A representation for exchangeable coalescent trees and generalized tree-valued Fleming-Viot processes

Stephan Gufler*

Abstract

We give a de Finetti type representation for exchangeable random coalescent trees (formally described as semi-ultrametrics) in terms of sampling iid sequences from marked metric measure spaces. We apply this representation to define versions of tree-valued Fleming-Viot processes from a Ξ-lookdown model. As state spaces for these processes, we use, besides the space of isomorphy classes of metric measure spaces, also the space of isomorphy classes of marked metric measure spaces and a space of distance matrix distributions. This allows to include the case with dust in which the genealogical trees have isolated leaves.

Contents

*Goethe-Universität, Institute of Mathematics, 60629 Frankfurt am Main, Germany
Current affiliation: Technion, Faculty of Industrial Engineering and Management, Haifa 3200003, Israel
E-mail: stephan.gufler@gmx.net

1 Introduction

1.1 Some background on coalescent trees, ultrametrics, and metric measure spaces

In population genetics, coalescents are common models for the genealogy of a sample from a population. The Kingman coalescent [34] is a partition-valued process in which each individual of the sample forms its own block at time 0, and as we look into the past, each pair of blocks merges independently at constant rate. These blocks stand for the families of individuals that have a common ancestor at given times in the past. Generalizations of the Kingman coalescent include the Λ-coalescent (Pitman [45], Sagitov [47], Donnelly and Kurtz [16]) where multiple blocks are allowed to merge to a single block at the same time, and the Ξ-coalescent (Möhle and Sagitov [42], Schweinsberg [48]) where several clusters of blocks may also merge simultaneously.

A (semi-)ultrametric ρ is a (semi-)metric that satisfies the strong triangle inequality $\max\{\rho(x,y), \rho(y,z)\} \geq \rho(x,z)$. A realization of a coalescent for an infinite sample can be expressed as a càdlàg path $(\pi_t, t \in \mathbb{R}_+)$ with values in the space of partitions of $\mathbb{N}$ such that π_t is a coarsening of π_s for all $s \leq t$. We assume that for each pair of integers, there is a time t such that the elements of this pair are in a common block of π_t. Then $(\pi_t, t \in \mathbb{R}_+)$ can equivalently be expressed as a semi-ultrametric ρ on $\mathbb{N}$ such that for all

$t \in \mathbb{R}_+$ and $i, j \in \mathbb{N}$,

$$\rho(i, j) \leq 2t \quad \text{if and only if } i \text{ and } j \text{ are in the same block of } \pi_t, \qquad (1.1)$$

and (1.1) yields a one-to-one correspondence between these càdlàg paths and the semi-ultrametrics on $\mathbb{N}$, cf. [21, Example 3.41] and [22, p. 262].

Evans [20] studies the completion of the random ultrametric space associated with the Kingman coalescent which he endows with a probability measure such that the mass on each ball is given by the asymptotic frequency of the corresponding family, and a class of more general coalescents is studied by Berestycki et al. [3].

Remark 1.1. Let us briefly recall the well-known correspondence between ultrametric spaces and real trees to which we will refer to explain main concepts in this article. A real tree is a metric space (T, d) that is tree-like in the sense that (i) no subspace is homeomorphic to the unit circle, and (ii) for each $x, y \in T$, there exists an isometry ι from the real interval $[0, d(x, y)]$ to T with $\iota(0) = x$ and $\iota(d(x, y)) = y$, see e. g. Evans [21] for an overview. An ultrametric space (X, ρ) can be isometrically embedded into the real tree (T, d) that is obtained by identifying the elements with distance zero of the semi-metric space $(\mathbb{R}_+ \times X, d)$ given by $d((s, i), (t, j)) = \max\{\rho(i, j) - s - t, |s - t|\}$. Then T equals the set $\mathbf{U}^0$ in [22, p. 262] with $X = \mathbb{N}$ and the metrics d here and in [22, p. 262] coincide up to a factor 2. Clearly, (X, ρ) is isometric to the subspace $\{0\} \times X$ of the leaves of (T, d). For a semi-ultrametric space (X, ρ), we identify the elements with distance zero to obtain an ultrametric space which we associate with a real tree (T, d) as above. A related embedding of an ultrametric space is given in [30, Section 6].

As in Remark 1.1, a semi-ultrametric on $\mathbb{N}$ can be considered as an infinite tree whose leaves are labeled by the elements of $\mathbb{N}$. Often these labels are not relevant, for instance, when they only record the order in which iid samples from a population are drawn. To remove the labels, we could pass to the isometry class. However, the asymptotic block frequencies in the coalescent given by an ultrametric on $\mathbb{N}$ are not determined by the isometry class, as one may apply an infinite permutation without changing the isometry class. To retain just this information besides the metric structure, we can take a measure-preserving isometry class of the completion of the ultrametric space that is endowed with a probability measure that charges each ball with the asymptotic frequency of the corresponding block, if such a probability measure exists. This probability measure can equivalently be described as the weak limit of the uniform probability measures on the individuals $1, \ldots, n$, as $n \to \infty$. Then we obtain the description by isomorphy classes of metric measure spaces of Greven, Pfaffelhuber, and Winter [25] that was applied to Λ-coalescents in the dust-free case. We speak of the dust-free case if the semi-ultrametric space has no isolated points, which means that the coalescent tree has no isolated leaves. Greven, Pfaffelhuber, and Winter [25] also show that their approach is not directly applicable to Λ-coalescents with dust. The most elementary example for the case with dust is the star-shaped coalescent which starts in the partition into singleton blocks which all merge into a single block at some instant. The associated ultrametric on $\mathbb{N}$ induces the discrete topology. Here the uniform probability measures on $1, \ldots, n$ do not converge weakly as they converge vaguely to the zero measure.

A triple (X, r, μ) that consists of a complete and separable metric space (X, r) and a probability measure μ on the Borel sigma algebra on X is called a metric measure space. For a metric measure space (X, r, μ), one can consider the matrix $(r(x(i), x(j)))_{i,j \in \mathbb{N}}$ of the distances between μ-iid samples $(x(i))_{i \in \mathbb{N}}$. The distribution of $(r(x(i), x(j)))_{i,j \in \mathbb{N}}$ is called the distance matrix distribution of (X, r, μ). By the Gromov reconstruction theorem (see Theorem 4 of Vershik [51]), there exists a measure-preserving isometry between the supports of the measures of any two metric measure spaces that have the same distance matrix distribution, in which case we call them isomorphic.

We view a random semi-metric ρ on $\mathbb{N}$ as the random matrix $(\rho(i,j))_{i,j\in\mathbb{N}}$, and we call it exchangeable if $(\rho(i,j))_{i,j\in\mathbb{N}}$ is distributed as $(\rho(p(i),p(j)))_{i,j\in\mathbb{N}}$ for each (finite) permutation p of $\mathbb{N}$. Under an appropriate condition which we interpret as dust-freeness in Remark 3.13, Vershik [51, Theorem 5] associates with any typical realization of an exchangeable (and ergodic) random semi-metric on $\mathbb{N}$ a metric measure space whose distance matrix distribution is the distribution of this semi-metric. In the next subsection, we discuss an extension of such a representation to the case with dust.

