



HAL
open science

Exchangeable coalescents, ultrametric spaces, nested interval-partitions: A unifying approach

Félix Foutel-Rodier, Amaury Lambert, Emmanuel Schertzer

► To cite this version:

Félix Foutel-Rodier, Amaury Lambert, Emmanuel Schertzer. Exchangeable coalescents, ultrametric spaces, nested interval-partitions: A unifying approach. 2018. hal-01841003v1

HAL Id: hal-01841003

<https://hal.science/hal-01841003v1>

Preprint submitted on 16 Jul 2018 (v1), last revised 12 Nov 2019 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Exchangeable coalescents, ultrametric spaces, nested interval-partitions: A unifying approach

Félix Foutel-Rodier^{1,2}, Amaury Lambert^{1,2}, and Emmanuel Schertzer^{1,2}

¹*Laboratoire de Probabilités, Statistiques et Modélisation (LPSM), Sorbonne Université, CNRS UMR 8001, Paris, France*

²*Center for Interdisciplinary Research in Biology (CIRB), Collège de France, CNRS UMR 7241, PSL Research University, Paris, France*

July 13, 2018

Abstract

Kingman's representation theorem [Kin78] 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 [Gne97].

Our first main result is that any exchangeable coalescent process (not necessarily Markovian) can be represented as a paintbox based on a random non-decreasing process valued in interval-partitions, called nested interval-partition, generalizing the notion of comb metric space introduced in [LUB17] 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 semi-group. This nested interval-partition directly relates to the flow of bridges of Bertoin and Le Gall [BLG03]. We also display a particularly simple description of the so-called evolving coalescent [PW06] by a comb-valued Markov process.

Next, we prove that any measured ultrametric 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 [GPW06], that was initially designed for separable metric spaces, to non-separable 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 complete backbone; 3) isometric to U if U is complete and separable.

KEYWORDS AND PHRASES: Combs; compositions; nested compositions; Lambda-coalescents; flow of bridges; metric measure spaces; Gromov-weak topology.

MSC2010 subject classification: Primary: 60G09; Secondary: 60C05; 60J35; 54E70.

1 Introduction

1.1 Ultrametric spaces and exchangeable coalescents

In this paper we extend earlier work from [LUB17] 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

$$\forall x, y, z \in U, d(x, y) \leq \max(d(x, z), d(z, y)).$$

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

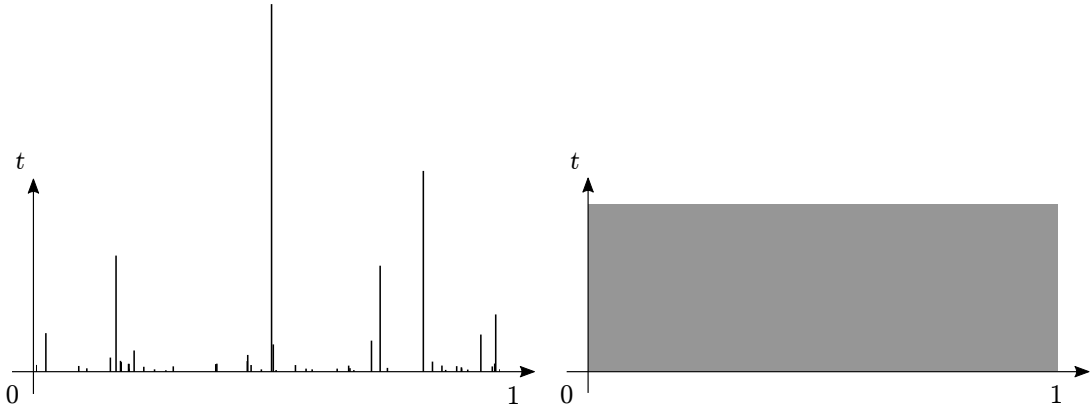


Figure 1: Representation of two nested-interval partitions. A point (x, t) is plotted in dark if $x \notin I_t$. Left panel: A realization of the Kingman comb, a tooth of size y at location x represents that $f(x) = y$. Right panel: The star-tree comb, an example of nested interval-partition that cannot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

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) on its Borel σ -field. A sample from the population 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

$$i \sim_t j \iff d(X_i, X_j) \leq t. \quad (1)$$

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

$$(\Pi_t)_{t \geq 0} \stackrel{(d)}{=} (\sigma(\Pi_t))_{t \geq 0}.$$

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 [LUB17] to represent *compact* ultrametric spaces. A comb is a function

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

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

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

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 [LUB17] shows that any compact ultrametric space is isometric to a properly completed and quotiented comb metric space (see Theorem 3.1 in [LUB17]).

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

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

For the sake of illustration, we recall the comb representation of the Kingman coalescent stated in [Kin82]. The Kingman comb is constructed out of an i.i.d. sequence $(e_i)_{i \geq 1}$ of exponential variables with parameter 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

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

See Figure 1 A. 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 [Ber06]).

More generally, the assumption that $\{f \geq \epsilon\}$ 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 [Lam17] 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 [LUB17] and Proposition 3.2 of [Lam17] to any ultrametric space (not only compact) and to any exchangeable 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 [LUB17]. This point will be discussed further in Section 2.1.

In order to deal with non-compact metric spaces, we need to generalize the definition of a comb by relaxing the condition on the finiteness of $\{f \geq \epsilon\}$. 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 noted $(I_i)_{i \geq 1}$. 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

$$d_H(I, \bar{I}) = \sup\{d(x, \bar{I}^c), x \notin I\} \vee \sup\{d(x, I^c), x \notin \bar{I}\}.$$

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

Definition 1.1. A *nested interval-partition* is a càdlàg function $(I_t)_{t \geq 0}$ taking values in the open subsets of $[0, 1]$ verifying

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

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 [LUB17]. Starting from a comb function f , we can build a nested interval-partition $(I_t)_{t \geq 0}$ as follows

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

and

$$I_0 = \text{int}(\{f = 0\})$$

where $\text{int}(A)$ 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 $f_I: [0, 1] \rightarrow \mathbb{R}_+$ as

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

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.

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

$$i \sim_{\Pi_t} j \iff X_i \text{ and } X_j \text{ belong to the same interval component of } I_t.$$

Remark 1.2. *The coalescent obtained through this sampling procedure is not càdlàg in general. As a coalescent is a non-decreasing process, we can always suppose that we work with a càdlàg modification of the coalescent.*

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 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 symmetric of a comb about 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 by this order but only with the genealogical structure of the comb. For this reason we introduce the following relation.

Definition 1.3. *Two generalized combs are paintbox-equivalent if their associated coalescents are identical in law. Being paintbox-equivalent is an equivalence relation, we note \mathfrak{J} the quotient space.*

Remark 1.4 (Combs as Ultrametric Measured Spaces (UMS)). *In Section 4, we show that any comb can be regarded as an UMS. More precisely, we show how to construct an UMS from any comb in a unique way. Further, any two paintbox-equivalent combs give rise to UMS's that are in weak isometry (see Proposition 4.11 for more details). Thus, in that sense, being paintbox-equivalent is equivalent to being weakly isometric, and the quotient space \mathfrak{J} can be obtained by quotienting using the weak isometry relation.*

Given $I \in \mathfrak{J}$ we note ρ_I the distribution on the space of coalescents of the paintbox based on any representative of \mathfrak{J} . We provide the following version of Kingman's representation theorem (e.g. see [Ber06] Theorem 2.1) for exchangeable coalescents.

Theorem 1.5. *Let $(\Pi_t)_{t \geq 0}$ be an exchangeable coalescent. There exists a unique distribution ν on \mathfrak{J} such that*

$$\mathbb{P}((\Pi_t)_{t \geq 0} \in \cdot) = \int_{\mathfrak{J}} \rho_I(\cdot) \nu(dI).$$

Remark 1.6. *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*

$$\beta_1 \geq \beta_2 \geq \dots \geq 0, \quad \sum_{i \geq 1} \beta_i \leq 1.$$

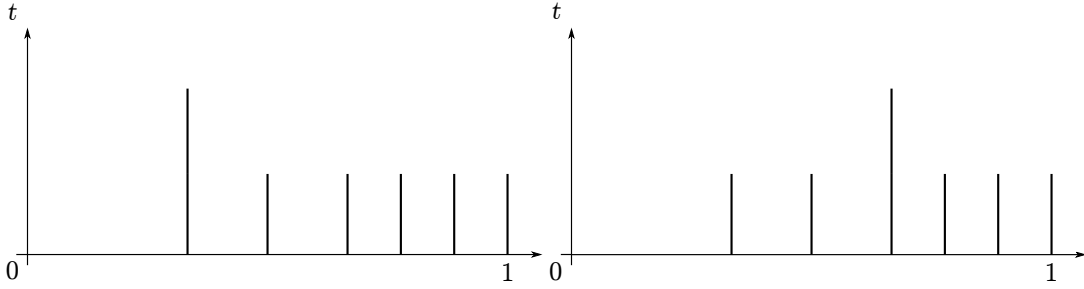


Figure 2: An example of two nested interval-partitions that have the same mass-coalescent but different coalescents. For both processes, the initial mass-partition is $(1/3, 1/6, 1/6, 1/9, 1/9, 1/9, 0, \dots)$, then $(1/3, 1/3, 1/3, 0, \dots)$ and finally $(1, 0, \dots)$. However, for the process on 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 $1/3$.

