to (\mathcal{S}, \mathscr{B}(\mathcal{S})), \\ x \mapsto (x, f(x)) \end{cases}$$

is measurable and we define $\mu_{\mathcal{S}} := \psi \star \mu$, *the pushforward measure* (*on* $(\mathcal{S}, \mathscr{B}(\mathcal{S}))$) *of* μ *by* ψ. *In particular*, *the support of* $\mu_{\mathcal{S}}$ *belongs to the subset of the backbone* $\{(x, t) \in \mathcal{S} : t = f(x)\}$.

(iii) *Consider an i.i.d. sequence* $(X_i)_{i \geq 1}$ *distributed according to* μ. *Then for all* $i, j \geq 1$, $\tilde{d}(X_i, X_j) = d(X_i, X_j)$ *a.s.*

PROOF. We start by proving (i). The fact that $\mathcal{S}$ is a real tree can be checked directly from the definition. We now show that it is separable. Let $t \in \mathbb{Q}_+$, there are only countably many balls of (U, d) of radius t and positive mass, let us label them $(B_i^t)_{i \geq 1}$. For any $t \in \mathbb{Q}_+$ and $i \geq 1$, let $x_i^t \in B_i^t$. Let us now consider the collection $((x_i^t, t); t \in \mathbb{Q}_+, i \geq 1)$. First, since $\mu(B(x_i^t, t)) > 0$, it follows from the definition that $t \geq f(x_i^t)$, and thus $((x_i^t, t); t \in \mathbb{Q}_+, i \geq 1)$ is a countable collection of $\mathcal{S}$ and it remains to show that this collection is dense in $\mathcal{S}$.

Let $\varepsilon > 0$ and let $(x, s) \in U \times \mathbb{R}_+$ be in $\mathcal{S}$. We can find $t \in \mathbb{Q}_+$ such that $t > s \geq f(x)$ and $t - s < \varepsilon$. By definition of f, $\mu(B(x, t)) > 0$, and we can find i such that $B(x, t) = B_i^t$. Then $d(x, x_i^t) < t$ and

$$d(x, x_i^t) - \frac{t + s}{2} < d(x, x_i^t) - t + \frac{\varepsilon}{2} < \frac{\varepsilon}{2}$$

and thus $d_T((x, s), (x_i^t, t)) < \varepsilon$. This shows that the collection is dense and that the space is separable.

We now turn to the proof of (ii). Let $(x,t) \in \mathcal{S}$, we denote by

$$C(x,t) = \{(y,s) \in \mathcal{S} : d_T((x,t),(y,t)) = 0\}$$

the *clade* generated by (x,t). In a genealogical interpretation, $C(x,t)$ is the progeny of (x,t) that is, the subtree that has (x,t) as its MRCA. Notice that this notion can be defined similarly on any rooted tree (here the root is an "infinite point" obtained by letting $t \to \infty$). It is clear that $\psi^{-1}(C(x,t)) = B(x,t)$. Our results is now immediate from the fact that the clades of a rooted separable tree induce the Borel σ-field of the tree. A proof of this fact is given in Appendix D.

We now prove (iii). It is sufficient to prove that a.s. $d(X,Y) = \tilde{d}(X,Y)$ for X and Y two independent variables distributed as μ. Notice that for any $x, y \in U$, $d(x,y) \le \tilde{d}(x,y)$. Thus the probability that $d(X,Y) \neq \tilde{d}(X,Y)$ can be written

$$\begin{aligned}\mathbb{P}(d(X,Y) \neq \tilde{d}(X,Y)) &= \iint \mathbb{1}_{\{d(x,y)<\tilde{d}(x,y)\}} \mu(\mathrm{d}x)\mu(\mathrm{d}y) \\ &= \int \mu(\mathrm{d}x) \int \mu(\mathrm{d}y) \mathbb{1}_{\{d(x,y)<f(x)\}} = 0,\end{aligned}$$

where the last equality can be seen by writing

$$\{x, y \in U : d(x,y) < f(x)\} = \bigcup_{\varepsilon>0} \{x, y \in U : d(x,y) < f(x) - \varepsilon\}$$

and noticing that each event of the union in the right-hand side has null mass. □

REMARK 4.10 (Backbone and marked metric measure space). An object similar to the backbone appears in Gufler (2018) using the framework of marked metric measure spaces introduced in Depperschmidt, Greven and Pfaffelhuber (2011). We can interpret the backbone as a marked metric measure space where the metric space is U endowed with the backbone metric

$$\bar{d}(x,y) = d_S((x, f(x)), (y, f(y)))$$

and the mark space is $\mathbb{R}_+$. According to this correspondence, backbones are examples of elements of the set $\hat{\mathbb{U}}$ defined in Gufler (2018). In Gufler (2018) the marked metric measure space corresponding to the backbone is either considered as given, or built as the completion of the ultrametric measure space on $\mathbb{N}$ corresponding to the distance matrix distribution. The novelty of the present work is that we start from a general UMS and simplify it to obtain the backbone. This approach requires to identify the measurability assumptions to be made on UMS to avoid the problems that are discussed in Section 4.1.

Moreover, the link between backbones and marked metric measure spaces enables us to use the work of Depperschmidt, Greven and Pfaffelhuber (2011). For instance, this provides a metric, the marked Gromov–Prohorov metric, that metrizes the Gromov-weak topology on UMS and ensures that the topology is separable.

4.4. *Isomorphism between backbones*.

The aim of this section is to introduce the notion of isomorphism between backbones and to prove our reformulation of the Gromov reconstruction theorem.

DEFINITION 4.11. Let $(U, d, \mathscr{U}, \mu)$ and $(U', d', \mathscr{U}', \mu')$ be two UMS with respective backbones $(\mathcal{S}, \mu_{\mathcal{S}})$ and $(\mathcal{S}', \mu'_{\mathcal{S}})$. We say that Φ is an isomorphism from $\mathcal{S}$ to $\mathcal{S}'$ if:

(i) The map Φ is a measure-preserving isometry from $\mathcal{S}$ to $\mathcal{S}'$.

(ii) For every $(x,t) \in \mathcal{S}$, there exists $x' \in U'$ such that $\Phi((x,t)) = (x',t)$, that is, Φ preserves the second coordinate.

We say that two UMS are in weak isometry when they have isomorphic backbones.

Recall Proposition 1.17 from the Introduction. We want to show the following reformulation of Proposition 1.17. In words, this states that having the same distance matrix distribution is equivalent to being weakly isometric.

PROPOSITION 4.12. *Let $(U,d,\mathscr{U},\mu)$ and $(U',d',\mathscr{U}',\mu')$ be two UMS with respective backbones $(\mathcal{S},\mu_{\mathcal{S}})$ and $(\mathcal{S}',\mu'_{\mathcal{S}})$. We suppose that the two backbones are complete metric spaces. Then the two spaces $(\mathcal{S},\mu_{\mathcal{S}})$ and $(\mathcal{S}',\mu'_{\mathcal{S}})$ are isomorphic iff the distance matrix distribution associated $(U,d,\mathscr{U},\mu)$ and $(U',d',\mathscr{U}',\mu')$ are identical.*

Let us compare this result to the original result from Gromov (1999). In the separable case, if two UMS share the same coalescent then the supports of their measures are in isometry. Thus two separable spaces that are indistinguishable in the Gromov-weak topology share the exact same metric structure. The situation is rather different in the general case. Even if two UMS share the same coalescent, they can have rather different metric structures, think of the bifurcating star-tree and the star-tree of Figure 5. What Proposition 4.12 states is that in this case there is only a correspondence between coarsenings of the UMS, that is, the backbones on which all the subtrees are replaced by star-trees. This result is not surprising as the distance matrix distribution only contains the information of a countable number of points, which is not enough to explore the fine metric structure of the UMS.