1.2 The sampling representation

We give a representation for all exchangeable random semi-ultrametrics on $\mathbb{N}$ in terms of sampling from random marked metric measure spaces. Marked metric measure spaces are introduced in Depperschmidt, Greven, and Pfaffelhuber [12]. A $(\mathbb{R}_+)$-marked metric measure space is a triple (X, r, m) that consists of a complete and separable metric space (X, r) and a probability measure m on the Borel sigma algebra on the product space $X \times \mathbb{R}_+$. The marked distance matrix distribution of a marked metric measure space (X, r, m) is defined as the distribution of $((r(x(i), x(j)))_{i,j\in\mathbb{N}}, (v(i))_{i\in\mathbb{N}})$ where $(x(i), v(i))_{i\in\mathbb{N}}$ is an m-iid sequence in $X \times \mathbb{R}_+$. Marked metric measure spaces with the same marked distance matrix distribution are called isomorphic.

In the present article, we use marked metric measure spaces to obtain from a random variable $(\tilde{r}, \tilde{v})$ that has the marked distance matrix distribution of a marked metric measure space an exchangeable semi-metric $\tilde{\rho}$ on $\mathbb{N}$ by

$$\tilde{\rho}(i,j) = (\tilde{r}(i,j) + \tilde{v}(i) + \tilde{v}(j))\, \mathbb{1}\{i \neq j\}.$$

We call the distribution of $(\tilde{\rho}(i,j))_{i,j\in\mathbb{N}}$ the distance matrix distribution of the marked metric measure space. The basic result in this article (stated in Theorem 3.9 below) is that every exchangeable semi-ultrametric ρ on $\mathbb{N}$ can be represented as the outcome of a two-stage random experiment, where we have the isomorphy class χ of a random marked metric measure space in the first stage, and we sample $(\rho(i,j))_{i,j\in\mathbb{N}}$ from this marked metric measure space according to its distance matrix distribution in the second stage.

We construct χ realization-wise from the exchangeable semi-ultrametric ρ: the key idea is to decompose the tree that is associated with a realization of ρ into the external branches and the remaining subtree. Here we define that an external branch consists only of the leaf if that leaf corresponds to an integer that has ρ-distance zero to another integer. In the marked metric measure space, the marks encode the external branch lengths, and the metric space describes the remaining subtree. We call the semi-ultrametric dust-free if the external branches all have length zero a.s. In this case, the marked metric measure space can also be replaced by a metric measure space (as in Corollary 3.12). We prove Theorem 3.9 in Section 10. In Section 2, we formulate the decomposition at the external branches in terms of semi-ultrametrics.

The representation for exchangeable semi-ultrametrics from Theorem 3.9 can also be seen in the more general but less explicit contexts of the ergodic decomposition (Section 3.5) and the Aldous-Hoover-Kallenberg representation (see e.g. [32, Section 7]). In the representation result outlined above, the distance matrix distribution of the isomorphy class χ of the marked metric measure space is the ergodic component in whose support the realization ρ lies. The ergodic component is also characterized by χ itself, or in the dust-free case by the isomorphy class of a metric measure space. The finite analog of the aforementioned ergodic decomposition is that a (discrete) random tree whose leaves are labeled exchangeably can be obtained by first drawing the random unlabeled tree and then sampling the labels of the leaves uniformly without replacement.

We mention that Evans, Grübel, and Wakolbinger [22] also decompose real trees into the external branches and the remaining subtree to give a representation of the elements

of the Doob-Martin boundary of Rémy's algorithm in terms of sampling from a weighted real tree and an additional structure. In [22, Section 7], a sampling representation for exchangeable ultrametrics is considered (see Remark 10.9).

1.3 Evolving genealogies

In Section 4, we lay the foundation for our study of evolving genealogies by considering a general time-homogeneous Markov process with values in the space of semi-ultrametrics on $\mathbb{N}$; this process describes evolving leaf-labeled trees. Assuming that the state at each time is exchangeable, we map this process realization-wise to the processes of the ergodic components. We express these ergodic components as (isomorphy classes of) metric measure spaces and marked metric measure spaces, and as distance matrix distributions, respectively. Here we use the representation result for exchangeable semi-ultrametrics. This approach characterizes the processes of the ergodic components up to null events only at countably many time points, i.e. as versions, as we discuss in Remark 4.4. Using the criterion of Rogers and Pitman [46, Theorem 2], we deduce that these image processes are also Markovian, and we describe them by well-posed martingale problems. This is an example of Markov mapping in the sense of Kurtz [37], and Kurtz and Nappo [38].

In Sections 5 – 6, we study a concrete Markov process with values in the space of semi-ultrametrics, namely the process given by the evolving genealogical trees in a lookdown model with simultaneous multiple reproduction events. Lookdown models were introduced by Donnelly and Kurtz [15, 16] to represent measure-valued processes along with their genealogy, see also e.g. Etheridge and Kurtz [18] and Birkner et al. [8]. A lookdown model can be seen as a (possibly) infinite population model in which each individual at each time is assigned a level. The role of this level is model-inherent, namely to order the individuals such that the restriction of the model to the first finitely many levels is well-behaved (i.e. only finitely many reproduction events are visible in bounded time intervals) and that the modeled quantity (e.g. types, genealogical distances) is exchangeable. In [16] and in the present article, the level is the rank among the individuals at the respective time according to the time of the latest descendant. Although the levels in finite restrictions of the lookdown model differ from the labels in the Moran model, the processes of the unlabeled genealogical trees coincide which is used to study the length of the genealogical trees in Pfaffelhuber, Wakolbinger, and Weisshaupt [44] and Dahmer, Knobloch, and Wakolbinger [11].

In Section 7, we remove the labels from the evolving genealogical trees in the infinite lookdown model by applying the result from Section 4 to the process from Sections 5–6. We call the processes of the ergodic components tree-valued Fleming-Viot processes, regardless which one of the three state spaces we use. The tree-valued Fleming-Viot process with values in the space of isomorphy classes of metric measure spaces is introduced in the case with binary reproduction events (which is associated with the Kingman coalescent) by Greven, Pfaffelhuber, and Winter [26] as the solution of a well-posed martingale problem that is the limit in distribution of corresponding processes read off from finite Moran models. In [26, Remark 2.20], a construction of (a version of) this process from the lookdown model of Donnelly and Kurtz [15] is outlined. The aim in the present article regarding tree-valued Fleming-Viot process is the generalization to the case with dust. We remark that tree-valued Fleming-Viot processes with mutation and selection are studied in Depperschmidt, Greven, and Pfaffelhuber [13, 14] where the states are isomorphy classes of marked metric measure spaces and the marks encode allelic types. In the present article, the marks encode lengths of external branches. We consider only the neutral case, and we describe genealogies without using types.

In Section 8, we show continuity properties of the semigroups of tree-valued Fleming-Viot processes and that the domains of the martingale problems for them are cores. In Section 9, we show that tree-valued Fleming-Viot processes converge in distribution to equilibrium.

While we construct versions of tree-valued Fleming-Viot processes in the present article using the representation result, the full sample paths are constructed by techniques specific to the lookdown model in the companion article [29].