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, i.e. 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.

Remark 1.7. A result very similar to Theorem 1.5 has been obtained in [FHP17] 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 [FHP17]. 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 Section A of the appendix to explain the links between the present work and [FHP17].

Λ -coalescents. Most of the efforts made in the study of exchangeable coalescents have been devoted to the special case of Λ -coalescents [Pit99, Sag99]. These coalescents are parametrized by a finite measure Λ on $[0, 1]$, and their restriction to $[n] := \{1, \dots, n\}$ 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

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

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 semi-group.

Proposition 1.8. *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 1.9 (Combs and the flow of bridges). *The flow of bridges introduced by [BLG03] 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 1.8, see Section 3.*

A coalescent process models the genealogy of a population living at a fixed observation time. Many works have been concerned with the dynamical genealogy obtain by varying the observation time of the population. For example, in [PW06, PWW11] the authors study some statistics of the dynamical genealogy, namely the time to the MRCA and the total length of the genealogy. In [GPW08] the genealogy is encoded as a metric space (a real tree, see [Eva07]) 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 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 transitions can be explicitly described. In the particular case of coalescents that come down from infinity, the transitions of the evolving comb take a particularly simple form in terms of sampling from an independent comb.

Remark 1.10. *There exists a natural extension of the Λ -coalescents called the coalescents with simultaneous multiple collisions or Ξ -coalescents [Sch00]. All our results carry over to Ξ -coalescents, however for the sake of clarity we will focus on the case of Λ -coalescents.*

1.5 Comb representation of ultrametric spaces

The second main aim of this paper is to provide a comb representation of ultrametric measured spaces in the same vein as Theorem 3.1 of [LUB17]. We will only state our results informally and refer to Section 4 for the precise statements.

Any nested interval-partition $(I_t)_{t \geq 0}$ encodes an ultrametric on $[0, 1]$. Recall the definition of the comb function associated to $(I_t)_{t \geq 0}$,

$$\forall x \in [0, 1], f_I(x) = \inf\{t \geq 0 : x \in I_t\}.$$

Similarly to the compact case, setting

$$\forall x, y \in [0, 1], d_I(x, y) = \sup_{[x \wedge y, x \vee y]} f_I$$

defines a pseudo-ultrametric on $[0, 1]$. The ultrametric space $([0, 1], d_I)$ is the *comb metric space* associated to $(I_t)_{t \geq 0}$. For later purpose we will note U_I the completion of the set $\{f_I = 0\}$ for the metric d_I (this completion can be realized explicitly by adding countably many “left” and “right” faces to the comb, see Section 4.4).

We now 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 of [GPW06, Gro07] which is focused on separable metric spaces. In short, starting from an UMS a coalescent can be obtained by sampling from it as described in Section 1.1. We say that a sequence of UMS converges to a limiting UMS in the Gromov-weak sense if the corresponding coalescents converge weakly as partition-valued stochastic processes (see Section 4 for a more precise definition). We are now ready to state our representation result, which is a direct application of Theorem 1.5.

Theorem 1.11. *For any UMS (U, d, μ) there exists a comb metric space that is indistinguishable in the Gromov-weak topology from (U, d, μ) .*

The comb representation given by Theorem 1.11 is rather weak, since it only ensures that we can find a comb that has the same sampling structure than 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 non-separable 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 their supports in isometry, see e.g.

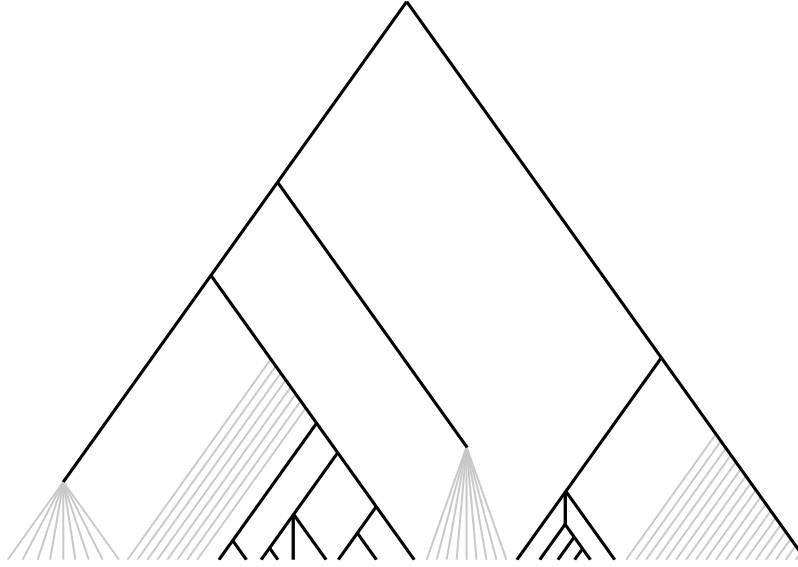


Figure 3: Illustration of the backbone decomposition. The dark thick lines represent the backbone. An element of the tree is represented in grey if its descendance has zero mass.

[Gro07] Section 3.1.5 or [GPW06] Proposition 2.6. Using the Gromov reconstruction theorem we will prove in Section 4.5 the following direct extension of Theorem 3.1 of [LUB17] to the separable case.

Corollary 1.12. *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) .*

Notice that the proof of the previous proposition is very different from the original proof of [LUB17] which is no longer valid for non-compact 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 non-separable 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.

An ultrametric space (X, d, μ) 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. Secondly, 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 have the sampling properties of a star-tree, in the sense that all sampled points are at the same distance. See Figure 3 for an illustration of this decomposition. An object very similar to the backbone is studied in [Guf18] 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 4.10. In the general case, we have the following version of Theorem 3.1 of [LUB17].

Corollary 1.13. *Let (U, d, μ) be an UMS with a complete backbone. There exists a comb metric space weakly isometric to (U, d, μ) .*

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 1.5 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 Λ -comb and study a family of nested compositions emerging from the Λ -coalescents. The proof of Proposition 1.8 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 1.5 on the comb representation of exchangeable coalescents. As was already mentioned in 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 [Gne97, DJ91] and then show how to extend them to nested compositions.

2.1 Exchangeable compositions

In combinatorics, a composition of $[n] := \{1, \dots, n\}$ (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, i.e. the blocks of π are denoted (B_1, B_2, \dots) and are such that for any $i, j \geq 1$,

$$i \leq j \iff \min(B_i) \leq \min(B_j).$$

Let σ be a finite permutation of \mathbb{N} , we note $\sigma(\mathcal{C})$ the composition whose blocks are $(\sigma(B_1), \sigma(B_2), \dots)$ and such that the order of the blocks is

$$\sigma(B_i) \leq \sigma(B_j) \iff B_i \leq B_j.$$

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

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

With our labeling convention, we have $B_1 = \{1, 4\}$, $B_2 = \{2, 3\}$ and $B_3 = \{5\}$ (B_1 needs not be the first block of \mathcal{C} for the order \leq). If $\sigma = (2, 1, 3, 5, 4)$, the composition $\sigma(\mathcal{C}^n)$ is given by

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

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

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

Gnedin [Gne97] provides a procedure to build an exchangeable composition of \mathbb{N} from any interval-partition I called the *ordered paintbox*. Let $(V_i)_{i \geq 1}$ 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

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

and the order of the block is

$$B \leq B' \iff V_i \leq V_j, \quad \forall i \in B, \forall j \in B'.$$

The main result of [Gne97] shows that any exchangeable composition of \mathbb{N} can be obtained as an ordered paintbox based on a random interval-partition. We now give a proof of this result different from the original one in [Gne97]. The result we prove is also slightly stronger because we provide a coupling between the composition and the interval-partition.

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 $(V_i)_{i \geq 1}$ of uniform $[0, 1]$ variables such that the ordered paintbox based on I by the sequence $(V_i)_{i \geq 1}$ is a.s. \mathcal{C} .*

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

$$i \preceq j \iff B_i \leq B_j$$

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

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

and the order \leq by

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

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 $(\xi_i)_{i \geq 1}$ such that*

$$i \preceq j \iff \xi_i \leq \xi_j.$$

Proof. Let A_i be the set of integers lower than i

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

It is immediate that the partition $(A_i \setminus \{i\}, \mathbb{N} \setminus \{i\} \setminus A_i)$ is an exchangeable partition of $\mathbb{N} \setminus \{i\}$. Thus Kingman's representation theorem [Ber06] ensures that the limit

$$\xi_i = \lim_{n \rightarrow \infty} \frac{1}{n} \text{Card}(A_i \cap [n])$$

exists a.s. For any n , the sequence $(\text{Card}(A_i \cap [n]))_{i \geq 1}$ is exchangeable. Taking the limit, the sequence $(\xi_i)_{i \geq 1}$ is also exchangeable.