The "only if" part of Proposition 4.12 is a direct consequence of the following lemma, which shows that the distance matrix distribution of a UMS can be recovered from an i.i.d. sequence of points of the backbone.

LEMMA 4.13. *Let $(X_i)_{i\geq 1}$ be an i.i.d. sequence in U sampled according to μ. Then a.s.*

$$\forall i,j\geq 1,\quad d(X_i,X_j)=d_{\mathcal{S}}\big((X_i,f(X_i)),(X_j,f(X_j))\big)+\frac{f(X_i)+f(X_j)}{2} \tag{6}$$

and

$$\forall i\geq 1,\quad f(X_i)=\inf\big\{t\geq 0:\{j:d(X_j,X_i)\leq t\}\text{ is infinite}\big\}. \tag{7}$$

PROOF. We know from Proposition 4.9 that for any $i,j\geq 1$, $\tilde{d}(X_i,X_j)=d(X_i,X_j)$ almost surely. Suppose that $(X_i,f(X_i))$ and $(X_j,f(X_j))$ lie at distance 0 in the backbone, then $\tilde{d}(X_i,X_j)=f(X_i)=f(X_j)$ and (6) holds. Otherwise notice that $d(X_i,X_j)\geq f(X_i)$ and $d(X_i,X_j)\geq f(X_j)$. Thus

$$d(X_i,X_j)-\frac{f(X_i)+f(X_j)}{2}\geq\frac{|f(X_i)-f(X_j)|}{2}$$

and

$$d_{\mathcal{S}}\big((X_i,f(X_i)),(X_j,f(X_j))\big)=d(X_i,X_j)-\frac{f(X_i)+f(X_j)}{2}.$$

The second point of the lemma is a direct consequence of the definition of f and of the observation that if $\mu(B(x,t))>0$, then a.s. there are infinitely many $(X_i)_{i\geq 1}$ that belong to this ball. □

It remains to show the converse proposition, that is, that if two UMS are sampling equivalent then they are in weak isometry. The proof we give is an adaptation of Gromov reconstruction theorem from Section $3.\frac{1}{2}.6$ of Gromov (1999).

PROOF OF PROPOSITION 4.12. We say that a sequence $(x_i, t_i)_{i\geq 1}$ in $\mathcal{S}$ is equidistributed if for any $A \in \mathcal{S}$,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{(x_i,t_i)\in A\}} = \mu_{\mathcal{S}}(A).$$

A well-known fact is that the empirical measure of an i.i.d. sample converges weakly to the sampling measure. Thus, a.s. an i.i.d. sequence is equidistributed.

Consider the map

$$D: \quad \begin{cases} \mathcal{S}^{\mathbb{N}} \to \mathbb{R}^{\mathbb{N}\times\mathbb{N}}, \\ (x_i, t_i)_{i\geq 1} \mapsto \left(d_{\mathcal{S}}\big((x_i,t_i),(x_j,t_j)\big) + \dfrac{t_i+t_j}{2}\right)_{i,j\geq 1} \end{cases}$$

and let D' be the analogous map for U'. Then Lemma 4.13 shows that the pushforward measure $D \star \mu_{\mathcal{S}}^{\otimes\mathbb{N}}$ is the distance matrix distribution associated to U. Similarly $D' \star {\mu'_{\mathcal{S}}}^{\otimes\mathbb{N}}$ is the distance matrix distribution associated to U'. As we have supposed that the two distance matrix distributions coincide, we can find a sequence $(x_i)_{i\geq 1}$ in U and a corresponding sequence $(x'_i)_{i\geq 1}$ in U' that have the same distance matrix, that is, such that

$$D\big((x_i, f(x_i))_{i\geq 1}\big) = D'\big((x'_i, f(x'_i))_{i\geq 1}\big).$$

We can suppose that these sequences are equidistributed and fulfill equalities (6) and (7) as all these events have probability 1. Using (7) we have

$$\forall i \geq 1, \quad f(x_i) = f(x'_i)$$

and then using (6) we obtain

$$\forall i, j \geq 1, \quad d_{\mathcal{S}}\big((x_i, f(x_i)),(x_j, f(x_j))\big) = d'_{\mathcal{S}}\big((x'_i, f(x'_i)),(x'_j, f(x'_j))\big).$$

We now extend this correspondence to an isomorphism between the backbones. Let $i \geq 1$ and $t \geq f(x_i)$, we set

$$\Phi\big((x_i, t)\big) = (x'_i, t).$$

It is clear that Φ is an isomorphism from $\{(x_i,t) \in \mathcal{S} : t \geq f(x_i), i \geq 1\}$ to $\{(x'_i,t) \in \mathcal{S}' : t \geq f(x'_i), i \geq 1\}$. It is now sufficient to show that this set is dense to end the proof, by extending Φ to $\mathcal{S}$ by continuity. To see that, let $(x,t) \in \mathcal{S}$. As $t \geq f(x)$, we know that $\mu(\{y \in U : d(x,y) \leq t+\varepsilon\}) > 0$ for any $\varepsilon > 0$. Writing

$$\{y \in U : d(x,y) \leq t+\varepsilon\} = \{y \in U : d_{\mathcal{S}}\big((x,t+\varepsilon),(y,t+\varepsilon)\big) = 0\},$$

as $(x_i, f(x_i))_{i\geq 1}$ is equidistributed, we see that we can find $(x_i, f(x_i))$ such that $(x_i, t+\varepsilon) = (x, t+\varepsilon)$. Moreover, it is immediate that $t+\varepsilon \geq f(x_i)$, and we have

$$d_{\mathcal{S}}\big((x_i, t+\varepsilon),(x,t)\big) = d_{\mathcal{S}}\big((x, t+\varepsilon),(x,t)\big) = \varepsilon.$$

The fact that Φ is measure preserving holds because we have chosen equidistributed sequences. $\square$

REMARK 4.14. According to the correspondence between backbones and marked metric measure spaces outlined earlier, Proposition 4.12 is similar to the more general Theorem 1 in Depperschmidt, Greven and Pfaffelhuber (2011), which is itself an adaptation of the Gromov reconstruction theorem. However, as we only address the case of backbones, we can be more specific. A direct application of Theorem 1 in Depperschmidt, Greven and Pfaffelhuber (2011) would only provide an isometry between the supports of the backbones whereas here we obtain a global isometry.

REMARK 4.15. The results of this section show that the backbone of a UMS contains the same information as the coalescent associated to that UMS. Thus properties of the coalescent can be read off from properties of the backbone. In particular, we can make precise an informal conjecture formulated in the context of exchangeable hierarchies in Forman, Haulk and Pitman (2018), and addressed in Forman (2020), concerning a nice decomposition of the sampling measure μ. Indeed, the sampling measure on the backbone is naturally decomposed into its atoms, its diffuse part on the set $\{(x,t) \in \mathcal{S} : t = 0\}$ of leaves of $\mathcal{S}$ at height 0 and the remaining diffuse part. This decomposition induces three qualitatively different behaviors of the coalescent. In short, points sampled in the atomic part form singletons of the coalescent that all merge at the same time, an event called "broom-like explosion" in Forman, Haulk and Pitman (2018). Second, points sampled in $\{(x,t) \in \mathcal{S} : t = 0\}$ always belong to an infinite block of the coalescent for $t > 0$, they form the "iterative branching part." Finally points sampled in the remaining part of the backbone are singletons of the coalescent that continuously merge with existing blocks. This behavior is referred to as "erosion."