1.4 Additional related literature

Aldous [1] represents consistent families of finite trees that satisfy a "leaf-tight" property by random measures on ℓ_1 (and random subsets of ℓ_1). Kingman's coalescent is given as an example in [1]. The "leaf-tight" property corresponds to the absence of dust. A representation for exchangeable hierarchies in terms of sampling from random weighted real trees is given by Forman, Haulk, and Pitman [23]. There are many other representation results for exchangeable structures in the literature. For instance, by the Dovbysh-Sudakov theorem, see Austin [2] for a proof based on a representation for exchangeable random measures, jointly exchangeable arrays that are non-negative definite can be represented in terms of sampling from the space $L_2[0,1] \times \mathbb{R}_+$.

The genealogy in the lookdown model is further studied in Pfaffelhuber and Wakolbinger [43]. Kliem and Löhr [35] further study marked metric measure spaces. In their article, tree-valued Λ-Fleming-Viot processes in the dust-free case is also mentioned. Kliem and Winter [36] use marked metric measure spaces to describe trait-dependent branching processes. In the context of measure-valued spatial Λ-Fleming-Viot processes with dust, Véber and Wakolbinger [50] work with a skeleton structure. Functionals of coalescents like external branch lengths have also been studied, see for example [41]. Also the time evolution of such functionals has been studied for evolving coalescents, see for example [33, 10].

Bertoin and Le Gall [5, 6, 7] represent Ξ-coalescents in terms of sampling from flows of bridges from which they also construct measure-valued Fleming-Viot processes. They also consider mass coalescents. Mass coalescents (see e.g. Chapter 4.3 in Bertoin [4]) also describe genealogies without labeling individuals. In Remark 7.2, we discuss that the Fleming-Viot process with values in the space of distance matrix distributions can be constructed from the dual flow of bridges. We also mention the work of Labbé [39] where relations between the lookdown model and flows of bridges are studied.

2 Distance matrices and their decompositions

We write $\mathbb{N} = \{1, 2, 3, \ldots\}$. Let $\mathfrak{U}$ denote the space of semi-ultrametrics on $\mathbb{N}$ and let $\mathfrak{D}$ denote the space of semimetrics on $\mathbb{N}$. We view $\mathfrak{U}$ and $\mathfrak{D}$ as subspaces of $\mathbb{R}^{\mathbb{N}^2}$ in that we do not distinguish between a semi-metric ρ and the distance matrix $(\rho(i,j))_{i,j \in \mathbb{N}}$. We endow $\mathbb{R}^{\mathbb{N}^2}$ with a complete and separable metric that induces the product topology when $\mathbb{R}$ is equipped with the Euclidean topology. Using the map

$$\alpha : \mathbb{R}_+^{\mathbb{N}^2} \times \mathbb{R}_+^{\mathbb{N}} \to \mathbb{R}_+^{\mathbb{N}^2}, \quad (r,v) \mapsto ((v(i) + r(i,j) + v(j)) \mathbb{1}\{i \neq j\})_{i,j \in \mathbb{N}},$$

we define the space

$$\hat{\mathfrak{U}} = \left\{ (r,v) \in \mathfrak{D} \times \mathbb{R}_+^{\mathbb{N}} : \alpha(r,v) \in \mathfrak{U} \right\}$$

whose elements we call decomposed semi-ultrametrics or marked distance matrices. As above, we view $\mathfrak{D} \times \mathbb{R}_+^{\mathbb{N}}$ and $\hat{\mathfrak{U}}$ as subspaces of $\mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}}$ which we endow with a complete and separable metric that induces the product topology.

We define the function

$$\Upsilon : \mathfrak{U} \to \mathbb{R}_+^{\mathbb{N}}, \quad \rho \mapsto \left(\tfrac{1}{2} \inf_{j \in \mathbb{N} \setminus \{i\}} \rho(i,j) \right)_{i \in \mathbb{N}},$$

and we denote by β the function that maps a semi-ultrametric $\rho \in \mathfrak{U}$ to the decomposed semi-ultrametric $(r,v) \in \hat{\mathfrak{U}}$ that is given by $v = \Upsilon(\rho)$ and

$$r(i,j) = (\rho(i,j) - v(i) - v(j)) \mathbb{1}\{i \neq j\}$$

for $i,j \in \mathbb{N}$. The interpretation of these functions is given in Remark 2.2 below from which it follows that r is a tree-like semi-metric (i. e., r is 0-hyperbolic, see e. g. [21]). Alternatively, it can be easily checked that r satisfies the triangle inequality.

The function α retrieves the semi-ultrametric from a decomposed semi-ultrametric. For instance, $\alpha \circ \beta$ is the identity map on $\mathfrak{U}$.

Remark 2.1. Let us agree on the following notation. When we identify the elements of a semi-metric space (X, ρ) that have ρ-distance zero to obtain a metric space (X', ρ), we refer by each element $x \in X$ also to the associated element of X'. Furthermore, we define the metric completion of the semi-metric space (X, ρ) as the metric completion of (X', ρ).

Remark 2.2. Let $\rho \in \mathfrak{U}$, $(r,v) = \beta(\rho)$, and let (T,d) be the real tree associated with ρ as in Remark 1.1 with $X = \mathbb{N}$. Then $v(i) = \Upsilon(\rho)(i)$ can be interpreted as the length, and $(i, v(i))$ as the starting vertex of the external branch that ends in the leaf $(i,0)$ of T. Here we define that this external branch consists only of the leaf if there exists $k \in \mathbb{N} \setminus \{i\}$ with $\rho(i,k) = 0$. Furthermore, the map $\varphi(i) = (i, v(i))$ from $(\mathbb{N}, r)$ to (T,d) is distance-preserving.

In this sense, the map $\beta : \rho \mapsto (r,v)$ decomposes the coalescent tree that is given by ρ into the external branches with lengths v and the subtree spanned by their starting vertices whose mutual distances are given by r. More generally, any element of $\hat{\mathfrak{U}}$ can be seen as a decomposed coalescent tree.

We call a semi-ultrametric $\rho \in \mathfrak{U}$ dust-free if $\Upsilon(\rho) = 0$, that is, if all external branches in the associated tree have length zero so that there are no isolated leaves.

3 Sampling from marked metric measure spaces

3.1 Preliminaries

Recall the definitions of metric measure spaces, marked metric measure spaces, and their (marked) distance matrix distributions from Sections 1.1 and 1.2. Also recall that two metric measure spaces are said to be isomorphic if they have the same distance matrix distributions. We denote the set of isomorphy classes of metric measure spaces by $\mathbb{M}$ and we endow it with the Gromov-weak topology in which metric measure spaces converge if and only if their distance matrix distributions converge. Greven, Pfaffelhuber, and Winter [25] showed that $\mathbb{M}$ is then a Polish space.

Analogously, two marked metric measure spaces are said to be isomorphic if they have the same marked distance matrix distributions. We denote the set of isomorphy classes of marked metric measure spaces by $\hat{\mathbb{M}}$ and we endow it with the marked Gromov-weak topology in which marked metric measure spaces converge if and only if their marked distance matrix distributions converge weakly. This makes $\hat{\mathbb{M}}$ a Polish space, as shown by Depperschmidt, Greven, and Pfaffelhuber [12].