We need to show that

$$i \preceq j \iff \xi_i \leq \xi_j.$$

The only difficulty here is to show that $\xi_i \leq \xi_j$ implies $i \preceq j$. Suppose that $i \not\preceq j$, we need to show that

$$\xi_i - \xi_j = \lim_{n \rightarrow \infty} \frac{1}{n} \text{Card}((A_i \setminus A_j) \cap [n]) > 0.$$

The partition $(A_j \setminus \{i, j\}, A_i \setminus \{i, j\} \setminus A_j, \mathbb{N} \setminus \{i, j\} \setminus A_i)$ 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. $A_i \setminus A_j$ 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 $A_i \setminus A_j$ is the reunion of all the blocks B such that $B_j < B \leq B_i$. Thus $A_i \setminus A_j$ 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 $(A_i \setminus A_j) \cap [n]$ is a singleton goes to 0 as n goes to infinity. \square

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:

$$F_\mu(V_1) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \mathbb{1}_{\{\xi_j \leq \xi_1\}} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \mathbb{1}_{\{j \leq 1\}} = \xi_1 \quad \mu\text{-a.s.}$$

In the terminology of [Gne97] 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 I_μ^c and coincides with the Lebesgue measure. The atomic part is supported by the right endpoints of the interval components of I_μ . If $J = (l, r)$ is an interval component of I_μ , the measure μ has an atom of mass $r - l$ located at r .

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

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

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. \square

We end this section with a technical result already present in [Gne97] which we will require. Let \mathcal{C} be an exchangeable composition of \mathbb{N} and \mathcal{C}^n its restriction to $[n]$. We denote n_i the size of the i -th block of \mathcal{C}^n . The *empirical interval-partition* associated to \mathcal{C}_n is given by

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

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*

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

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

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

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 [Gne97], the weak convergence of μ_n implies the convergence of I_{μ_n} to I in the Hausdorff topology. \square

Remark 2.4. *This also shows that the representation obtained through Theorem 2.1 is unique in distribution up to the points 0 and 1 that do not contribute to the paintbox. 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 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

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

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 .

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

Theorem 2.5. *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 $(\mathcal{C}_t)_{t \geq 0}$. Moreover, if we identify all the interval partitions that coincide on $(0, 1)$, the interval-partition representation 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

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

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.

$$i \sim_t j \iff V_i \text{ and } V_j \text{ are in the same interval of } I_t \quad (2)$$

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 (2) 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 (2) 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 points 0 and 1 play no role in the paintbox construction, thus in order to have a uniqueness result we need to choose a convention. Here we impose that $0 \notin I_t$ and $1 \notin I_t$ and show that in this case uniqueness holds. Notice that the interval-partition built in the proof of Theorem 2.1 and the empirical interval-partitions fulfill this property. The uniqueness will come from the following convergence result

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

We start by showing the convergence. Let $\epsilon > 0$, we can split $[0, 1]$ into a finite number of pairwise disjoint intervals of length smaller than ϵ noted 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

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

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

We now show that a.s.

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

Let $x \notin I_t$, and $J_x = (l_x, r_x)$ be the interval such that $x \in J$ (in case x is the boundary of two intervals, we chose the left interval). First suppose that $l_x = 0$ or $r_x = 1$. By construction $0, 1 \notin I_t^n$, thus $d(x, 0) < \epsilon$ or $d(x, 1) < \epsilon$. In the other case, the variables $(V_i)_{i \geq 1}$ which are in $[0, l_x]$ 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 $[0, l_x]$ (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 $y \in [l_x - \eta, r_x + \eta]$ and thus $d(x, y) \leq \eta + \epsilon$. This shows

$$\forall t \geq 0, \forall n \geq N, \sup_{x \notin I_t} d(x, (I_t^n)^c) \leq \eta + \epsilon.$$

Similarly consider $x_n \notin I_t^n$. If $x_n \in \{0, 1\}$, as we have imposed $0, 1 \notin I_t$, $d(x_n, I_t^c) = 0$. 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 $J_{y_n} = (l_{y_n}, r_{y_n})$ be the interval to which y belongs, and f_1^n, f_2^n be as above the frequency of the n first variables in $[0, l_{y_n}]$ and $[0, r_{y_n}]$. As $l_{y_n} \leq y_n$, we know that $f_1^n \leq x_n$, and similarly $x_n \leq f_2^n$. Thus for $n \geq N$, $x_n \in [l_{y_n} - \eta, r_{y_n} + \eta]$ and $d(x_n, y_n) \leq \eta + \epsilon$. This shows

$$\forall t \geq 0, \forall n \geq N, \sup_{x \notin I_t^n} d(x, I_t^c) \leq \eta + \epsilon.$$

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

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

Remark 2.6. *This also proves Proposition 2.3 in a more detailed way.*

2.3 Uniform nested compositions, proof of Theorem 1.5

We recall that \mathfrak{J} 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{J} . By definition of \mathfrak{J} 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{J} 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 [GPW06], more on this can be found in Section 4.

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

Definition 2.7. *Let $(C_t)_{t \geq 0}$ be an exchangeable nested composition of \mathbb{N} and $(\Pi_t)_{t \geq 0}$ be the associated coalescent. We say that $(C_t)_{t \geq 0}$ is uniform if for any $n \geq 1$, conditionally on $(\Pi_t^n)_{t \geq 0}$, the order of the blocks of $(C_t^n)_{t \geq 0}$ is uniform among all the possible orderings, i.e. all the orderings such that $(C_t^n)_{t \geq 0}$ is a nested composition.*

Let $(\Pi_t)_{t \geq 0}$ be an exchangeable coalescent. We now describe how to build a uniform nested composition from $(\Pi_t)_{t \geq 0}$. 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 $(\Pi_t^n)_{t \geq 0}$ 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 $(\Pi_t^{n+1})_{t \geq 0}$ 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 $(\Pi_t)_{t \geq 0}$ a nested composition of \mathbb{N} whose blocks merge according to $(\Pi_t)_{t \geq 0}$. It is easily checked from the construction that $(C_t)_{t \geq 0}$ is a uniform nested composition. A straightforward calculation shows that this nested composition is exchangeable.

Proof of Theorem 1.5. Let $(\Pi_t)_{t \geq 0}$ be an exchangeable coalescent. Let $(C_t)_{t \geq 0}$ be the uniform nested compositions obtained as above. Invoking Theorem 2.5 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. \square

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 $(\Pi_t^n)_{t \geq 0}$ to $[n]$ is a Markov process such that starting from a partition with b blocks, any k blocks coalesce at rate

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

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 semi-group. Recall from the last section the path followed to obtain a comb associated to an exchangeable coalescent. The first step is to order the block of the coalescent to get a nested composition, and then to use Theorem 2.5 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 $(\mathcal{C}_t^n)_{t \geq 0}$ 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 $(\Pi_t^n)_{t \geq 0}$ 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 $(\mathcal{C}_t^n)_{t \geq 0}$ 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.

It is usual for Λ -coalescents to formulate the semi-group of the process in terms of coagulation of partitions. For the composition $(\mathcal{C}_t^n)_{t \geq 0}$, we have the following identity. Let $t, s \geq 0$, then

$$\mathcal{C}_{t+s}^n \stackrel{(d)}{=} \text{Coag}(\mathcal{C}_t^n, \mathcal{C}'_s), \quad (3)$$

where \mathcal{C}'_s is an independent nested composition of $[n]$ distributed as \mathcal{C}_s^n . The precise definition of the coagulation operator Coag for compositions, which is a straightforward extension of the notion of coagulation of partitions, is given in Section 3.2.

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, it is not hard to see that the sequence of processes $(\mathcal{C}_t^n, t \geq 0)_{n \geq 1}$, with this initial distribution, is sampling consistent, i.e. 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, \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 Q_n be the generator of $(\mathcal{C}_t^n)_{t \geq 0}$ and \hat{Q}_n be the generator of a Λ -coalescent on $[n]$. The result will follow by using a Markov projection theorem from [RP81]. 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 partitions of $[n]$ to \mathbb{R} ,

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

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}_π be the random composition of $[n]$ obtained by ordering the blocks of π uniformly. We set

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

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. \square

Using Theorem 2.5, 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 remaining of the section we want to show that the Λ -comb is a Markov process and give its semi-group. For a finite $n \geq 1$ we have a description of the semi-group of $(\mathcal{C}_t^n)_{t \geq 0}$ given by the identity (3). Roughly speaking, we want to take the limit as $n \rightarrow \infty$ in this identity to obtain the semi-group of $(I_t)_{t \geq 0}$. In the framework of interval-partitions, the notion of coagulation is replaced by that of composition of bridges that we now introduce.

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

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

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

$$I(B) = \text{int} \left([0, 1] \setminus B([0, 1]) \right)$$

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.

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 $t \geq 0$, I_t , the Λ -comb at time t , has a uniform order. We will denote B^{I_t} the bridge associated to I_t through Lemma 3.3. We are now in position to provide the semi-group of the Λ -comb, which is the direct translation of identity (3) for interval-partitions and bridges.

Proposition 3.4. *Let $(I_t)_{t \geq 0}$ be the Λ -comb. The process $(I_t)_{t \geq 0}$ is Markovian and its semi-group is given by*

$$\forall t, s \geq 0, I_{t+s} \stackrel{(d)}{=} I(B^{I_t} \circ B'_s) \tag{4}$$

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 semi-group 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 [BLG03]. A flow of bridges is a collection $(B_{s,t})_{s \leq t}$ of bridges which fulfills the following three conditions

1. for any $s < r < t$, $B_{s,t} = B_{s,r} \circ B_{r,t}$ (cocycle property);
2. for any $t_1 < \dots < t_p$, 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);

3. 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 [BLG03] 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 [BLG03] 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 [BLG03]). We call it the Λ -flow of bridges. Using this link, 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*

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

Proof. Let $p \geq 1$ and $0 \leq t_1 < \dots < t_p$. Using the Markov property of $(I_t)_{t \geq 0}$ and the semi-group (4) we know that

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

where (B'_1, \dots, 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

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

Moreover from the independence and stationarity of the increments, $(B_{t_1,t_2}, \dots, B_{t_{p-1},t_p})$ are independent bridges with the same distribution as above. \square

Remark 3.7. *In this section we have defined the Λ -comb using the framework of compositions developed earlier. Another approach could be to start from the Λ -flow of bridges and define the Λ -comb to be the process $(I(B_{0,t}))_{t \geq 0}$. However, it is rather difficult then to prove that the Λ -comb is a Markov process. Using the notion of composition greatly simplifies this task here.*

3.1 Proof of Lemma 3.3

We will need the following continuity result.