4.5. *Comb metric measure space, completion of the backbone.* An important assumption of Proposition 4.12 is that the backbones of the UMS we consider are complete metric spaces. We will show in this section that the UMS associated to a comb enjoys this property up to the addition of a countable number of points. Let us start with two examples of combs illustrating that the backbone of a comb metric measure space is not in general complete.

First, consider the comb associated to the diadic space. Let $0 < t < 1$ and let k be the only integer such that $t \in [2^{-(k+1)}, 2^{-k})$. We set

$$I_t^2 = \bigcup_{0 \le i \le 2^{k+1}-1} \left(i2^{-(k+1)}, (i+1)2^{-(k+1)}\right)$$

and for $t \geq 1$ we set

$$I_t^2 = (0,1).$$

The diadic comb is illustrated in Figure 6. Now consider the comb metric d_I^2 associated to this comb, and let $x = 2^{-k}$ for some $k \geq 1$. Consider a nondecreasing sequence $(x_n)_{n\geq 1}$ that converges to x. It is not hard to see that $(x_n)_{n\geq 1}$ is Cauchy for d_I^2 but does not admit a limit.

Let us discuss a second example which is not separable. Consider the following comb:

$$I'_t = \begin{cases} \varnothing & \text{if } t < 1/2, \\ I^2_{t-1/2} & \text{otherwise.} \end{cases}$$

This comb is illustrated in Figure 6. It is rather clear that the backbone associated to $(I'_t)_{t\geq 0}$ is isometric to the backbone obtained from $(I_t^2)_{t\geq 0}$ (notice that here the isometry is not an isomorphism, as the backbone associated to $(I'_t)_{t\geq 0}$ is "shifted above by 1/2" from that of $(I_t^2)_{t\geq 0}$). The backbone is not complete for the same reason as above. The following proposition shows that up to the addition of a countable number of points, we can assume that the backbone associated to a comb metric space is complete.

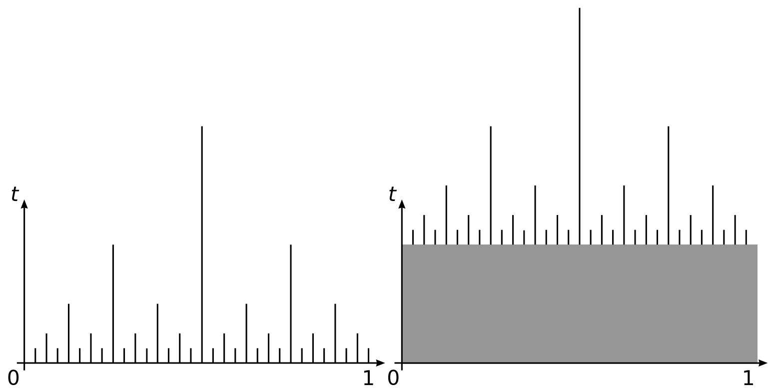

FIG. 6. *Left panel*: *The diadic comb. Right panel*: *The comb* $(I'_t)_{t\geq 0}$.

PROPOSITION 4.16. *Consider the comb metric* d_I *associated to a comb* $(I_t)_{t\geq 0}$. *We can find a countable set* F *and an extension* $\bar{d}_I$ *of* d_I *to* $[0,1]\cup F$ *such that* $\bar{d}_I$ *is ultrametric and the backbone associated to* $([0,1]\cup F, \bar{d}_I, \mathscr{I}, \mathrm{Leb})$ *is complete*.

REMARK 4.17. Here we have implicitly extended the Lebesgue measure to $[0,1]\cup F$ by giving zero mass to F.

A proof of this result is given in Appendix E. The proof of Corollary 1.18 now directly follows from the various results we have shown.

PROOF OF COROLLARY 1.18. Let $(U, d, \mathscr{U}, \mu)$ be a UMS with complete backbone, and let $(\Pi_t)_{t\geq 0}$ be the associated coalescent. Using Theorem 1.8 we can find a nested interval-partition whose associated coalescent is $(\Pi_t)_{t\geq 0}$. We can now use Proposition 4.16 to find a comb metric measure space whose backbone is complete which has the same distance matrix distribution as $(U, d, \mathscr{U}, \mu)$. Using Proposition 4.12 ends the proof. □

4.6. *The separable case*. In this section we consider the case of separable UMS and prove Corollary 1.15 and Proposition 1.16. The former result states that the weak isometry between backbones can be reinforced to an isometry between the supports of the measures in the case of separable complete UMS. The latter states that any complete separable ultrametric space is isometric to a properly completed comb metric space. Let us start with Corollary 1.15.

PROOF OF COROLLARY 1.15. Let $(I_t)_{t\geq 0}$ be a nested interval-partition without dust, and consider the corresponding comb metric measure space $([0,1], d_I, \mathscr{I}, \mathrm{Leb})$. The quotient space of $\{f_I = 0\}$ by the equivalence relation $x \sim y$ iff $d_I(x,y) = 0$ is a separable ultrametric space. Moreover, it is isometric to the subset $\{(x,t)\in \mathcal{S} : t = 0\}$ of the backbone $\mathcal{S}$ of $([0,1], d_I, \mathscr{I}, \mathrm{Leb})$. Thus the quotient space of $(\{f_I = 0\}, d_I)$ can be turned into a complete ultrametric space by adding a countable number of points as in Proposition 4.16, we denote this completion by (U_I, d_I) as in the Introduction. As $(I_t)_{t\geq 0}$ has no dust, we have $\mathrm{Leb}(\{f_I = 0\}) = 1$. Thus U_I can be endowed with the pushforward measure of the restriction of Leb to $\{f_I = 0\}$, defined on the Borel σ-field of (U_I, d_I). It is a probability measure, let us denote it by Leb. The space (U_I, d_I, Leb) is a separable complete Borel UMS that has the same distance matrix distribution as the original comb metric measure space $([0,1], d_I, \mathscr{I}, \mathrm{Leb})$.

Let $(U, d, \mathscr{U}, \mu)$ be a complete separable UMS. By restricting our attention to $\mathrm{supp}(\mu)$ we can assume without loss of generality that $\mathrm{supp}(\mu) = U$. According to Theorem 1.14 we can find a nested interval-partition $(I_t)_{t\geq 0}$ and a corresponding comb metric measure space $([0, 1], d_I, \mathscr{I}, \mathrm{Leb})$ whose distance matrix distribution is equal to that of $(U, d, \mathscr{U}, \mu)$. As $\mathrm{supp}(\mu) = U$, for each $t > 0$ we have $\mu(B(x, t)) > 0$. If $(\Pi_t)_{t\geq 0}$ denotes the coalescent obtained by sampling from $(U, d, \mathscr{U}, \mu)$, this shows that for each $t > 0$ all the blocks of Π_t have positive asymptotic frequency. Thus $(I_t)_{t\geq 0}$ has no dust, and we let (U_I, d_I, Leb) be the completion of the comb metric measure space as above. Then (U, d, μ) and (U_I, d_I, Leb) are two complete separable metric measure spaces (in the usual sense) whose distance matrix distributions are equal. Thus, the Gromov reconstruction theorem (see Section $3.\frac{1}{2}.6$ of Gromov (1999)) proves that we can find a measure-preserving isometry between (U_I, d_I, Leb) and (U, d, μ), which ends the proof. □

We now turn to the proof of Proposition 1.16. We will need the following lemma.