We denote the distance matrix distribution of the isomorphy class of a metric measure space $\chi \in \mathbb{M}$ by ν^χ. We denote the marked distance matrix distribution of $\chi' \in \hat{\mathbb{M}}$ by $\nu^{\chi'}$, so that $\alpha(\nu^{\chi'})$ is the distance matrix distribution of χ', in accordance with the definition

in Section 1.2. (We denote by $\varphi(\xi) = \xi \circ \varphi^{-1}$ the pushforward measure of a measure ξ on a measurable space E under a measurable function φ on E.)

Remark 3.1. We call a marked metric measure space (X, r, m) dust-free if the probability measure m is of the form $m = \mu \otimes \delta_0$ for a probability measure μ on the Borel sigma algebra on X. Then the distance matrix distribution $\alpha(\nu^{(X,r,\mu\otimes\delta_0)})$ equals the distance matrix distribution $\nu^{(X,r,\mu)}$ of the metric measure space (X, r, μ). We call (X, r, μ) the metric measure space associated with the dust-free marked metric measure space $(X, r, \mu \otimes \delta_0)$.

Let S_∞ denote the group of finite permutations on $\mathbb{N}$. We define the action of S_∞ on $\mathfrak{D}$ and $\mathfrak{D} \times \mathbb{R}_+^\mathbb{N}$, respectively, by $p(\rho) = (\rho(p(i), p(j)))_{i,j\in\mathbb{N}}$ and

$$p(r, v) = \big((r(p(i), p(j)))_{i,j\in\mathbb{N}}, (v(p(i)))_{i\in\mathbb{N}}\big)$$

for $p \in S_\infty$, $\rho \in \mathfrak{D}$, $(r, v) \in \mathfrak{D} \times \mathbb{R}_+^\mathbb{N}$. A random variable, for instance with values in $\mathfrak{D}$ or $\mathfrak{D} \times \mathbb{R}_+^\mathbb{N}$, is called exchangeable if its distribution is invariant under the action of the group S_∞.

Remark 3.2. Exchangeable random variables with values in $\mathfrak{D}$ or $\mathfrak{D} \times \mathbb{R}_+^\mathbb{N}$ can be seen as jointly exchangeable arrays, see e. g. [32, Section 7]. Also recall that the definition of exchangeability does not change when S_∞ is replaced with the group of all bijections from $\mathbb{N}$ to itself, as the finite restrictions determine the distribution of a random variable in $\mathfrak{D}$ or $\mathfrak{D} \times \mathbb{R}_+^\mathbb{N}$.

Remark 3.3. The coalescents associated by (1.1) with the exchangeable semi-ultra-metrics on $\mathbb{N}$ form a larger class of processes than the so-called exchangeable coalescents defined in e. g. Section 4.2.2 of Bertoin [4]. For example, the coalescent process associated with an exchangeable semi-ultrametric on $\mathbb{N}$ needs not be Markovian.

3.2 Tree-like marked metric measure spaces

We consider the space

$$\mathbb{U} = \{\chi \in \mathbb{M} : \nu^X(\mathfrak{U}) = 1\}$$

of ultrametric measure spaces which is a closed subspace of $\mathbb{M}$, as shown in [26, Lemma 2.3]. By the same argument, the space

$$\hat{\mathbb{U}} = \left\{\chi \in \hat{\mathbb{M}} : \alpha\left(\nu^X\right)(\mathfrak{U}) = 1\right\}.$$

is a closed subspace of $\hat{\mathbb{M}}$. It contains the marked metric measure spaces with ultrametric distance matrix distribution. Following e. g. [25, 26] and Remark 1.1, we call the elements of $\mathbb{U}$ trees. Also the elements of $\hat{\mathbb{U}}$ may be called trees (as in Remark 10.8 below).

Proposition 3.4 below states that a. e. realization of a $\hat{\mathfrak{U}}$-valued random variable with the marked distance matrix distribution of a marked metric measure space in $\hat{\mathbb{U}}$ is the decomposition of a semi-ultrametric by the map β from Section 2. As a consequence, the isomorphy class of a marked metric measure space in $\hat{\mathbb{U}}$ is determined already by its distance matrix distribution.

Proposition 3.4. Let (X, r', m) be a marked metric measure space with $\alpha(\nu^{(X,r',m)})(\mathfrak{U}) = 1$. Let (r, v) be a $\hat{\mathfrak{U}}$-valued random variable with distribution $\nu^{(X,r',m)}$. Then

$$(r, v) = \beta \circ \alpha(r, v) \quad \text{a. s.}$$

The proof is deferred to Section 10.1.

Remark 3.5. We call a semi-ultrametric $\rho \in \mathfrak{U}$ dust-free if $\Upsilon(\rho) = 0$. It can be seen as a consequence of Proposition 3.4 that (the isomorphy class of) a marked metric measure

space (X, r, m) in $\hat{\mathbb{U}}$ is dust-free (as defined in Remark 3.1) if and only if a random variable with distribution $\alpha(\nu^{(X,r,m)})$ is a. s. dust-free. In particular, a random variable with the distance matrix distribution of a metric measure space is a. s. dust-free.

3.3 Marked metric measure spaces from marked distance matrices

In this subsection, we define functions by which we construct a (marked) metric measure space from a (marked) distance matrix. An interpretation of these functions is given in Remark 3.8 below. In Remark 3.16, we state their role in the context of the ergodic decomposition.

First we define the function $\psi : \mathfrak{D} \to \mathbb{M}$ that maps $\rho \in \mathfrak{D}$ to the isomorphy class of the metric measure space (X, ρ, μ), given as follows: (X, ρ) is the metric completion of $(\mathbb{N}, \rho)$. The probability measure μ is defined as the weak limit of the probability measures $n^{-1} \sum_{i=1}^{n} \delta_i$ as n tends to infinity, if this weak limit exists. If the limit does not exist, we define m arbitrarily, let us set $\mu = \delta_1$. Furthermore, we denote by $\mathfrak{D}^*$ the subset of distance matrices $\rho \in \mathfrak{D}$ such that the weak limit in the definition above exists.

Analogously, we define the function $\hat{\psi} : \mathfrak{D} \times \mathbb{R}_+^{\mathbb{N}} \to \hat{\mathbb{M}}$ that maps (r, v) to the isomorphy class of the marked metric measure space (X, r, m), where (X, r) is the metric completion of the semi-metric space $(\mathbb{N}, r)$ and m is the weak limit of the probability measures $n^{-1} \sum_{i=1}^{n} \delta_{(i,v(i))}$ on $X \times \mathbb{R}_+$ if this weak limit exists, else we set $m = \delta_{(1,0)}$. We denote by $\hat{\mathfrak{D}}^*$ the subset of marked distance matrices $(r, v) \in \mathfrak{D} \times \mathbb{R}_+^{\mathbb{N}}$ such that the weak limit in the definition above exists.

We call μ and m in the definitions of ψ and $\hat{\psi}$ also sampling measures.

Remark 3.6. Let $(r, v) \in \mathfrak{D} \times \mathbb{R}_+^{\mathbb{N}}$. Then $(r, v) \in \hat{\mathfrak{D}}^*$ implies $r \in \mathfrak{D}^*$. For a representative (X, r, m) of $\hat{\psi}(r, v)$, the isomorphy class of $(X, r, m(\cdot \times \mathbb{R}_+))$ equals $\psi(r)$.