Lemma 3.8. *The map $I: B \mapsto I(B)$ 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 $\lim_{n \rightarrow \infty} \|\lambda_n - \text{Id}\|_\infty = 0$ and $\lim_{n \rightarrow \infty} \|B - B^n \circ \lambda_n\|_\infty = 0$. 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

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

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 non-decreasing, 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 blocks and no dust. The fact that the order of the blocks of the composition \mathcal{C} are 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 $V_1^* < \dots < V_K^*$ 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

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

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

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

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

Let us turn to the general case. Let $n \geq 1$ 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. [Ber06]). We can now make use of another continuity result, namely Lemma 1 from [BLG03], to show that the sequence of bridges $(B^n)_{n \geq 1}$ converges in distribution to a bridge B obtained from the mass-partition β . Using Lemma 3.8, we know that $I(B^n)$ converges in distribution to $I(B)$. By uniqueness of the limit, we get that

$$I \stackrel{(d)}{=} I(B),$$

and that B is unique. □

3.2 Proof of Proposition 3.4

Proposition 3.4 will follow by taking the limit in the equality (3). First, let us define more carefully the notion of coagulation.

Coagulation of partitions. Let π and π' be two partitions of $[n]$ or \mathbb{N} . Let $(A_i)_{i \geq 1}$ and $(A'_i)_{i \geq 1}$ be the blocks of π and π' respectively, labelled in the order of their least element. We note $\text{Coag}(\pi, \pi')$, the coagulation of π by π' , the partition whose blocks are

$$\bigcup_{i \in A'_j} A_i, \quad j \geq 1.$$

In words the blocks of π whose labels belong to the same block of π' are merged together. For example if

$$\pi = \{\{1, 4\}, \{2, 5\}, \{3\}\}, \quad \pi' = \{\{1, 3\}, \{2, 4, 5\}\}$$

then

$$\text{Coag}(\pi, \pi') = \{\{1, 3, 4\}, \{2, 5\}\}.$$

Composition of n . Let c be a composition of $[n]$ with b blocks and n_i be the size of the i -th block, in the order of the composition. The vector (n_1, \dots, n_b) is called the *composition of n* associated to c , we will denote it $l(c)$. The compositions of n are the ordered collections (n_1, \dots, n_b) such that $n_1 + \dots + n_b = n$. For example the composition of 5 associated to

$$c = (\{3, 5\}, \{1\}, \{2, 4\})$$

is $l(c) = (2, 1, 2)$. Alternatively, we can represent a composition of n as a partition of $[n]$: the blocks of the partitions are $[n_i] \setminus [n_{i-1}]$ for $1 \leq i \leq b$ with the convention $[0] = \emptyset$. In our example the representation of $(2, 1, 2)$ is

$$\{\{1, 2\}, \{3\}, \{4, 5\}\}.$$

Coagulation of compositions. The last definition is an extension of the notion of coagulation to compositions. Let c be a composition of $[n]$ with b blocks. We will first define the notion of coagulation of c by a composition of b . Let c' be a composition of b represented as a partition. We note $A_1 \leq \dots \leq A_b$ the blocks of c labelled according to their rank, i.e. such that A_1 is the first block etc, and A'_1, \dots, A'_k the blocks of c' labelled according to their least elements. Let π be the partition whose blocks are

$$C_j = \bigcup_{i \in A'_j} A_i, \quad j \leq k.$$

By construction, the blocks of π are formed of adjacent blocks of the composition c . Thus π inherits a natural order on its blocks from c . Let C_i and C_j be two blocks of π , we define

$$C_i \leq' C_j \iff \forall r \in A'_i, \forall l \in A'_j, A_r \leq A_l.$$

The composition (π, \leq') is called the *coagulation of c by c'* and is noted $\text{Coag}(c, c')$. This notion is better understood with an example. Let

$$c = (\{5, 6\}, \{1, 4\}, \{3\}, \{2\}), \quad c' = (1, 2, 1) = \{\{1\}, \{2, 3\}, \{4\}\},$$

then

$$\text{Coag}(c, c') = (\{5, 6\}, \{1, 3, 4\}, \{2\}).$$

Now consider c a composition of $[n]$ with b blocks and c' another composition of $[n]$. We set

$$\text{Coag}(c, c') = \text{Coag}(c, l(c \cap [b])).$$

Proof of Proposition 3.4. The first step to prove Proposition 3.4 is to rephrase the equality (3) in terms of empirical interval-partitions. Let $(\mathcal{C}_t^n)_{t \geq 0}$ be the Markovian nested composition whose transition rates are given by the array $(\tilde{\lambda}_{b,k}, 2 \leq k \leq b < \infty)$.

Let $t, s \geq 0$ and suppose that \mathcal{C}_t^n has b blocks. Let I_t^n be the empirical interval-partition associated to \mathcal{C}_t^n and I_s' be an independent interval-partition distributed as I_s . We now build a new interval partition as follows. We label the blocks of I_t^n from left to right by $1, \dots, b$. Consider $V_1^* < \dots < V_b^*$ the order statistics of independent uniform variables V_1, \dots, V_b . Similarly as above, we can define a bridge B^n from V_1^*, \dots, V_b^* and I_t^n such that $I(B^n) = I_t^n$. Let I be the interval-partition obtained by merging intervals i and j of I_t^n if V_i^* and V_j^* are in the same interval of I_s' , for all i, j . It should be clear that if $B^{I_s'}$ is such that $I(B^{I_s'}) = I_s'$ then

$$I = I(B^n \circ B^{I_s'}).$$

Now notice that the coagulation mechanism we have described to obtain I is exactly the coagulation of \mathcal{C}_t^n by the composition of b obtained from the ordered paintbox using the variables V_1^*, \dots, V_b^* and the interval-partition I_s' . As the latter composition of b is distributed as \mathcal{C}_s' , expression (3) translates into

$$I_{t+s}^n \stackrel{(d)}{=} I(B^n \circ B^{I_s'}). \quad (5)$$

We now take the limit in equation (5). We know that a.s. I_t^n converges to I_t . A similar argument as in the proof of Lemma 3.3 shows that B_t^n converges in distribution to a bridge B_t , and that $I(B_t)$ is distributed as I_t . Using the facts that 1) composition (in the ordinary sense of the word) is continuous in the Skorohod topology and 2) the map I is continuous, shows that the right-hand side of (5) converges in distribution to

$$I(B^{I_t} \circ B'_s).$$

Moreover I_{t+s}^n converges a.s. to I_{t+s} , thus

$$I_{t+s} \stackrel{(d)}{=} I(B^{I_t} \circ B'_s).$$

Let us now show the Markov property of the Λ -comb. Let $p \geq 1$ and $0 \leq t_1 < \dots < t_p$. Using the Markov property of $(C_t^n)_{t \geq 0}$ and repeating the above argument shows that

$$(I_{t_1}^n, \dots, I_{t_p}^n) \stackrel{(d)}{=} (I(B^{I_{t_1}^n}), I(B^{I_{t_1}^n} \circ B'_2), \dots, I(B^{I_{t_1}^n} \circ B'_2 \circ \dots \circ B'_p))$$

where $B^{I_{t_1}^n}$ is such that $I_{t_1}^n = I(B^{I_{t_1}^n})$ and (B'_2, \dots, B'_p) are independent bridges, such that for $k = 2, \dots, p$, B'_k is distributed as $B^{I_{t_k} - t_{k-1}}$. Taking the limit as above shows that

$$(I_{t_1}, \dots, I_{t_p}) \stackrel{(d)}{=} (I(B^{I_{t_1}}), I(B^{I_{t_1}} \circ B'_2), \dots, I(B^{I_{t_1}} \circ B'_2 \circ \dots \circ B'_p))$$

where $I_{t_1} = I(B^{I_{t_1}})$. □

3.3 Dynamical combs

As mentioned in 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, [PW06, PWW11] have studied statistics of the evolving coalescent using a look-down representation, [GPW08] 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 e.g. [KSW14] for the case of Beta-coalescents and [Sch12] 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 yields an evolving 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-Hausdorff topology as in [GPW08]. 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 semi-group. 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 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 $i \geq 2$, we set

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

The Kingman comb is given by

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

It is known [LS18] that the above construction generates the comb associated to the flow of bridges, i.e. the Λ -comb associated to the Kingman coalescent. There are only finitely many teeth of f_K that are larger than s , i.e. such that $T_i \geq s$, say N_s . Let σ be their order, e.g. $\sigma(1)$ is the label of the left-most tooth. Consider $V_1^* < \dots < V_{N_s+1}^*$ the order statistics of $N_s + 1$ i.i.d. uniform variables. This defines N_s intervals, for $1 \leq k \leq N_s$ let M_k be the greatest tooth of f in (V_k^*, V_{k+1}^*) , i.e.