LEMMA 4.18. *Any separable ultrametric space* (U, d) *can be endowed with a measure* μ *on its Borel* σ*-field such that* $\mathrm{supp}(\mu) = U$.

PROOF. We build the measure by induction. For $n = 1$, as the space is separable there are only countably many balls of radius 1. If there are finitely many such balls, say k balls $B_1, \dots, B_k$, we define

$$\mu(B_i) = \frac{1}{k}.$$

Else we can find an enumeration of the balls, $(B_i)_{i\geq 1}$, and we define

$$\mu(B_i) = \left(\frac{1}{2}\right)^i.$$

Suppose that we have defined $\mu(B)$ for any ball of radius $1/n$. Given a ball B^n of radius $1/n$ there are at most countably many balls $(B_i^{n+1})_{i\geq 1}$ of radius $1/(n+1)$ such that $B_i^{n+1} \subset B^n$. Similarly if there are k balls we define

$$\mu(B_i^{n+1}) = \frac{\mu(B^n)}{k}$$

and if there are countably many balls we define

$$\mu(B_i^{n+1}) = \mu(B^n)\left(\frac{1}{2}\right)^i.$$

A simple application of Caratheodory's extension theorem now provides a probability measure μ defined on the Borel σ-field of (U, d) that extends this measure. It is straightforward from the construction that $\mathrm{supp}(\mu) = U$. □

REMARK 4.19. Note that a similar construction was mentioned in Lambert and Uribe Bravo (2017), where the resulting measure was referred to as the "visibility measure."

PROOF OF PROPOSITION 1.16. Let (U, d) be a separable complete UMS. Using Lemma 4.18 we can find a measure μ such that $\mathrm{supp}(\mu) = U$. An appeal to Corollary 1.15 now proves the result. □

APPENDIX A: EXCHANGEABLE HIERARCHIES

The aim of this section is to recall some results derived in Forman, Haulk and Pitman (2018) and discuss the link they have with the current results. Again, we recall that the present work should not be viewed as stemming from the work of Forman, Haulk and Pitman (2018), but should be viewed as an independent approach bearing similarities that we now expose.

Let X be an infinite space. A hierarchy on X is a collection $\mathcal{H}$ of subsets of X such that:

(i) for $x \in X$, $\{x\} \in \mathcal{H}$, $X \in \mathcal{H}$ and $\varnothing \in \mathcal{H}$;
(ii) given $A, B \in \mathcal{H}$, then $A \cap B$ is either A, B or $\varnothing$.

Any ultrametric space encodes a hierarchy that is obtained by "forgetting the time." More precisely, if (U, d) is an ultrametric space, then

$$\mathcal{H} = \{B(x,t), x \in X, t \geq 0\} \cup \{\{x\}, x \in X\} \cup \{X, \varnothing\}$$

is a hierarchy. The hierarchy $\mathcal{H}$ encodes the genealogical structure of (U, d), that is, the order of coalescence of the families, but not the coalescence times.

REMARK A.1. The converse does not hold, there exist hierarchies that cannot be obtained as the collection of balls of an ultrametric space. For example, consider a space X with cardinality greater than the continuum, endowed with a total order $\leq$, and define

$$\mathcal{H} = \{\{y : y \leq x\} : x \in X\} \cup \{\{x\}, x \in X\} \cup \{X, \varnothing\}.$$

The main object studied in Forman, Haulk and Pitman (2018) are exchangeable hierarchies on $\mathbb{N}$. Let σ be a permutation of $\mathbb{N}$, and $\mathcal{H}$ be a hierarchy on $\mathbb{N}$. Then σ naturally acts on $\mathcal{H}$ as

$$\sigma(\mathcal{H}) = \{\sigma(A), A \in \mathcal{H}\}.$$

A random hierarchy on $\mathbb{N}$ (see Forman, Haulk and Pitman (2018) for a definition of the σ-field associated to hierarchies) is called exchangeable if for any permutation σ,

$$\sigma(\mathcal{H}) \overset{(d)}{=} \mathcal{H}.$$

In a similar way that exchangeable coalescents are obtained by sampling in UMS, exchangeable hierarchies are obtained by sampling in hierarchies on measure spaces. Let (X, μ) be a probability space, and consider a hierarchy $\mathcal{H}$ on X. An exchangeable hierarchy $\mathcal{H}'$ can be generated out of an i.i.d. sequence $(X_i)_{i \geq 1}$ by defining

$$\mathcal{H}' = \{\{i \geq 1 : X_i \in A\}, A \in \mathcal{H}\}.$$

Again, an exchangeable hierarchy can be obtained from an exchangeable coalescent by forgetting the time. Let $(\Pi_t)_{t \geq 0}$ be an exchangeable coalescent. Then

$$\mathcal{H} = \{B, B \text{ is a block of } \Pi_t, t \geq 0\}$$

is an exchangeable hierarchy.

The main results in Forman, Haulk and Pitman (2018) show that any exchangeable hierarchy can be obtained by sampling from (1) a random "interval hierarchy" on $[0, 1)$ and (2) a random real-tree. The link with our results now seems straightforward.

An interval hierarchy on $[0, 1)$ is a hierarchy $\mathcal{H}$ on $[0, 1)$ such that all nonsingleton elements of $\mathcal{H}$ are intervals. Again, an interval hierarchy can be obtained from a nested interval-partition $(I_t)_{t \geq 0}$ by forgetting the time. The family of sets

$$\begin{aligned} \mathcal{H} = \{&I : I \text{ is an interval component of } I_t, t \geq 0\} \\ &\cup \{\{x\}, x \in [0, 1)\} \\ &\cup \{[0, 1), \varnothing\} \end{aligned}$$

is an interval hierarchy. Theorem 4 in Forman, Haulk and Pitman (2018) states that any exchangeable hierarchy on $\mathbb{N}$ can be obtained by sampling in a random interval hierarchy. This is the direct equivalent of our Theorem 1.8 that states that any exchangeable coalescent can be obtained by sampling in a random nested interval-partition.

Consider a measure rooted real-tree (T, d, ρ, μ), it can be endowed with a partial order $\preceq$ such that $y \preceq x$ if x is an ancestor of y (see Evans (2008)). Then, the fringe subtree of T rooted at $x \in T$ is defined as the set

$$F_T(x) = \{y \in T : y \preceq x\},$$

it is the set of the offspring of x. The natural hierarchy associated to (T, d, ρ) is

$$\mathcal{H} = \{F_T(x), x \in T\}.$$

Theorem 5 in Forman, Haulk and Pitman (2018) states that any exchangeable hierarchy can be obtained by sampling in the hierarchy associated to a random measure rooted real-tree. In our framework, we have seen that a nested interval-partition can be seen as an ultrametric space, and in Section 4.5 we have seen how this ultrametric space is embedded in a real-tree. Again we have proved here the reformulation of Theorem 5 from Forman, Haulk and Pitman (2018).

In a subsequent work, one of the authors has introduced the notion of mass-structural isomorphism (Forman (2020)). In a nutshell, two trees that are mass-structural isomorphic induce the same exchangeable hierarchy. In our framework, two spaces have the same coalescent iff their backbones are isomorphic. Thus, the mass-structural isomorphism is replaced here by the simpler notion of isomorphism.