Proposition 3.7. *The functions ψ and $\hat{\psi}$ are measurable.*

The proof, in which we write ψ and $\hat{\psi}$ as limits of continuous functions, is deferred to Section 10.2.

Remark 3.8 (An interpretation of ψ and $\hat{\psi}$)**.** For $\rho \in \mathfrak{D}^* \cap \mathfrak{U}$, the probability measure in the ultrametric metric measure space $\psi(\rho)$ charges each ball with the asymptotic frequency of the corresponding block of the coalescent which is associated with ρ by (1.1).

Similarly, for $(r, v) \in \hat{\mathfrak{D}}^* \cap \hat{\mathfrak{U}}$, let (X, r, m) be the representative of $\hat{\psi}(r, v)$ from the definition of $\hat{\psi}$. We consider the completion $(\bar{T}, d)$ of the real tree (T, d) associated with (r, v) as in Remark 2.2, and the extension $\varphi : X \to \bar{T}$ of the isometry φ from Remark 2.2. Then the image measure $\mu := \varphi(m(\cdot \times \mathbb{R}_+))$ charges each region of $\bar{T}$ with the asymptotic frequency of the integers that label the leaves of T that are the endpoints of external branches that begin in that region.

3.4 The sampling representation

The basic result in this paper is stated in Theorem 3.9 below. Here we consider an exchangeable random semi-ultrametric ρ on $\mathbb{N}$, and we assert existence of a random variable χ with values in the space of isomorphy classes of marked metric measure spaces that has the following property: Let ρ' be a random variable whose conditional distribution given χ is the distance matrix distribution of χ. Then the random variables ρ and ρ' have the same (unconditional) distribution. (In the language of the theory of random measures, this means that the distribution of ρ is equal to the first moment measure $\mathbb{E}[\alpha(\nu^{\chi})]$. That is, $\mathbb{E}[\int \alpha(\nu^{\chi})(d\rho') \phi(\rho')] = \mathbb{E}[\phi(\rho)]$ for each bounded measurable ϕ.)

Theorem 3.9. *Let ρ be an exchangeable $\mathfrak{U}$-valued random variable. Let $\chi = \hat{\psi} \circ \beta(\rho)$. Let ρ' be a $\mathfrak{U}$-valued random variable whose conditional distribution given χ is $\alpha(\nu^{\chi})$.*

Then:

(i) $\beta(\rho) \in \hat{\mathfrak{D}}^*$ *a. s.*

(ii) ρ *and* ρ' *are equal in distribution.*

(iii) $\chi = \hat{\psi} \circ \beta(\rho')$ *a. s.*

Assertion (i) above states that for a typical realization of ρ and its decomposition $\beta(\rho)$, the sampling measure m in the definition of $\hat{\psi}(\beta(\rho))$ in Subsection 3.3 is the weak limit of the uniform probability measures therein. Assertion (iii) states that the realization of χ can typically be reconstructed from the realization of ρ'. We interpret the reconstruction map $\hat{\psi} \circ \beta$ in terms of the ergodic decomposition in Remark 3.16. We prove Theorem 3.9 in Section 10.4. We give two proofs of Theorem 3.9(i). In one of them, the de Finetti theorem yields the aforementioned sampling measure m as the directing measure of an exchangeable sequence.

Remark 3.10. In the context of Theorem 3.9, $(\rho, \alpha(\nu^\chi))$ and $(\rho', \alpha(\nu^\chi))$ are equal in distribution. Hence, $\alpha(\nu^\chi)$ is a regular conditional distribution of ρ given $\alpha(\nu^\chi)$.

We also note the following uniqueness property which is proved in Section 10.3.

Proposition 3.11. *Let* χ *and* χ' *be* $\hat{\mathbb{U}}$*-valued random variables. Let* ρ *be a* $\mathfrak{U}$*-valued random variable with conditional distribution* $\alpha(\nu^\chi)$ *given* χ*, and let* ρ' *be another* $\mathfrak{U}$*-valued random variable with conditional distribution* $\alpha(\nu^{\chi'})$ *given* χ'*. Then* ρ *and* ρ' *are equal in distribution if and only if* χ *and* χ' *are equal in distribution.*

(In terms of first-moment measures, Proposition 3.11 says that χ and χ' are equal in distribution if and only if $\mathbb{E}[\alpha(\nu^\chi)] = \mathbb{E}[\alpha(\nu^{\chi'})]$.)

The aim of the present paper is the treatment of the case with dust. In the dust-free case, we need not decompose the semi-metric ρ by the map β. Instead, we can work directly with the map ψ from Subsection 3.3. Theorem 3.9 then reduces to the setting of metric measure spaces as follows:

Corollary 3.12. *Let* ρ *be an exchangeable* $\mathfrak{U}$*-valued random variable that is a. s. dust-free. Let* $\chi = \psi(\rho)$*. Let* ρ' *be a* $\mathfrak{U}$*-valued random variable whose conditional distribution given* χ *is* ν^χ*. Then:*

(i) $\rho \in \mathfrak{D}^*$ *a. s.*

(ii) ρ *and* ρ' *are equal in distribution.*

(iii) $\chi = \psi(\rho')$ *a. s.*

Proof. This is immediate from Theorem 3.9 and Remarks 3.1, 3.5, and 3.6. □

Remark 3.13. The assertions of Corollary 3.12 are closely related to Vershik [51]: Condition (4) in [51, Theorem 5] is a necessary and sufficient condition for an exchangeable (and ergodic) random semi-metric to have the distance matrix distribution of a metric measure space. By Remark 3.5, the marked metric measure space χ in Theorem 3.9 is a. s. dust-free if and only if ρ is a. s. dust-free. Hence, for a semi-ultrametric ρ, condition (4) in [51] is equivalent to dust-freeness. In the dust-free case, the metric measure space associated with χ as in Remark 3.1 is the completion of a typical realization of the semi-metric, endowed with the probability measure given by the asymptotic block frequencies of the associated coalescent (as in Remark 3.8). This can also be deduced from [51, Equation (9)]. Assertion (iii) can be proved by Proposition 10.5 below which is related to [51] as stated in Remark 10.6.

3.5 Interpretation as ergodic decomposition

In this subsection, we interpret the representation from Theorem 3.9 as the ergodic decomposition of an exchangeable distribution on the semi-ultrametrics on $\mathbb{N}$.

We denote by $\mathcal{U}$ the space of exchangeable probability distributions on $\mathfrak{U}$, and we endow $\mathcal{U}$ with the Prohorov metric d_{P} which is complete and separable. We will also consider the subspace

$$\mathcal{U}^{\mathrm{erg}} = \left\{ \xi \in \mathcal{U} : \xi = \alpha\left(\nu^{(X,r,m)}\right) \text{ for some marked metric measure space } (X,r,m) \right\}$$

of distance matrix distributions of marked metric measure spaces. The sets $\mathcal{U}^{\mathrm{erg}}$ and $\hat{\mathbb{U}}$ are in one-to-one correspondence by Proposition 3.4. Hence, also the elements of $\mathcal{U}^{\mathrm{erg}}$ can be seen as trees.