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

We define new variables $(\hat{T}_i)_{i \geq 2}$ as follows

$$\forall i > N_s + 1, \hat{T}_i = T'_i,$$

and

$$\forall i \leq N_s, \hat{T}_i = M_{\sigma(i)} + s.$$

(Notice that as the labels start from 2, σ maps $\{1, \dots, N_s\}$ to $\{2, \dots, N_s + 1\}$ and the sequence $(\hat{T}_i)_{i \geq 2}$ is well-defined.) We define

$$\hat{f}_K = \sum_{i \geq 2} \hat{T}_i \mathbb{1}_{U_i}.$$

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

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

for all continuous bounded functions F . We will show below that the family of operators $(P_t)_{t \geq 0}$ is a semi-group. 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 semi-group. 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 at 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. \square

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

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

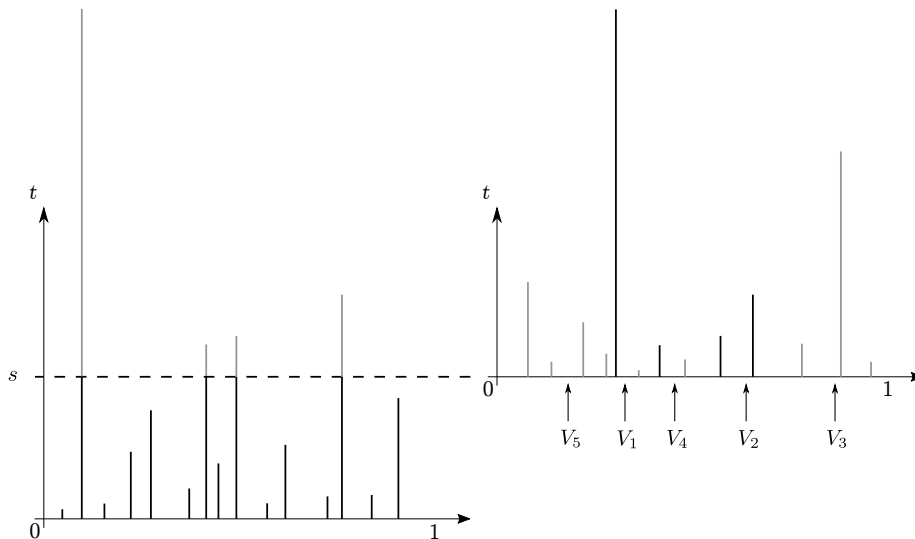


Figure 4: Transition of the evolving Kingman comb. The comb at time s , \hat{f}_K , is represented on the left, and the initial comb f_K is on the right. To obtain \hat{f}_K , one has first to erase the part of the left comb lying above level s . Here we have erased $N_s = 4$ teeth. Then throw $N_s + 1$ uniform variables V_1, \dots, 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 in each of these intervals, represented in black on the right-hand comb, and paste it in place of the erased tooth, as represented in grey on the left-hand comb.

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 spaces have already been considered in [GPW06, Gro07]. 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 4.5). While separability is a very natural hypothesis when considering measured metric 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 spaces to account for non-separable UMS.

Let us recall the heuristic of our approach and give a short outline of this section. Under minimal assumptions which are described in Section 4.1, we can obtain a coalescent by sampling from an UMS. We can define a topology on the space of UMS by saying that a sequence of UMS converges if the associated sequence of coalescents converges weakly as probability measures. In the separable case, the Gromov reconstruction theorem ensures that spaces that are indistinguishable have their supports in isometry. In general this result does not hold, we want to obtain a similar result for general UMS. In order to that, we introduce in Section 4.2 the notion of backbone of an UMS. An 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 4.3, 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 4.11 for a rigorous statement). Finally Section 4.4 is dedicated to showing Corollary 1.13, i.e. that we can always find a comb metric space weakly isometric to a given UMS, and Section 4.5 is devoted to showing Corollary 1.12, i.e. that in the separable case we can always find a comb metric space isometric to a given UMS.

4.1 General UMS and the Gromov-weak topology

Naively, an UMS is any metric measured space (U, d, μ) where d is an ultrametric and μ is a probability measure defined on the Borel σ -field of (U, d) . However, two problems arise from this definition. First, the space U can be very general. We need to impose some measurability conditions in order to make sense of the coalescent associated to an UMS. Second, roughly speaking, the Borel σ -field of a non-separable metric space tends to be large, and fewer measures can be defined on it. For example, by restricting our attention to Borel measure, we cannot find an UMS that is associated to a coalescent with dust. (More precisely, this question is related to a measure theoretic problem known as the Banach-Ulam problem, and a positive answer to it would require additional axioms, see Section F of the Appendix.) These two considerations lead us to the following definitions.

Definition 4.1. *Let (U, d) be an ultrametric space. The genealogical σ -field \mathcal{E}_{gen} associated to (U, d) is defined as*

$$\mathcal{E}_{gen} = \sigma(\{B(x, t), x \in U, t > 0\})$$

where $B(x, t)$ is the open ball of radius t centered at x .

Example 4.2. *Consider any set U as above endowed with the metric*

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

In this case \mathcal{E}_{gen} is the countable-cocountable σ -field.

Definition 4.3. *An UMS is an ultrametric measured space (U, d, \mathcal{F}, μ) , where μ is a probability measure defined on \mathcal{F} and such that*

- (i) $\mathcal{E}_{gen} \subseteq \mathcal{F}$;
- (ii) the ultrametric d is a measurable map from $U \times U$ to \mathbb{R}_+ when $U \times U$ is endowed with the product σ -field $\mathcal{F} \otimes \mathcal{F}$.

Remark 4.4. *The second assumption has rather drastic consequences on the space U . For example, it implies that the cardinal of U is not greater than that of \mathbb{R} .*

The inclusion $\mathcal{E}_{gen} \subseteq \mathcal{B}(U)$ always holds. Thus we allow ourselves to define measures on a σ -field smaller than the Borel σ -field as desired. Moreover, as we will see below, point (ii) is necessary to obtain a coalescent out of an i.i.d. sequence in U .

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

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

(Notice that here our measurability assumption is key.) Alternatively, we can see this coalescent as a random pseudo-ultrametric on \mathbb{N} defined as

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

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

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

The distribution of this pseudo-ultrametric is called the *distance matrix distribution* of the UMS.

Remark 4.5. *From a topological point of view, a pseudo-ultrametric on \mathbb{N} can be seen as an element of $\mathbb{R}_+^{\binom{\mathbb{N}}{2}}$ endowed with its product topology. Notice that in this case, the correspondence between pseudo-ultrametrics on \mathbb{N} and coalescents outlined above is an homeomorphism.*

We use distance matrix distributions to define a topology on the space of UMS. Consider a sequence $(U_n, d_n, \mu_n)_{n \geq 1}$ of UMS, and denote $(\nu_n)_{n \geq 1}$ the associated sequence of distance matrix distributions. We say that the sequence $(U_n, d_n, \mu_n)_{n \geq 1}$ converges in the Gromov-weak topology to (U, d, μ) 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 $\mathbb{R}^{\binom{\mathbb{N}}{2}}$.

Remark 4.6. *In the separable case, $\mathcal{E}_{gen} = \mathcal{B}(U)$ always hold, and d is measurable. Thus in the separable case, Definition 4.3 coincides with the “naive” definition of an UMS. We recover the usual framework of metric measured spaces.*

4.2 Backbone

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

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

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

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

Then the space (T, d_T) is a real tree [Eva07] whose leaves can be identified with (U, d) .

Definition 4.7 (Backbone of T). *Define*

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

(note that f is measurable since $\mathcal{E}_{gen} \subseteq \mathcal{F}$) 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 4.8 states that even if the underlying UMS is not separable, the backbone is always a separable tree. Secondly, 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 have 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

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

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 $\{0, 1\}$. Consider two independent random variables (X, A) and (Y, B) distributed according to the above measure. We see that these

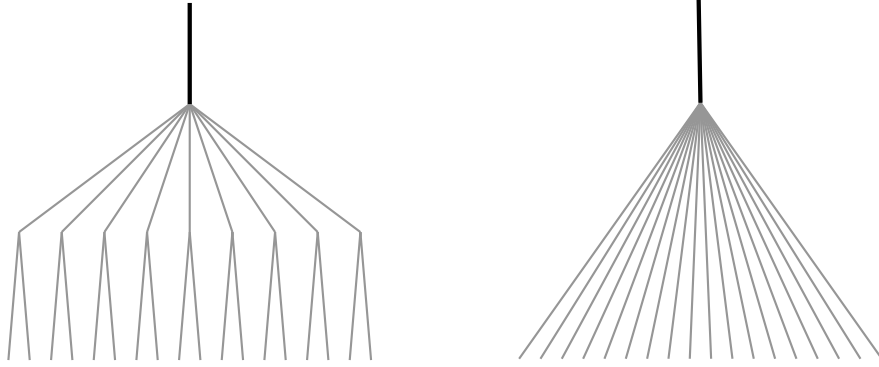


Figure 5: Left panel: The bifurcating star-tree. Right panel: The bifurcating star-tree simplified according to the metric \tilde{d} . In both cases, the backbone is illustrated with a bold black line and the subtrees attached to it with thin grey lines.

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, i.e. the bifurcating star-tree is a star-tree (see Figure 5 right panel).