Overall, the two works are very similar in the sense that they obtain the same kind of representation results for exchangeable hierarchies and exchangeable coalescents. However, the techniques used in the proofs are different, for example, the work of Forman, Haulk and Pitman (2018) relies on spinal decomposition whereas the present work relies on nested compositions. Moreover, as an ultrametric space contains "more information" than a hierarchy, our results are not trivially implied by the results in Forman, Haulk and Pitman (2018), but constitute an extension of their work.

Finally, we wish to stress two things. First, most of the difficulties that Section 4 deals with stem from the fact that we consider nonseparable metric spaces. These issues and the work that is done here heavily relies on the theory of metric spaces. Seeing genealogies as metric spaces is only possible if we keep the information on the times of coalescence, which is not the case when considering hierarchies.

Second, keeping this information allows us to study genealogies as time-indexed stochastic processes. It is a necessary step to study the Markov property of the combs associated to Λ-coalescents as in Section 3. This creates a direct link between the present work and the very rich literature on Λ-coalescents and coalescence theory that is not present in Forman, Haulk and Pitman (2018). Moreover, this provides a new approach to the question of dynamical genealogies, with the introduction of the dynamical comb.

APPENDIX B: INDEPENDENCE OF THE NESTED INTERVAL-PARTITIONS AND THE SAMPLING VARIABLES

Consider an exchangeable nested composition $(\mathcal{C}_t)_{t\geq 0}$, and let $(I_t)_{t\in\mathbb{Q}_+}$ be the nested interval-partition obtained by applying Theorem 2.1 distinctly for any $t \in \mathbb{Q}_+$, and $(V_i)_{i\geq 1}$ be the sequence of i.i.d. uniform variables obtained from Theorem 2.1 applied at time 0. The aim of this section is to show that $(V_i)_{i\geq 1}$ is independent from $(I_t)_{t\in\mathbb{Q}_+}$.

Let $0 = t_0 < t_1 < \cdots < t_p$. We can build a collection of sequences $(\xi_i^{(k)})_{i\geq 1, k=0,\dots,p}$ where for $k = 0, \dots, p$ and $i \geq 1$,

$$\xi_i^{(k)} = \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \mathbb{1}_{\{j \preceq_k i\}},$$

and $\preceq_k$ is the partial order on $\mathbb{N}$ representing $\mathcal{C}_{t_k}$ as in Section 2.1. The sequence of vectors $(\xi_i^{(0)}, \dots, \xi_i^{(p)})_{i\geq 1}$ is exchangeable. Thus by applying a vectorial version of de Finetti's theorem we know that there exists a measure μ on $[0,1]^{p+1}$ such that conditionally on μ the sequence of vectors is i.i.d. distributed as μ. We can now "spread" the variables $(\xi_i^{(0)})_{i\geq 1}$ using an independent i.i.d. uniform sequence as in the proof of Theorem 2.1 to obtain a sequence $(V_i)_{i\geq 1}$ that is i.i.d. uniform conditionally on μ. Thus the sequence $(V_i)_{i\geq 1}$ is independent of μ. The interval-partitions $(I_{t_0}, \dots, I_{t_p})$ can be recovered from the push-forward measures of μ by the coordinate maps on $\mathbb{R}^{p+1}$. Thus $(V_i)_{i\geq 1}$ is independent from $(I_t)_{t\in\mathbb{Q}_+}$.

APPENDIX C: GENERATOR CALCULATION

Let $n \geq 1$ and let $\hat{Q}_n$ denote the generator of the nested composition $(\mathcal{C}_t^n)_{t\geq 0}$ defined from the transition rates $(\tilde{\lambda}_{b,k}; 2 \leq k \leq b < \infty)$. Let Q_n be the generator of the restriction to $[n]$ of a Λ-coalescent. Here we show that for any function f, from the space of compositions of $[n]$ to $\mathbb{R}$,

$$\forall \pi, \quad \hat{Q}_n L_n f(\pi) = L_n Q_n f(\pi),$$

where L_n is the operator defined in the proof of Lemma 3.1.

We will need additional notation. The space of partitions and compositions of $[n]$ will be denoted by P_n and S_n respectively. For $\pi, \pi' \in P_n$, we denote by $q_{\pi,\pi'}$ the transition rate from π to π', that is, $q_{\pi,\pi'} = \lambda_{b,k}$ if π has b blocks and π' is obtained by merging k blocks of π, and $q_{\pi,\pi'} = 0$ otherwise. Similarly for $c, c' \in S_n$ we define $q_{c,c'}$ to be the transition rate from c to c'. Finally, we denote by $O(\pi)$ the set of compositions of $[n]$ whose blocks are given by the partition π, and $\mathrm{Card}(\pi)$ the number of blocks of π. Let $\pi \in P_n$ and denote by b the number of blocks of π, we have

$$\begin{aligned}
\hat{Q}_n L_n f(\pi) &= \sum_{\pi'\in P_n} q_{\pi,\pi'} \big(L_n f(\pi') - L_n f(\pi)\big) \\
&= \sum_{\pi'\in P_n} q_{\pi,\pi'} \bigg(\sum_{c'\in O(\pi')} \frac{1}{\mathrm{Card}(\pi')!} f(c') - \sum_{c\in O(\pi)} \frac{1}{\mathrm{Card}(\pi)!} f(c) \bigg) \\
&= \sum_{\pi'\in P_n} \sum_{c'\in O(\pi')} q_{\pi,\pi'} \frac{1}{\mathrm{Card}(\pi')!} f(c') - \sum_{c\in O(\pi)} \sum_{k=2}^{b} \frac{1}{\mathrm{Card}(\pi)!} \binom{b}{k} \lambda_{b,k} f(c).
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
L_n Q_n f(\pi) &= \sum_{c\in O(\pi)} \frac{1}{\mathrm{Card}(\pi)!} Q_n f(c) \\
&= \sum_{c\in O(\pi)} \frac{1}{\mathrm{Card}(\pi)!} \sum_{c'\in S_n} q_{c,c'} \big(f(c') - f(c)\big) \\
&= \sum_{c\in O(\pi)} \frac{1}{\mathrm{Card}(\pi)!} \sum_{c'\in S_n} q_{c,c'} f(c') - \sum_{c\in O(\pi)} \sum_{k=2}^{b} \frac{1}{\mathrm{Card}(\pi)!} \binom{b}{k} \lambda_{b,k} f(c).
\end{aligned}$$

We will end the calculation by showing that for any $c' \in S_n$, the coefficient in front of the term $f(c')$ in the left sum is the same for both expression. Let π' be the partition associated to c'. If π' is not obtained by merging k blocks of π for some k, then the coefficient of the term $f(c')$ in the sum is 0 in both expressions. Now suppose that π' is obtained by merging k blocks of π. In the first expression, we first choose the blocks of π that merge to get π' and then order the resulting partition to get the composition c'. There is only one possible way to do that and obtain a given c'. Thus the coefficient in front of $f(c')$ is $\lambda_{b,k}/(b-k+1)!$. In the second expression, we first choose an order to obtain a composition c, and then merge its blocks to get the composition c'. There are $k!$ possible orderings of π, and then exactly one merger of c that lead to c' (we can take any permutation of the k blocks that merge). Thus the coefficient in front of term $f(c')$ is

$$\frac{k!}{b!}\tilde{\lambda}_{b,k} = \frac{k!}{b!}\frac{1}{b-k+1}\frac{b!}{k!(b-k)!}\lambda_{b,k} = \frac{1}{(b-k+1)!}\lambda_{b,k}.$$

APPENDIX D: MEASURABILITY OF SEPARABLE ROOTED TREES

In this section we prove the claim made in the proof of Proposition 4.9 that the Borel σ-field of a separable rooted tree is induced by the clades of the tree. Let us be more specific.