We define the invariant sigma algebra $\mathcal{I}$ on $\mathfrak{U}$ as the sigma algebra that is generated by those Borel sets $B \subset \mathfrak{U}$ that satisfy $B = \{(\rho(p(i),p(j)))_{i,j \in \mathbb{N}} : \rho \in B\}$ for all finite permutations $p \in S_\infty$. A distribution ξ on $\mathfrak{U}$ is called ergodic (with respect to the action of the group S_∞ of finite permutations) if $\xi(I) \in \{0,1\}$ for all $I \in \mathcal{I}$.

Proposition 3.14. *The distance matrix distribution $\alpha(\nu^{(X,r,m)})$ of a marked metric measure space (X,r,m) is invariant and ergodic with respect to the action of the group of finite permutations.*

Proof. This is analogous to [51, Lemma 7]. For $I \in \mathcal{I}$, the Borel set $\tilde{I} \subset (X \times \mathbb{R}_+)^{\mathbb{N}}$ that given by

$$\tilde{I} = \left\{ (x(i),v(i))_{i \in \mathbb{N}} \in (X \times \mathbb{R}_+)^{\mathbb{N}} : ((v(i) + r(x(i),x(j)) + v(j)) \mathbb{1}\{i \neq j\})_{i,j \in \mathbb{N}} \in I \right\},$$

is invariant under finite permutations, that is,

$$\tilde{I} = \left\{ (x(p(i)),v(p(i)))_{i \in \mathbb{N}} : (x,v) \in \tilde{I} \right\} \quad \text{for all } p \in S_\infty.$$

From the ergodicity of an m-iid sequence $(x(i),v(i))_{i \in \mathbb{N}}$, we obtain

$$\alpha\left(\nu^{(X,r,m)}\right)(I) = \mathbb{P}\left((x,v) \in \tilde{I}\right) \in \{0,1\}. \qquad \square$$

Proposition 3.15. *The subset $\mathcal{U}^{\mathrm{erg}} \subset \mathcal{U}$ consists of the ergodic distributions.*

Proof. By Theorem 3.9(ii), each element of $\mathcal{U}$ is a mixture of elements of $\mathcal{U}^{\mathrm{erg}}$. The assertion follows by Proposition 3.14 and as the ergodic distributions in $\mathcal{U}$ are extreme in the convex set $\mathcal{U}$ (see e. g. [32, Lemma A1.2]). $\qquad \square$

Remark 3.16. Theorem 3.9 decomposes the distribution of the exchangeable $\mathfrak{U}$-valued random variable ρ' into ergodic components in the sense of e. g. Theorem A1.4 in Kallenberg [32]. The function

$$\zeta : \mathfrak{U} \to \mathcal{U}^{\mathrm{erg}}, \quad \tilde{\rho} \mapsto \alpha\left(\nu^{\hat{\psi} \circ \beta(\tilde{\rho})}\right)$$

is a decomposition map in the sense of Varadarajan [49, Section 4] so that typically, $\zeta(\rho')$ is the ergodic component in whose support the realization ρ' lies. Note that this ergodic component is characterized by the isomorphy class $\chi = \hat{\psi} \circ \beta(\rho')$ of a marked metric measure space, and in the dust-free case also by the isomorphy class $\psi(\rho')$ of a metric measure space. Some further references on the ergodic decomposition are given e. g. in [32, p. 475].

By the following proposition, $(\mathcal{U}^{\mathrm{erg}}, d_{\mathrm{P}})$ is Polish which will be applied in [28].

Proposition 3.17. *The subspace $\mathcal{U}^{\mathrm{erg}}$ is closed in $(\mathcal{U}, d_{\mathrm{P}})$.*

Proof. Let $(\rho^n, n \in \mathbb{N})$ be a sequence of $\mathfrak{U}$-valued random variables that converges in distribution to some $\mathfrak{U}$-valued random variable ρ. Assume that for each $n \in \mathbb{N}$, the distribution of ρ^n lies in $\mathcal{U}^{\mathrm{erg}}$. Then ρ^n has ergodic distribution by Proposition 3.14. Lemma 7.35 of [32] says that ρ^n is dissociated, which means that for any disjoint $I_1, \ldots, I_k \subset \mathbb{N}$, the restrictions $(\rho^n(i,j))_{i,j \in I_1}, \ldots, (\rho^n(i,j))_{i,j \in I_k}$ are independent. As this property is preserved under the limit in distribution, it also holds for ρ, and another application of Lemma 7.35 of [32] and yields that ρ has ergodic distribution. The assertion follows by Proposition 3.15. $\qquad\square$

4 Application to tree-valued processes

Using the function $\hat{\psi}$ from Section 3.3, we map a Markov process $(\rho_t, t \in \mathbb{R}_+)$ whose states are exchangeable $\mathfrak{U}$-valued random variables to a process with values in the space of isomorphy classes of marked metric measure spaces. At each time, the state of the image process is the marked metric measure space from the representation (Theorem 3.9) of the state of the $\mathfrak{U}$-valued process. We also consider the process of the distance matrix distributions of these marked metric measure spaces. In the dust-free case, we can also work with isomorphy classes of metric measure spaces and the map ψ as in Corollary 3.12.

In the proof of Theorem 4.1 below, we use the criterion of Rogers and Pitman [46, Theorem 2] to show that also the image processes are Markovian. A martingale problem for the $\mathfrak{U}$-valued process $(\rho_t, t \in \mathbb{R}_+)$ or the $\hat{\mathfrak{U}}$-valued process $(\beta(\rho_t), t \in \mathbb{R}_+)$ yields a martingale problem for the respective image process.

The so-called polynomials and marked polynomials, introduced in [25, 12] have been used as domains of martingale problems in e.g. [26, 13, 14]. We recall them here, adapting the definition to our present use of the marks. The uniform continuity of the derivative in the definitions of $\mathcal{C}_n$ and $\hat{\mathcal{C}}_n$ below will turn out useful in [28]. For $n \in \mathbb{N}$, we write $[n] = \{1, \ldots, n\}$ for $n \in \mathbb{N}$, and we denote by γ_n the restriction from $\mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}}$ to $\mathbb{R}^{n^2} \times \mathbb{R}^n$, $\gamma_n(r,v) = ((r(i,j))_{i,j \in [n]}, (v(i))_{i \in [n]})$. We denote also by γ_n the restriction from $\mathbb{R}^{\mathbb{N}^2}$ to $\mathbb{R}^{n^2}$, $\gamma_n(\rho) = (\rho(i,j))_{i,j \in [n]}$. Let $\mathcal{C}_n$ denote the set of bounded differentiable functions $\mathbb{R}^{n^2} \to \mathbb{R}$ with bounded uniformly continuous derivative. For $\phi \in \mathcal{C}_n$, we denote also by ϕ the function $\phi \circ \gamma_n : \mathbb{R}^{\mathbb{N}^2} \to \mathbb{R}$, and we call the function $\mathbb{U} \to \mathbb{R}$, $\chi \mapsto \nu^\chi \phi$ the polynomial associated with ϕ. (Here and at other places, we use the notation $\xi f = \int \xi(dx) f(x)$ for a measure ξ and an integrable function f, and we view measures also as functionals on spaces of integrable functions.) Similarly, we denote by $\hat{\mathcal{C}}_n$ the set of bounded differentiable functions $\mathbb{R}^{n^2} \times \mathbb{R}^n \to \mathbb{R}$ with uniformly continuous derivative. For $\phi \in \hat{\mathcal{C}}_n$, we denote also by ϕ the function $\phi \circ \gamma_n : \mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$, and we call the function $\hat{\mathbb{U}} \to \mathbb{R}$, $\chi \mapsto \nu^\chi \phi$ the marked polynomial associated with ϕ. (Usually, the argument (r,v) of a function $\phi \in \hat{\mathcal{C}}$ will be a marked distance matrix.) We write $\mathcal{C} = \bigcup_{n \in \mathbb{N}} \mathcal{C}_n$ and $\hat{\mathcal{C}} = \bigcup_n \hat{\mathcal{C}}_n$. We denote the set of polynomials by