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 an UMS (U, d, \mathcal{F}, μ) . We introduce the distance

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

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, \mathcal{F}, μ) is the same as the coalescent obtained by sampling from $(U, \tilde{d}, \mathcal{F}, \mu)$.

Proposition 4.8. (i) $(\mathcal{S}, d_{\mathcal{S}})$ is a separable tree.

(ii) The map

$$\psi: \begin{cases} (U, \mathcal{F}) \rightarrow (\mathcal{S}, \mathcal{B}(\mathcal{S})) \\ x \mapsto (x, f(x)) \end{cases}$$

is measurable and we define $\mu_{\mathcal{S}} := \psi \star \mu$, the pushforward measure (on $(\mathcal{S}, \mathcal{B}(\mathcal{S}))$) 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 a.s. for all $i, j \geq 1$, $d(X_i, X_j) = \tilde{d}(X_i, X_j)$.

Proof. We start by proving (i). The fact that \mathcal{S} is a 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 $\epsilon > 0$ and let $(x, s) \in U \times \mathbb{R}_+$ in \mathcal{S} . We can find $t \in \mathbb{Q}_+$ such that $t > s \geq f(x)$ and $t - s < \epsilon$. 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

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

and thus $d_T((x, s), (x_i^t, t)) < \epsilon$. 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 note

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

the *clade* generated by (x, t) . In a genealogical interpretation, $C(x, t)$ is the progeny of (x, t) i.e. 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 $t \rightarrow \infty$). 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 $x, y \in U$, $d(x, y) \leq \tilde{d}(x, y)$. Thus the probability that $d(X, Y) \neq \tilde{d}(X, Y)$ can be written

$$\mathbb{P}(d(X, Y) \neq \tilde{d}(X, Y)) = \iint \mathbb{1}_{\{d(x, y) < \tilde{d}(x, y)\}} \mu(dx) \mu(dy) = \int \mu(dx) \int \mu(dy) \mathbb{1}_{\{d(x, y) < f(x)\}} = 0,$$

where the last equality can be seen by writing

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

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

Remark 4.9 (Backbone and marked metric space). *An object similar to the backbone appears in [Guf18] using the framework of marked metric spaces introduced in [DGP11]. We can interpret the backbone as a marked metric 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{U} defined in [Guf18]. The novelty of the present work is that we do not consider as given the backbone of the space, but starting from the UMS, we 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 F of the Appendix.

Moreover, the link between backbones and marked metric spaces enables us to use the work of [DGP11]. 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.

4.3 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 4.10. *Let (U, d, μ) and (U', d', μ') be two UMS with respective backbones (\mathcal{S}, μ_S) and (\mathcal{S}', μ'_S) . We say that Φ is an isomorphism from \mathcal{S} to \mathcal{S}' if*

- (i) 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 $\Phi((x, t)) = (x', t)$, i.e. Φ preserves the second coordinate.

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

We want to show the following result. In words, having the same distance matrix distribution is equivalent to being weakly isometric.

Proposition 4.11. *Let (U, d, \mathcal{F}, μ) and $(U', d', \mathcal{F}', \mu')$ be two UMS with respective backbones (\mathcal{S}, μ_S) and (\mathcal{S}', μ'_S) . We suppose that the two backbones are complete metric spaces. Then the two spaces (\mathcal{S}, μ_S) and (\mathcal{S}', μ'_S) are isomorphic iff the distance matrix distribution associated (U, d, \mathcal{F}, μ) and $(U', d', \mathcal{F}', \mu')$ are identical.*

Let us compare this result to the original result from [Gro07]. In the separable case, if two UMS share the same coalescent then their supports 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 4.11 states is that in this case there is only a correspondence between coarsenings of the UMS, i.e. 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 4.11 is a direct consequence of the following lemma, which shows that the distance matrix distribution of an UMS can be recovered from an i.i.d. sequence of points of the backbone.

Lemma 4.12. *Let $(X_i)_{i \geq 1}$ be an i.i.d. sequence in U sampled according to μ . Then a.s.*

$$\forall i, j \geq 1, d(X_i, X_j) = d_S((X_i, f(X_i)), (X_j, f(X_j))) + \frac{f(X_i) + f(X_j)}{2} \quad (6)$$

and

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

Proof. We know from Proposition 4.8 that a.s. for any $i, j \geq 1$, $\tilde{d}(X_i, X_j) = d(X_i, X_j)$. 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

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

and

$$d_S((X_i, f(X_i)), (X_j, f(X_j))) = d(X_i, X_j) - \frac{f(X_i) + f(X_j)}{2}.$$

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. \square

It remains to show the converse proposition, i.e. 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 [Gro07].

Proof of Proposition 4.11. We say that a sequence $(x_i, t_i)_{i \geq 1}$ in \mathcal{S} is equidistributed if for any $A \in \mathcal{S}$,

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

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

$$D: \begin{cases} \mathcal{S}^{\mathbb{N}} \rightarrow \mathbb{R}^{\binom{\mathbb{N}}{2}} \\ (x_i, t_i)_{i \geq 1} \mapsto (d_S((x_i, t_i), (x_j, t_j)) + \frac{t_i + t_j}{2})_{i, j \geq 1}. \end{cases}$$

and let D' be the analogous map for U' . Then Lemma 4.12 shows that the pushforward measure $\mu_{\mathcal{S}}^{\otimes \mathbb{N}} \star D$ is the distance matrix distribution associated to U . Similarly $\mu_{\mathcal{S}'}^{\otimes \mathbb{N}} \star D'$ 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, i.e. such that

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

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

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

and then using (6) we obtain

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

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

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

It should be 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 these sets are 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 $\mu(B(x, t + \epsilon)) > 0$ for any $\epsilon > 0$. Writing

$$B(x, t + \epsilon) = \{y \in U : d_{\mathcal{S}}((x, t + \epsilon), (y, t + \epsilon))\},$$

as $(x_i, f(x_i))_{i \geq 1}$ is equidistributed, we see that we can find $(x_i, f(x_i))$ such that $(x_i, t + \epsilon) = (x, t + \epsilon)$.

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

Remark 4.13. *According to the correspondence between backbones and marked metric spaces outlined earlier, Proposition 4.11 is similar to the more general Theorem 1 in [DGP11], 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 [DGP11] would only provide an isometry between the supports of the backbones whereas here we obtain a global isometry.*

Remark 4.14. *The results of this section show that the backbone of an 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 [FHP17], and addressed in [For18], 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 [FHP17]. 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’.*

4.4 UMS associated to a comb, completion

In this section we construct the UMS associated to a comb. An important assumption of Proposition 4.11 is that the backbones of the UMS are complete metric spaces. We will show that the UMS associated to a comb enjoys this property up to the addition of a countable number of points.

Let $(I_t)_{t \geq 0}$ be a nested interval-partition. Recall that we have defined the comb function f_I associated to $(I_t)_{t \geq 0}$ as

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

and that we have naively defined the ultrametric distance associated to $(I_t)_{t \geq 0}$ as

$$\forall x, y \in [0, 1], d_I(x, y) = \mathbf{1}_{\{x \neq y\}} \sup_{[x \wedge y, x \vee y]} f_I.$$

The natural sampling measure associated to $([0, 1], d_I)$ is the Lebesgue measure Leb defined on the usual Borel σ -field on $[0, 1]$. First we need to check that $([0, 1], d_I, \mathcal{B}([0, 1]), \text{Leb})$ fulfills the assumptions of Definition 4.3.

First, let $x \in [0, 1]$ and $t \geq 0$. If $x \in I_t$, then $B(x, t)$ is the interval component of I_t to which x belongs. Otherwise, $B(x, t) = \{x\}$. In both case $B(x, t) \in \mathcal{B}([0, 1])$ and thus $\mathcal{E}_{\text{gen}} \subseteq \mathcal{B}([0, 1])$. Moreover, it is straightforward to check that d_I is $\mathcal{B}([0, 1]) \otimes \mathcal{B}([0, 1])$ measurable. The comb metric space associated to $(I_t)_{t \geq 0}$ is the UMS $([0, 1], d, \mathcal{B}([0, 1]), \text{Leb})$.

Remark 4.15. *If \mathcal{B}_I denotes the Borel σ -field on $[0, 1]$ induced by the metric d_I , it is not true in general that $\mathcal{B}_I \subseteq \mathcal{B}([0, 1])$. It is crucial to define the sampling measure on \mathcal{E}_{gen} rather than on \mathcal{B}_I . Note that this holds here since the sampling measure Leb is defined on $\mathcal{B}([0, 1])$ and $\mathcal{E}_{\text{gen}} \subseteq \mathcal{B}([0, 1])$.*

A key assumption to use Proposition 4.11 is that we consider metric spaces that have a complete backbone. However, in general the backbone associated to a comb metric space does not enjoy this property. Let us see two examples of combs that are not associated to a complete backbone.

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

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

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 non-decreasing 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

$$I_t' = \begin{cases} \emptyset & \text{if } t < 1/2 \\ I_{t-1/2}^2 & \text{otherwise.} \end{cases}$$

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” 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.

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 $([0, 1] \cup F, \bar{d}_I)$ 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 1.13 now directly follows from the various results we have shown.