We consider a separable real-tree (T, d) with a particular point $\rho \in T$ that we call the root. For $x, y \in T$, we denote by $[x, y]$ the unique geodesic with endpoints x and y (see Evans (2008)). Recall from Appendix A the fringe subtree of T rooted at x equivalently defined as the *clade*

$$C(x) = \{y \in T : x \in [\rho, y]\};$$

see Figure 7 for an illustration. The claim is that

$$\sigma(\{C(x), x \in T\}) = \mathscr{B}(T).$$

REMARK D.1. Our goal in the proof of Proposition 4.9 is to apply the result to the backbone whose root should be such that clades are the balls of U. This can be done by seeing the backbone as having a root "at infinity."

Let $x \in T$ and $\varepsilon > 0$, we assume that $\varepsilon < d(x, \rho)$. We denote by $B(x, \varepsilon)$ the open ball centered in x with radius ε, and $S(x, \varepsilon)$ the sphere of center x and radius ε, that is,

$$S(x, \varepsilon) = \{y \in T : d(x, y) = \varepsilon\}.$$

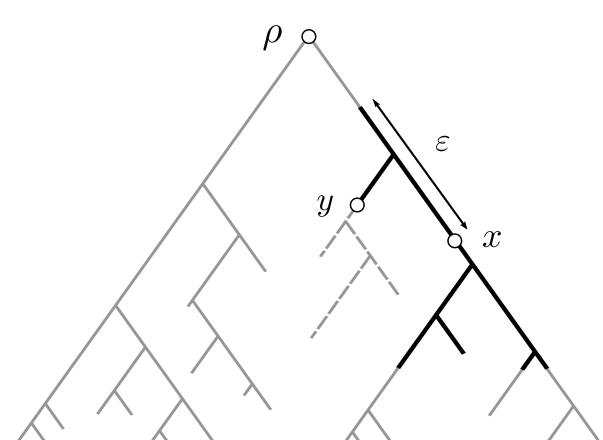

FIG. 7. *A tree rooted at ρ. The ball of radius ε and center x is represented by the black bold lines. An example of $y \in S(x, \varepsilon)$ is given, and its corresponding clade $C(y)$ is represented by grey dashed lines.*

There is a unique point in $a \in [\rho, x] \cap S(x, \varepsilon)$. It is clear that

$$B(x, \varepsilon) = C(a) \setminus \bigcup_{y \in S(x,\varepsilon)\setminus\{a\}} C(y).$$

Let $y \in S(x, \varepsilon)$, and $0 < \eta < \varepsilon$, we denote by y_η the only point in $[y, x]$ such that $d(y_\eta, y) = \eta$. We can write

$$\bigcup_{y \in S(x,\varepsilon)\setminus\{a\}} C(y) = \bigcap_{\eta > 0} \bigcup_{y \in S(x,\varepsilon)\setminus\{a\}} C(y_\eta).$$

The claim is proved if we can show that the union on the right-hand side is countable. This holds due to the separability of (T, d). To see that notice that by uniqueness of the geodesic, if y and y' are such that $y_\eta \neq y'_\eta$, then $d(y, y') > \eta$. Thus if the set $\{y_\eta : y \in S(x, \varepsilon) \setminus \{a\}\}$ is not countable, we can find an uncountable subset of $S(x, \varepsilon)$ such that any two points lie at distance at least η. This is not possible due to separability.

APPENDIX E: COMB COMPLETION

In this section we prove Proposition 4.16, that is, that the backbone of a comb is complete up to the addition of a countable number of points. We start from a nested interval-partition $(I_t)_{t \geq 0}$. We define

$$\mathcal{R} = \{x \in [0, 1] : \exists s_x, t_x \text{ s.t. } x \text{ is the right endpoint of an interval component of } I_u \text{ for } u \in [s_x, t_x]\}$$

and

$$\mathcal{L} = \{x \in [0, 1] : \exists s_x, t_x \text{ s.t. } x \text{ is the left endpoint of an interval component of } I_u \text{ for } u \in [s_x, t_x]\}.$$

We now work with a subset of $[0, 1] \times \{0, r, \ell\}$. Let

$$\bar{I} = ([0, 1] \times \{0\}) \cup (\mathcal{R} \times \{r\}) \cup (\mathcal{L} \times \{\ell\}).$$

We will simply write x for $(x, 0)$, x_r for (x, r) if $x \in \mathcal{R}$ and x_ℓ for (x, ℓ) if $x \in \mathcal{L}$. We extend d_I to $\bar{I}$ in the following way. Let $x < y$, we define

$$\begin{aligned}
&\bar{d}_I(x, y) = \bar{d}_I(x, y_\ell) = \bar{d}_I(x_r, y_\ell) = \bar{d}_I(x_r, y) = \sup_{[x,y]} f_I, \\
&\bar{d}_I(x, y_r) = \bar{d}_I(x_r, y_r) = \sup_{[x,y)} f_I, \\
&\bar{d}_I(x_\ell, y) = \bar{d}_I(x_\ell, y_\ell) = \sup_{(x,y]} f_I, \\
&\bar{d}_I(x_\ell, y_r) = \sup_{(x,y)} f_I
\end{aligned}$$

and $\bar{d}_I(x_r, x_\ell) = f(x)$. We use symmetrized definitions if $x > y$. It is straightforward to check that $\bar{d}_I$ is a pseudo-ultrametric. We will denote by $\mathcal{S}_I$ the backbone associated to this UMS, and $d_{\mathcal{S}_I}$ the restriction of the tree metric to $\mathcal{S}_I$, that is,

$$\forall (x', t), (y', s) \in \mathcal{S}_I, \quad d_{\mathcal{S}_I}((x', t), (y', s)) = \max\left\{\bar{d}_I(x', y') - \frac{t+s}{2}, \frac{|t-s|}{2}\right\}.$$

LEMMA E.1. *The backbone* $(\mathcal{S}_I, d_{\mathcal{S}_I}, \text{Leb})$ *associated to* $(\bar{I}, \bar{d}_I, \text{Leb})$ *is a complete metric space.*

PROOF. Consider $(x_n', t_n)_{n\geq 1}$ a Cauchy sequence in $\bar{I}$ for the metric $d_{\mathcal{S}_I}$. As

$$\frac{|t_n - t_m|}{2} \leq d_{\mathcal{S}_I}\big((x_n', t_n), (x_m', t_m)\big),$$

the sequence $(t_n)_{n\geq 1}$ is Cauchy and converges to a limit that we denote by t. Each point x_n' can be written as $x_n' = (x_n, a_n)$ with $x_n \in [0,1]$ and $a_n \in \{0, r, \ell\}$. The sequence $(x_n)_{n\geq 1}$ admits a subsequence that converges to a limit x for the usual topology in $[0,1]$. Without loss of generality we can assume that $(x_n)_{n\geq 1}$ is nondecreasing and converges to x.