$$\Pi = \left\{ \mathbb{U} \to \mathbb{R}, \chi \mapsto \nu^\chi \phi : \phi \in \mathcal{C} \right\},$$

the set of marked polynomials by

$$\hat{\Pi} = \left\{ \hat{\mathbb{U}} \to \mathbb{R}, \chi \mapsto \nu^\chi \phi : \phi \in \hat{\mathcal{C}} \right\},$$

and we define the set of test functions

$$\mathscr{C} = \left\{ \mathcal{U}^{\mathrm{erg}} \to \mathbb{R}, \xi \mapsto \xi \phi : \phi \in \mathcal{C} \right\}.$$

For a metric space E, let $M_{\mathrm{b}}(E)$ denote the set of bounded measurable functions $E \to \mathbb{R}$. For a subset $\mathscr{D} \subset M_{\mathrm{b}}(E)$ and an operator $G : \mathscr{D} \to M_{\mathrm{b}}(E)$, we mean by a solution of the martingale problem $(G, \mathscr{D})$ a progressive E-valued process $(X_t, t \in \mathbb{R}_+)$ such that for every $f \in \mathscr{D}$, the process

$$f(X_t) - \int_0^t Gf(X_s)ds$$

is a martingale with respect to the filtration induced by $(X_t, t \in \mathbb{R}_+)$, cf. Ethier and Kurtz [19, p. 173].

Theorem 4.1. *Let $(\rho_t, t \in \mathbb{R}_+)$ be a $\mathfrak{U}$-valued time-homogenous Markov process. Assume that for each $t \in \mathbb{R}_+$, the random variable ρ_t is exchangeable. Let $A : \mathcal{C} \to M_{\mathrm{b}}(\mathbb{R}^{\mathbb{N}^2})$ and $\hat{A} : \hat{\mathcal{C}} \to M_{\mathrm{b}}(\mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}})$ be operators. Define the $\mathbb{U}$-valued process $(\chi_t, t \in \mathbb{R}_+) := (\psi(\rho_t), t \in \mathbb{R}_+)$, the $\hat{\mathbb{U}}$-valued process $(\hat{\chi}_t, t \in \mathbb{R}_+) := (\hat{\psi}(\beta(\rho_t)), t \in \mathbb{R}_+)$, and the $\mathcal{U}^{\mathrm{erg}}$-valued process $(\xi_t, t \in \mathbb{R}_+) := (\alpha(\nu^{\hat{\chi}_t}), t \in \mathbb{R}_+)$. Then the following two assertions hold:*

(i) The process $(\hat{\chi}_t, t \in \mathbb{R}_+)$ is Markovian. If the $\hat{\mathfrak{U}}$-valued process $(\beta(\rho_t), t \in \mathbb{R}_+)$ solves the martingale problem $(\hat{A}, \hat{\mathcal{C}})$, then $(\hat{\chi}_t, t \in \mathbb{R}_+)$ solves the martingale problem $(\hat{B}, \hat{\Pi})$, given by

$$\hat{B}\Phi(\chi) = \nu^\chi\left(\hat{A}\phi\right)$$

for all $\phi \in \hat{\mathcal{C}}$ with associated polynomial Φ, and all $\chi \in \hat{\mathbb{U}}$.

(ii) The process $(\xi_t, t \in \mathbb{R}_+)$ is Markovian. If $(\rho_t, t \in \mathbb{R}_+)$ solves the martingale problem $(A, \mathcal{C})$, then $(\xi_t, t \in \mathbb{R}_+)$ solves the martingale problem $(C, \mathscr{C})$, given by

$$C\Psi(\xi) = \xi(A\phi)$$

for all $\xi \in \mathcal{U}^{\mathrm{erg}}$ and $\phi \in \mathcal{C}$, and the function $\Psi \in \mathscr{C}$, $\xi' \mapsto \xi'\Psi$.

Assertion (iii) below holds under the additional assumption that ρ_t is a. s. dust-free for each $t \in \mathbb{R}_+$.

(iii) The process $(\chi_t, t \in \mathbb{R}_+)$ is Markovian. If $(\rho_t, t \in \mathbb{R}_+)$ solves the martingale problem $(A, \mathcal{C})$, then $(\chi_t, t \in \mathbb{R}_+)$ solves the martingale problem (B, Π), given by

$$B\Phi(\chi) = \nu^\chi(A\phi)$$

for all $\phi \in \mathcal{C}$ with associated polynomial Φ, and all $\chi \in \mathbb{U}$.

The proof of Theorem 4.1 can be found in Section 10.5.

Remark 4.2. The process $(\beta(\rho_t), t \in \mathbb{R}_+)$ in Theorem 4.1 is Markov. This follows as $(\rho_t, t \in \mathbb{R}_+)$ is Markov by assumption and as ρ_t is determined by $\beta(\rho_t)$ via $\rho_t = \alpha(\beta(\rho_t))$ so that

$$\mathbb{E}[f(\beta(\rho_u))|(\beta(\rho_s))_{s \le t}] = \mathbb{E}[f(\beta(\rho_u))|(\rho_s)_{s \le t}] = \mathbb{E}[f(\beta(\rho_u))|\rho_t] = \mathbb{E}[f(\beta(\rho_u))|\beta(\rho_t))] \quad \text{a. s.}$$

for all $s \le t \le u$ and bounded measurable $f : \hat{\mathfrak{U}} \to \mathbb{R}$. This is an example for Dynkin's criterion [17, Theorem 10.13] for a function of a Markov process to be Markov.

Remark 4.3. In Theorem 4.1, if ρ_t is dust-free for some $t \in \mathbb{R}_+$, then $\hat{\chi}_t$ is (by Theorem 3.9 and Remark 3.5 the isomorphy class of a) dust-free marked metric measure space, χ_t is the (isomorphy class of the) metric measure space associated (as in Remark 3.1) with (any representative of) $\hat{\chi}_t$, and we have $\xi_t = \nu^{\chi_t}$. The process $(\chi_t, t \in \mathbb{R}_+)$ is relevant only in the dust-free case: If ρ_t is not dust-free, then $\psi(\rho_t)$ is just the arbitrary element of $\mathbb{M}$ from the definition of ψ in Section 3.3.