Proof of Corollary 1.13. Let (U, d, \mathcal{F}, μ) be an UMS, and let $(\Pi_t)_{t \geq 0}$ be the associated coalescent. Using Theorem 1.5 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 comb metric space whose backbone is complete which has the same distance matrix distribution than (U, d, \mathcal{F}, μ) . Using Proposition 4.11 ends the proof. \square

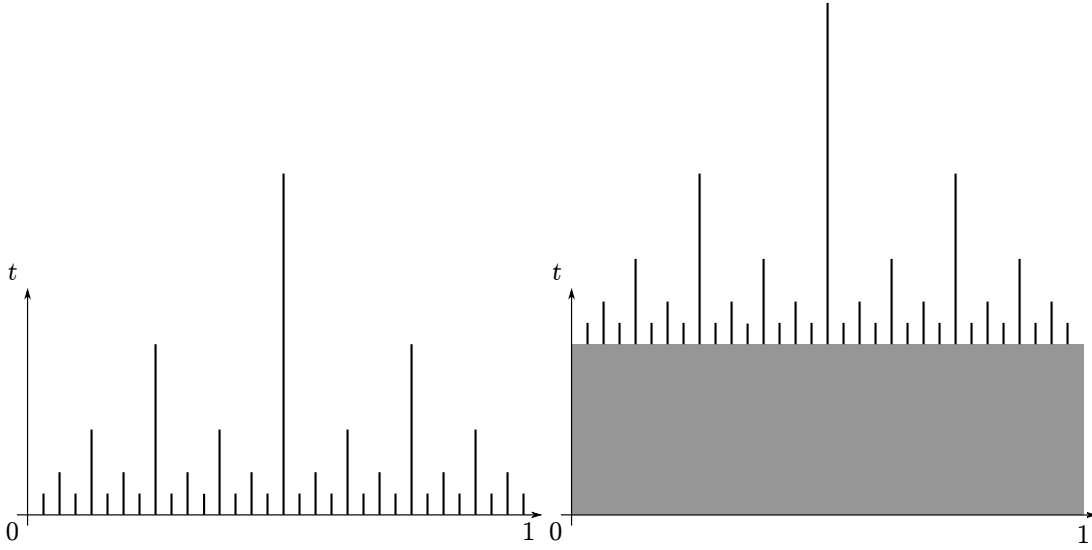


Figure 6: Left panel: The diadic comb. Right panel: The comb $(I'_t)_{t \geq 0}$.

4.5 The separable case

In this section we consider the case of separable UMS and prove Corollary 1.12 that shows that in the separable case we can find a direct isometry between any UMS and a comb metric space.

Let (U, d, μ) be a separable UMS where $U = \text{supp}(\mu)$. The key point is that in this case $f \equiv 0$. Thus, the application ψ defined in Proposition 4.8 is a bijective isometry from (U, d) to the set of leaves of the corresponding backbone. Moreover the backbone of (U, d) is complete iff (U, d) is complete. We will need the following lemma to prove Corollary 1.12.

Lemma 4.18. *Any separable ultrametric space (U, d) can be endowed with a measure μ such that $\text{supp}(\mu) = U$.*

Proof of Corollary 1.12. Let (U, d) be a separable UMS and μ be the measure obtained from Lemma 4.18. Consider $(\Pi_t)_{t \geq 0}$ the coalescent obtained from sampling in U according to μ . For any $t > 0$, (U, d) only has countably many balls of radius t , and as $U = \text{supp}(\mu)$ each of them has positive mass. Thus by the law of large numbers, all the blocks of Π_t are infinite. The coalescent $(\Pi_t)_{t \geq 0}$ has no dust.

Using Theorem 1.5 we can find a nested interval-partition without dust $(I_t)_{t \geq 0}$ such that the distance matrix distribution associated to $([0, 1], d_I, \text{Leb})$ is similar to that of (U, d, μ) . Let $(U \cup F, \bar{d}_I)$ be the completion of Proposition 4.16. It is immediate that the support of the extension of the Lebesgue measure is the completion of the set $\{f_I = 0\}$. Let ψ and ψ_I be the map of Proposition 4.8 applied to (U, d) and $([0, 1] \cup F, \bar{d}_I)$ respectively. Using Proposition 4.11 we can find an isometry Φ between the respective backbones of these spaces. The isometry between (U, d) and the comb metric space is obtained by composing ψ , Φ and ψ_I . \square

Proof of Lemma 4.18. 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 $\text{supp}(\mu) = U$. \square

Remark 4.19. Note that a similar construction was mentioned in [LUB17], where the resulting measure was referred to as the "visibility measure".

Acknowledgements. The authors thank *Center for Interdisciplinary Research in Biology* (CIRB, Collège de France) for funding.

References

- [Ber06] Jean Bertoin. *Random Fragmentation and Coagulation Processes*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2006.
- [BLG03] Jean Bertoin and Jean-François Le Gall. Stochastic flows associated to coalescent processes. *Probability Theory and Related Fields*, 126(2):261–288, 2003.
- [DGP11] Andrej Depperschmidt, Andreas Greven, and Peter Pfaffelhuber. Marked metric measure spaces. *Electronic Communications in Probability*, 16:174–188, 2011.
- [DJ91] Peter Donnelly and Paul Joyce. Consistent ordered sampling distributions: Characterization and convergence. *Advances in Applied Probability*, 23(2):229–258, 1991.
- [Eva07] Steven N Evans. *Probability and Real Trees: École d'Été de Probabilités de Saint-Flour XXXV-2005*. Springer, 2007.
- [FHP17] Noah Forman, Chris Haulk, and Jim Pitman. A representation of exchangeable hierarchies by sampling from random real trees. *Probability Theory and Related Fields*, 2017.
- [For18] Noah Forman. Mass-structure of weighted real trees. *ArXiv e-prints*, 2018.
- [Fre93] David H Fremlin. Real-valued-measurable cardinals. In *Israel Mathematical Conference Proceedings*, volume 6, pages 151–304, 1993.
- [Gne97] Alexander V. Gnedin. The representation of composition structures. *The Annals of Probability*, pages 1437–1450, 1997.
- [GPW06] Andreas Greven, Peter Pfaffelhuber, and Anita Winter. Convergence in distribution of random metric measure spaces: Lambda-coalescent measure trees. *Probability Theory and Related Fields*, 2006.
- [GPW08] Andreas Greven, Peter Pfaffelhuber, and Anita Winter. Tree-valued resampling dynamics: Martingale problems and applications. *Probability Theory and Related Fields*, 2008.
- [Gro07] Mikhail Gromov. *Metric structures for Riemannian and non-Riemannian spaces*. Springer Science & Business Media, 2007.

- [Guf18] Stephan Gufler. A representation for exchangeable coalescent trees and generalized tree-valued Fleming-Viot processes. *Electronic Journal of Probability*, 23:42 pp., 2018.
- [Kin78] John FC Kingman. The representation of partition structures. *Journal of the London Mathematical Society*, 2(2):374–380, 1978.
- [Kin82] John FC Kingman. The coalescent. *Stochastic Processes and their Applications*, 13(3):235 – 248, 1982.
- [KSW14] Götz Kersting, Jason Schweinsberg, and Anton Wakolbinger. The evolving beta coalescent. *Electronic Journal of Probability*, 19:27 pp., 2014.
- [Lam17] Amaury Lambert. Random ultrametric trees and applications. *ESAIM: Proceedings and Surveys*, 60:70–89, 2017.
- [LS18] Amaury Lambert and Emmanuel Schertzer. Recovering the Brownian coalescent point process from the Kingman coalescent by conditional sampling. *Bernoulli*, Eprint arXiv:1611.01323, 2018.
- [LUB17] Amaury Lambert and Gerónimo Uribe Bravo. The comb representation of compact ultrametric spaces. *p-Adic Numbers, Ultrametric Analysis and Applications*, 9(1):22–38, 2017.
- [Pit99] Jim Pitman. Coalescents with multiple collisions. *Annals of Probability*, pages 1870–1902, 1999.
- [PW06] Peter Pfaffelhuber and Anton Wakolbinger. The process of most recent common ancestors in an evolving coalescent. *Stochastic Processes and their Applications*, 116(12):1836–1859, 2006.
- [PWW11] Peter Pfaffelhuber, Anton Wakolbinger, and Heinz Weisshaupt. The tree length of an evolving coalescent. *Probability theory and related fields*, 151(3):529–557, 2011.
- [RP81] Chris Rogers and Jim Pitman. Markov functions. *The Annals of Probability*, 9(4):573–582, 08 1981.
- [Sag99] Serik Sagitov. The general coalescent with asynchronous mergers of ancestral lines. *Journal of Applied Probability*, 36(4):1116–1125, 12 1999.
- [Sch00] Jason Schweinsberg. Coalescents with simultaneous multiple collisions. *Electronic Journal of Probability*, 5, 2000.
- [Sch12] Jason Schweinsberg. Dynamics of the evolving Bolthausen-Sznitman coalescent. *Electronic Journal of Probability*, 17:50, 2012.

A Exchangeable hierarchies

The aim of this section is to recall some results derived in [FHP17] 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 [FHP17], but could 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

1. for $x \in X$, $\{x\} \in \mathcal{H}$, $X \in \mathcal{H}$ and $\emptyset \in \mathcal{H}$;
2. given $A, B \in \mathcal{H}$, then $A \cap B$ is either A , B or \emptyset .