Using the fact that the sequence is Cauchy, we know that

$$\lim_{n\to\infty} \sup_{m\geq n} \bar{d}_I(x_n', x_m') - \frac{t_n + t_m}{2} \leq 0,$$

which directly implies that

$$\lim_{\varepsilon\to 0} \sup_{[x-\varepsilon, x)} f_I \leq t.$$

Suppose that $x \in \mathcal{R}$. By definition of $\bar{d}_I$ and the above remark,

$$\lim_{n\to\infty} \bar{d}_I(x_n', x_r) - \frac{t_n + t}{2} \leq 0.$$

Thus the sequence $(x_n', t_n)_{n\geq 1}$ converges to (x_r, t).

Now suppose that $x \notin \mathcal{R}$. We claim that

$$\lim_{\varepsilon\to 0} \sup_{[x-\varepsilon, x)} f_I = f(x).$$

As $x \notin \mathcal{R}$ we directly know that

$$\lim_{\varepsilon\to 0} \sup_{[x-\varepsilon, x)} f_I \geq f(x).$$

Suppose that the above limit is strictly greater than $f(x)$. Then we can find a nondecreasing sequence $(y_n)_{n\geq 1}$ converging to x in the usual topology such that $f(y_n) \downarrow \lambda > f(x)$ as n goes to infinity. Let $\eta < \lambda - f(x)$. Notice that the set $\{y \in [0,1] : f(y) > \lambda - \eta\}$ is closed in the usual topology, as it is the complement of $I_{\lambda-\eta}$. This shows that x belongs to this set, which is a contradiction. Our claim is proved. Similar to above, it is now immediate that

$$\lim_{n\to\infty} \bar{d}_I(x_n', x) - \frac{t_n + t}{2} \leq 0$$

and that $(x_n', t_n)_{n\geq 1}$ converges to (x, t). □

REMARK E.2. This completion is already present in the compact case in Lambert and Uribe Bravo (2017). In this case, we have $\mathcal{R} = \mathcal{L} = \{f_I > 0\}$.

APPENDIX F: THE LINK BETWEEN DUST AND THE BANACH–ULAM PROBLEM

In this section we prove Proposition 4.2. We prove this result by constructing a solution to the so-called Banach–Ulam problem. This problem can be formulated as follows: is it possible to find a space X with a probability measure μ on the power-set $\mathcal{P}(X)$ of X such that $\mu(\{x\}) = 0$ for all $x \in X$?

Recall that a UMS $(U, d, \mathscr{U}, \mu)$ is called a Borel UMS if $\mathscr{U}$ is the Borel σ-field of (U, d). The support of the measure μ, $\mathrm{supp}(\mu)$, is defined as the intersection of all balls with positive mass. Equivalently, it can be defined as

$$\mathrm{supp}(U) = \big\{x \in U : \forall t > 0, \mu\big(B(x,t)\big) > 0\big\}.$$

We start with the following lemma, which gives a necessary and sufficient condition for the coalescent sampled from U to have dust in terms of the support of μ.

LEMMA F.1. *Let* $(U, d, \mathscr{U}, \mu)$ *be a UMS, and let* $(\Pi_t)_{t\geq 0}$ *be the associated coalescent. Then* $(\Pi_t)_{t\geq 0}$ *has dust iff* $\mu(\mathrm{supp}(\mu)) < 1$.

PROOF. Let $(X_i)_{i\geq 1}$ be an i.i.d. sequence in U distributed as μ and let $(\Pi_t)_{t\geq 0}$ be the coalescent obtained as above. We say that i is in the dust of the coalescent if there exists $t > 0$ such that $\{i\}$ is a singleton block of Π_t. We show that a.s.

$$i \text{ is in the dust} \quad \Longleftrightarrow \quad X_i \notin \mathrm{supp}(\mu).$$

Suppose that $X_i \in \mathrm{supp}(\mu)$. Then for any $t > 0$, $\mu(B(X_i, t)) > 0$, thus a.s. there are infinitely many other variables $(X_j)_{j\geq 1}$ in $B(X_i, t)$. Thus X_i is in an infinite block of Π_t. Conversely suppose that i is not in the dust, that is, that for any $t > 0$, $\{i\}$ is not a singleton block. Using Kingman's representation theorem for exchangeable partitions, we know that the block of i is a.s. infinite and has a positive asymptotic frequency f_i. The law of large numbers shows that $f_i = \mu(B(X_i, t)) > 0$. □

PROOF OF PROPOSITION 4.2. Let us start by showing that (i) implies (iii). Let $(U, d, \mathscr{U}, \mu)$ be a Borel UMS with associated coalescent $(\Pi_t)_{t\geq 0}$. Suppose that $(\Pi_t)_{t\geq 0}$ has dust. According to Lemma F.1, we know that $\mu(\mathrm{supp}(\mu)) < 1$. Consider $t > 0$ and let $(B^t_\alpha)_{\alpha\in A_t}$ be the collection of open balls of radius t with zero mass, where A_t is just an index set. We know that

$$\bigcup_{t>0} \bigcup_{\alpha\in A_t} B^t_\alpha = U \setminus \mathrm{supp}(\mu).$$

Using the continuity from below of the measure μ, we can find an $\varepsilon > 0$ such that $\mu(\bigcup_{\alpha\in A_\varepsilon} B^\varepsilon_\alpha) > 0$. We now consider the equivalence relation

$$x \sim y \quad \Longleftrightarrow \quad d(x, y) < \varepsilon$$

and denote by X the quotient space of $\bigcup_{\alpha\in A_\varepsilon} B^\varepsilon_\alpha$ for the relation $\sim$. We define the quotient map as

$$\varphi: \quad \begin{cases} U \to X, \\ x \mapsto \{y \in U : d(x, y) < \varepsilon\}. \end{cases}$$

We claim that φ is continuous when U is equipped with the metric topology induced by d, and X is equipped with the discrete topology $\mathcal{P}(X)$. Let $C \subset X$, then

$$\varphi^{-1}(C) = \bigcup_{x\in\varphi^{-1}(C)} B(x, \varepsilon)$$

which is an open subset of U. We call μ_X the push-forward measure of μ by the map φ. The measure $\mu_X/\mu_X(X)$ is a diffuse probability measure defined on $\mathcal{P}(X)$ as required. Thus, $(X, \mathcal{P}(X), \mu_X)$ is a solution to the Banach–Ulam problem.

Using the terminology from Fremlin (1993), this proves that the cardinality of X is a real-valued cardinal (see Notation 1C in Fremlin (1993)). According to Ulam's theorem (see Theorem 1D in Fremlin (1993)), real-valued cardinals fall into two classes: atomlessly-measurable cardinals and two-valued-measurable cardinals. The cardinal of X is atomlessly-measurable. To see this, one can, for example, notice that our measurability assumption on d implies that the cardinality of U (and thus that of X) is not larger than the continuum. (If this does not hold, then the diagonal does not belong to the product σ-field $\mathcal{P}(U) \otimes \mathcal{P}(U)$ and the metric d is not measurable.) Finally, using Theorem 1D of Fremlin (1993) proves (iii).

The fact that (ii) implies (i) is obvious, it remains to show that (iii) implies (ii). Suppose that there exists an extension of the Lebesgue measure to all subsets of $\mathbb{R}$, let us denote by $\overline{\text{Leb}}$ its restriction to $[0, 1]$. Let $(\Pi_t)_{t\geq 0}$ be any coalescent with dust. By Theorem 1.8 we can find a nested interval-partition $(I_t)_{t\geq 0}$ such that the paintbox based on $(I_t)_{t\geq 0}$ is distributed as $(\Pi_t)_{t\geq 0}$. Let d_I be the corresponding comb metric on $[0, 1]$. Then $([0, 1], d_I, \mathscr{B}_I([0, 1]), \overline{\text{Leb}})$ is a UMS, where $\mathscr{B}_I([0, 1])$ refers to the Borel σ-field induced by d_I and $\overline{\text{Leb}}$ is restricted to that σ-field. The coalescent obtained by sampling from this UMS is distributed as $(\Pi_t)_{t\geq 0}$. $\square$

Acknowledgements. The authors thank *Center for Interdisciplinary Research in Biology* (CIRB, Collège de France) for funding. They also greatly thank two anonymous reviewers for helpful comments.

REFERENCES

BERTOIN, J. (2006). *Random Fragmentation and Coagulation Processes*. *Cambridge Studies in Advanced Mathematics* **102**. Cambridge Univ. Press, Cambridge. MR2253162 https://doi.org/10.1017/CBO9780511617768

BERTOIN, J. and LE GALL, J.-F. (2003). Stochastic flows associated to coalescent processes. *Probab. Theory Related Fields* **126** 261–288. MR1990057 https://doi.org/10.1007/s00440-003-0264-4

BILLINGSLEY, P. (1995). *Probability and Measure*, 3rd ed. *Wiley Series in Probability and Mathematical Statistics*. Wiley, New York. MR1324786

BILLINGSLEY, P. (1999). *Convergence of Probability Measures*, 2nd ed. *Wiley Series in Probability and Statistics*: *Probability and Statistics*. Wiley, New York. MR1700749 https://doi.org/10.1002/9780470316962

DEPPERSCHMIDT, A., GREVEN, A. and PFAFFELHUBER, P. (2011). Marked metric measure spaces. *Electron. Commun. Probab.* **16** 174–188. MR2783338 https://doi.org/10.1214/ECP.v16-1615

DONNELLY, P. and JOYCE, P. (1991). Consistent ordered sampling distributions: Characterization and convergence. *Adv. in Appl. Probab.* **23** 229–258. MR1104078 https://doi.org/10.2307/1427746

EVANS, S. N. (2008). *Probability and Real Trees*. *Lecture Notes in Math.* **1920**. Springer, Berlin. MR2351587 https://doi.org/10.1007/978-3-540-74798-7

FORMAN, N. (2020). Exchangeable hierarchies and mass-structure of weighted real trees. *Electron. J. Probab.* **25** Paper no. 131. MR4169172 https://doi.org/10.1214/20-ejp522

FORMAN, N., HAULK, C. and PITMAN, J. (2018). A representation of exchangeable hierarchies by sampling from random real trees. *Probab. Theory Related Fields* **172** 1–29. MR3851828 https://doi.org/10.1007/s00440-017-0799-4

FREMLIN, D. H. (1993). Real-valued-measurable cardinals. In *Set Theory of the Reals* (*Ramat Gan*, 1991). *Israel Math. Conf. Proc.* **6** 151–304. Bar-Ilan Univ., Ramat Gan. MR1234282

GNEDIN, A. V. (1997). The representation of composition structures. *Ann. Probab.* **25** 1437–1450. MR1457625 https://doi.org/10.1214/aop/1024404519

GREVEN, A., PFAFFELHUBER, P. and WINTER, A. (2009). Convergence in distribution of random metric measure spaces (Λ-coalescent measure trees). *Probab. Theory Related Fields* **145** 285–322. MR2520129 https://doi.org/10.1007/s00440-008-0169-3

GREVEN, A., PFAFFELHUBER, P. and WINTER, A. (2013). Tree-valued resampling dynamics martingale problems and applications. *Probab. Theory Related Fields* **155** 789–838. MR3034793 https://doi.org/10.1007/s00440-012-0413-8

GROMOV, M. (1999). *Metric Structures for Riemannian and Non-Riemannian Spaces*. *Progress in Mathematics* **152**. Birkhäuser, Boston, MA. MR1699320

GUFLER, S. (2018). A representation for exchangeable coalescent trees and generalized tree-valued Fleming–Viot processes. *Electron. J. Probab.* **23** Paper no. 41. MR3806409 https://doi.org/10.1214/18-ejp153

JECH, T. (2003). *Set Theory*, 3rd millennium ed., revised and expanded. *Springer Monographs in Mathematics*. Springer, Berlin. MR1940513

KALLENBERG, O. (2002). *Foundations of Modern Probability*, 2nd ed. *Probability and Its Applications* (*New York*). Springer, New York. MR1876169 https://doi.org/10.1007/978-1-4757-4015-8

KEMENY, J. G. and SNELL, J. L. (1976). *Finite Markov Chains*. *Undergraduate Texts in Mathematics*. Springer, New York. MR0410929

KERSTING, G., SCHWEINSBERG, J. and WAKOLBINGER, A. (2014). The evolving beta coalescent. *Electron. J. Probab.* **19** Paper no. 64. MR3238784 https://doi.org/10.1214/EJP.v19-3332

KINGMAN, J. F. C. (1978). The representation of partition structures. *J. Lond. Math. Soc.* (2) **18** 374–380. MR0509954 https://doi.org/10.1112/jlms/s2-18.2.374

KINGMAN, J. F. C. (1982). The coalescent. *Stochastic Process. Appl.* **13** 235–248. MR0671034 https://doi.org/10.1016/0304-4149(82)90011-4

LAMBERT, A. (2017). Random ultrametric trees and applications. *ESAIM Proc. Surv.* **60** 70–89. https://doi.org/10.1051/proc/201760070

LAMBERT, A. and SCHERTZER, E. (2019). Recovering the Brownian coalescent point process from the Kingman coalescent by conditional sampling. *Bernoulli* **25** 148–173. MR3892315 https://doi.org/10.3150/17-bej971

LAMBERT, A. and URIBE BRAVO, G. (2017). The comb representation of compact ultrametric spaces. *P-Adic Numbers Ultrametric Anal. Appl.* **9** 22–38. MR3607645 https://doi.org/10.1134/S2070046617010034

PFAFFELHUBER, P. and WAKOLBINGER, A. (2006). The process of most recent common ancestors in an evolving coalescent. *Stochastic Process. Appl.* **116** 1836–1859. MR2307061 https://doi.org/10.1016/j.spa.2006.04.015

PFAFFELHUBER, P., WAKOLBINGER, A. and WEISSHAUPT, H. (2011). The tree length of an evolving coalescent. *Probab. Theory Related Fields* **151** 529–557. MR2851692 https://doi.org/10.1007/s00440-010-0307-6

PITMAN, J. (1999). Coalescents with multiple collisions. *Ann. Probab.* **27** 1870–1902. MR1742892 https://doi.org/10.1214/aop/1022677552

ROGERS, L. C. G. and PITMAN, W. (1981). Markov functions. *Ann. Probab.* **9** 573–582. MR0624684

SAGITOV, S. (1999). The general coalescent with asynchronous mergers of ancestral lines. *J. Appl. Probab.* **36** 1116–1125. MR1742154 https://doi.org/10.1017/s0021900200017903

SCHWEINSBERG, J. (2000). Coalescents with simultaneous multiple collisions. *Electron. J. Probab.* **5** Paper no. 12. MR1781024 https://doi.org/10.1214/EJP.v5-68

SCHWEINSBERG, J. (2012). Dynamics of the evolving Bolthausen–Sznitman coalecent. *Electron. J. Probab.* **17** Paper no. 91. MR2988406 https://doi.org/10.1214/EJP.v17-2378