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

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

is a hierarchy. The hierarchy \mathcal{H} encodes the genealogical structure of (U, d) , i.e. 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.*

The main object studied in [FHP17] 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 [FHP17] for a definition of the σ -field associated to hierarchies) is called exchangeable if

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

In a similar way that exchangeable coalescents are obtained by sampling in UMS, exchangeable hierarchies are obtained by sampling in hierarchies on measured 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 [FHP17] 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 non-singleton 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

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

is an interval hierarchy. Theorem 4 in [FHP17] states that any exchangeable hierarchy on \mathbb{N} can be obtained by sampling in an interval hierarchy. This is the direct equivalent of our Theorem 1.5 that states that any exchangeable coalescent can be obtained by sampling in a random nested interval-partition.

Consider a measured 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 [Eva07]). 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 [FHP17] states that any exchangeable hierarchy can be obtained by sampling in the hierarchy associated to a random measured rooted real-tree. In our framework, we have seen that a

nested interval-partition can be seen as an ultrametric space, and in Section 4.4 we have seen how this ultrametric space is embedded in a real-tree. Again we have proved here the reformulation of Theorem 5 from [FHP17].

In a subsequent work, one of the authors has introduced the notion of mass-structural isomorphism [For18]. 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, e.g. the work of [FHP17] 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 [FHP17], but constitute an extension of their work.

Finally, we wish to stress two things. First, most of the difficulties that Section 4 deal with stem from the fact that we consider non-separable 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 [FHP17]. Moreover, this provides a new approach to the question of dynamical genealogies, with the introduction of the dynamical comb.

B Details of the proof of Theorem 2.5

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

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

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 $(I_t)_{t \in \mathbb{Q}_+}$.

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 from the space of compositions of $[n]$ to \mathbb{R} ,

$$\forall \pi, \hat{Q}L f(\pi) = LQ f(\pi),$$

where L is the operator defined in Section 3.

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

$$\begin{aligned} \hat{Q}Lf(\pi) &= \sum_{\pi' \in P_n} q_{\pi, \pi'} (Lf(\pi') - Lf(\pi)) \\ &= \sum_{\pi' \in P_n} q_{\pi, \pi'} \left(\sum_{c' \in O(\pi')} \frac{1}{\text{Card}(\pi')!} f(c') - \sum_{c \in O(\pi)} \frac{1}{\text{Card}(\pi)!} f(c) \right) \\ &= \sum_{\pi' \in P_n} \sum_{c' \in O(\pi')} q_{\pi, \pi'} \frac{1}{\text{Card}(\pi')!} f(c') - \sum_{c \in O(\pi)} \sum_{k=2}^b \frac{1}{\text{Card}(\pi)!} \binom{b}{k} \lambda_{b, k} f(c). \end{aligned}$$

Similarly, we have

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

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

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

D Measurability of separable rooted trees

In this section we prove the claim made in the proof of Proposition 4.8 that the Borel σ -field of a rooted tree is induced by the clades of the tree. Let us be more specific.

We consider a real-tree (T, d) with a particular point $\rho \in T$ that we call the root. For $x, y \in T$, we note $[x, y]$ the unique geodesic with endpoints x and y (see [Eva07]). 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

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

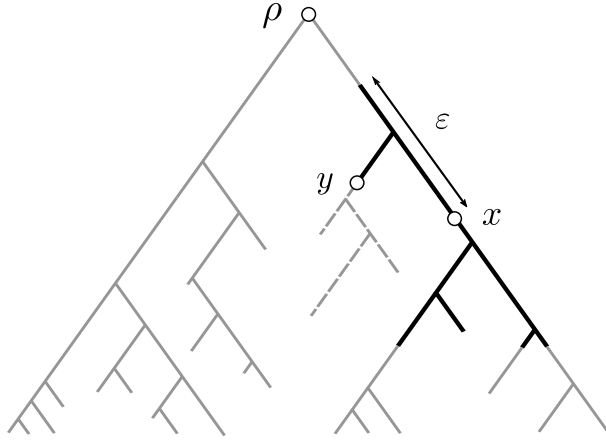


Figure 7: A tree rooted at ρ . The ball of radius ϵ and center x is represented by the black bold lines. An example of $y \in S(x, \epsilon)$ is given, and its corresponding clade $C(y)$ is represented by grey dashed lines.

Remark D.1. *Our goal in the proof of Proposition 4.8 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 $\epsilon > 0$, we assume that $\epsilon < d(x, \rho)$. We denote $B(x, \epsilon)$ the open ball centered in x with radius ϵ , and $S(x, \epsilon)$ the sphere of center x and radius ϵ , i.e.

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

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

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

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

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

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 $\{y_\eta : y \in S(x, \epsilon) \setminus \{a\}\}$ is not countable, we can find an uncountable subset of $S(x, \epsilon)$ such that any two points lie at distance at least η . This is not possible due to separability.

E Comb completion

In this section we prove Proposition 4.16, i.e. 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

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

and

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

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

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

We will simply note 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

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

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

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

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 \bar{d}_I^* . As

$$\frac{|t_n - t_m|}{2} \leq d_{\mathcal{S}_I}((x'_n, t_n), (x'_m, t_m)),$$

the sequence $(t_n)_{n \geq 1}$ is Cauchy and converges to a limit that we note t . Each point x'_n can be written as $x'_n = (x_n, a_n)$ with $x_n \in [0, 1]$ and $a_n \in \{0, r, l\}$. 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 non-decreasing and converges to x .

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

$$\lim_{n \rightarrow \infty} \sup_{m \geq n} \bar{d}_I(x'_n, x'_m) - \frac{t_n + t_m}{2} \leq 0,$$

which directly implies that

$$\lim_{\epsilon \rightarrow 0} \sup_{[x-\epsilon, x]} f_I \leq t.$$

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

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

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

$$\lim_{\epsilon \rightarrow 0} \sup_{[x-\epsilon, x]} f_I = f(x).$$

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

$$\lim_{\epsilon \rightarrow 0} \sup_{[x-\epsilon, x]} f_I \geq f(x).$$

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

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

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

Remark E.2. *This completion is already present in [LUB17] in the compact case. In this case, we have $\mathcal{R} = \mathcal{L} = \{f_I > 0\}$.*

F The link between dust and the Banach-Ulam problem

In this section we discuss in more details the reasons that pushed us to allow the sampling measure to be defined on \mathcal{E}_{gen} rather than on the Borel σ -field. We will try to answer the following question: can we find an ultrametric space (U, d) and a Borel sampling measure μ such that the coalescent obtained by sampling from (U, d, μ) has dust? We will show that a positive answer to this question requires additional axioms than the usual axioms ZFC (Zermelo-Fraenkel-Choice) of set theory. Thus, as we want to provide an extension of the theory of metric measured spaces that accounts for dust, it is natural to allow the measures to be defined on \mathcal{E}_{gen} .

Our strategy is to link this question with a deep measure-theoretic problem known as the *Banach-Ulam problem*. The Banach-Ulam problem can be phrased as follows: can we find a set X and a diffuse probability measure μ defined on the power set $\mathcal{P}(X)$ of X ($\mathcal{P}(X)$ is the set of all subsets of X)? More precisely, we will show the following result.

Lemma F.1. *Consider (U, d, μ) an ultrametric measured space where μ is defined on the Borel σ -field of (U, d) . Suppose that the coalescent associated to (U, d, μ) has dust. Then we can find a set X and a diffuse probability measure ν defined on $\mathcal{P}(X)$.*

A treatment of the Banach-Ulam problem requires advanced set theoretic tools, we refer the interested reader to [Fre93]. We simply mention that if a solution to the Banach-Ulam problem exists, then its cardinal is inaccessible. The existence of inaccessible cardinals cannot be deduced from the usual axioms ZFC of set theory and requires additional axioms.

Let (U, d) be an ultrametric space and μ be a Borel measure. We recall that the support $\text{supp}(\mu)$ of μ is defined as the intersection of all the closed subsets of mass 1. Equivalently, the support can be defined as

$$\text{supp}(\mu) = \{x \in U : \forall \epsilon > 0, \mu(B(x, \epsilon)) > 0\}.$$

We will need the following lemma that relates the presence of dust to the mass of the support.

Lemma F.2. *Let (U, d, μ) be an UMS, and $(\Pi_t)_{t \geq 0}$ be the associated coalescent. Then $(\Pi_t)_{t \geq 0}$ has dust iff $\mu(\text{supp}(\mu)) < 1$.*

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.

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

Suppose that $X_i \in \text{supp}(\mu)$. Then for any $t > 0$, $\mu(B(X_i, t)) > 0$, thus a.s. there are infinitely many other $(X_j)_{j \geq 1}$ 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, i.e. 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 $f_i = \mu(B(X_i, t)) > 0$. \square

We are now ready to prove the link with the Banach-Ulam problem.

Proof of Lemma F.1. As $(\Pi_t)_{t \geq 0}$ has dust, Lemma F.2 ensures that $\mu(\text{supp}(\mu)) < 1$. Consider $t > 0$ and let $(B_\alpha^t)_{\alpha \in A_t}$ be the collection of open balls of radius t with zero mass, where A_t is just an index set. We know that

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

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

$$x \sim y \iff d(x, y) < \epsilon$$

and denote X the quotient space of $\bigcup_{t>\epsilon} \bigcup_{\alpha \in A_t} B_\alpha^t$ for the relation \sim . We define the quotient map

$$\phi: \begin{cases} U \rightarrow X \\ x \mapsto \{y \in U : d(x, y) < \epsilon\} \end{cases}$$

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

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

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. \square