Remark 4.4. In Theorem 4.1, we characterize only versions of the processes $(\hat{\chi}_t, t \in \mathbb{R}_+)$, $(\chi_t, t \in \mathbb{R}_+)$, and $(\xi_t, t \in \mathbb{R}_+)$. That is, we do not make assertions on the full sample paths but only on the states at countably many times.

From Theorem 3.9, we obtain $\beta(\rho_t) \in \hat{\mathfrak{D}}^*$ (and in the dust-free case also $\rho_t \in \mathfrak{D}^*$ by Corollary 3.12) only for a fixed time t (or countably many t) on an event of probability 1. This means that the uniform probability measures on the starting vertices of the external branches that end in the first n leaves of the tree associated with the semi-ultrametric ρ_t are shown to converge only at countably many times t on an event of probability 1. For $\beta(\rho_t) \in \hat{\mathfrak{D}}^*$, a realization $\hat{\chi}_t = \hat{\psi}(\beta(\rho_t))$ can be considered as an ergodic component. At the other times t, we do not exclude that $\hat{\chi}_t = \hat{\psi}(\beta(\rho_t))$ is just the arbitrary element of $\hat{\mathbb{M}}$ with probability measure $\delta_{(1,0)}$ in the definition of $\hat{\psi}$ in Section 3.3.

Theorem 4.1 yields in particular the semigroups of the processes $(\hat{\chi}_t, t \in \mathbb{R}_+)$, $(\chi_t, t \in \mathbb{R}_+)$, and $(\xi_t, t \in \mathbb{R}_+)$. Also the martingale problems in Theorem 4.1 characterize only versions of these processes.

For the particular example of the process $(\rho_t, t \in \mathbb{R}_+)$ in Sections 5 – 9, it is shown in [29] that $\beta(\rho_t) \in \hat{\mathfrak{D}}^*$ (and $\rho_t \in \mathfrak{D}^*$ in the dust-free case) also holds simultaneously for all $t \in \mathbb{R}_+$ on an event of probability 1 (see Theorems 3.1(i) and 3.10(i), and Remarks 4.4 and 4.13 in [29]). This allows to construct the full sample paths (Section 4 in [29]). These results are obtained in [29] by techniques specific to the lookdown model.

Remark 4.5. Theorem 4.1 is an example for Markov mapping. To show that the image processes $(\hat{\psi}(\beta(\rho_t)), t \in \mathbb{R}_+)$, $(\xi_t, t \in \mathbb{R}_+)$, and $(\psi(\rho_t), t \in \mathbb{R}_+)$ are Markovian, we use the simple criterion of Rogers and Pitman [46, Theorem 2] as this criterion is formulated in terms of the abstract semigroups of the processes, which fits to our assumption that $(\rho_t, t \in \mathbb{R}_+)$ is a general time-homogenous Markov process whose states ρ_t are exchangeable.

A criterion for the Markov property of the image processes in terms of martingale problems is given in Corollary 3.5 of Kurtz [37] which requires more assumptions, including uniqueness for the martingale problem for $(\rho_t, t \in \mathbb{R}_+)$ and existence of solutions of the martingale problems for the image processes. Corollary 3.5 of [37] would also yield uniqueness for the martingale problems for the image processes.

In the present paper, we use martingale problems only to provide additional characterizations of the processes under consideration. In Proposition 7.1, we show uniqueness for the martingale problems for the image processes directly by duality for the concrete examples from Section 7.

Remark 4.6. In particular in Sections 8 – 9, 11.3 and in [28], we need convergence determining (or at least separating) sets of test functions. As in [40, 25, 12], the sets Π and $\hat{\Pi}$ are convergence determining in $\mathbb{U}$ and $\hat{\mathbb{U}}$, respectively. The argument from [40, Corollary 2.8] also applies for $\mathscr{C}$: The algebra $\mathcal{C}$ generates the product topology on $\mathbb{R}^{\mathbb{N}^2}$. By a theorem due to Le Cam, see e.g. [40, Theorem 2.7] and the references therein, it follows that $\mathcal{C}$ is convergence determining in $\mathfrak{U}$. Hence, $\mathscr{C}$ generates the weak topology on $\mathcal{U}^{\mathrm{erg}}$. As $\hat{\Pi}$ is an algebra (see [25, 12]) and by definition of $\mathcal{U}^{\mathrm{erg}}$, also $\mathscr{C}$ is an algebra. Again by [40, Theorem 2.7], it follows that $\mathscr{C}$ is convergence determining in $\mathcal{U}^{\mathrm{erg}}$.

Remark 4.7. The set of polynomials $\Pi' = \{\hat{\mathbb{U}} \to \mathbb{R}, \chi \mapsto \alpha(\nu^\chi)\phi : \phi \in \mathcal{C}\}$ is separating on $\hat{\mathbb{U}}$. This follows from Propositions 3.4 and 10.5 as in the proof of Proposition 3.11. Nevertheless, we work with the space $\hat{\Pi}$ of test functions on $\hat{\mathbb{M}}$ as Π' is not convergence determining, a counterexample can be constructed from [25, Example 2.12(ii)].

5 Genealogy in the lookdown model

In this section, we define a Markov process $(\rho_t, t \in \mathbb{R}_+)$ to which we will later apply Theorem 4.1. In Subsection 5.1, we read off a realization of such a process from a

population model that is driven by a deterministic point measure η. In Subsection 5.2, we let η be a Poisson random measure, and we study further properties of $(\rho_t, t \in \mathbb{R}_+)$ in Subsection 5.3. We remark that for the lookdown model of Donnelly and Kurtz [15], the process of the evolving genealogical distances and its martingale problem are considered in Remark 2.20 of Greven, Pfaffelhuber, and Winter [26].

5.1 The deterministic construction

We denote by $\mathcal{P}$ the set of partitions of $\mathbb{N}$. We endow $\mathcal{P}$ with the topology in which a sequence of partitions converges if and only if the sequences of their finite restrictions converge. For $n \in \mathbb{N}$, we denote by $\mathcal{P}_n$ the set of partitions of $[n] = \{1, \ldots, n\}$. We denote the restriction map from $\mathcal{P}$ to $\mathcal{P}_n$ by γ_n, that is, $\gamma_n(\pi) = \{B \cap [n] : B \in \pi\} \setminus \{\emptyset\}$. Recall that other restriction maps, e.g. from $\mathbb{R}^{\mathbb{N}^2} \to \mathbb{R}^{n^2}$ are also denoted by γ_n. Moreover, we denote by $\mathbf{0}_n = \{\{1\}, \ldots, \{n\}\}$ the partition in $\mathcal{P}_n$ that consists of singletons only, and by $\mathcal{P}^n = \{\pi \in \mathcal{P} : \gamma_n(\pi) \neq \mathbf{0}_n\}$ the set of partitions of $\mathbb{N}$ in which the first n integers are not all in different blocks. Furthermore, for $\pi \in \mathcal{P}$, we denote by $B_1(\pi), B_2(\pi), \ldots$ the enumeration of the blocks of π with $\min B_1(\pi) < \min B_2(\pi) < \ldots$. For $i \in \mathbb{N}$, we denote by $\pi(i)$ the integer j that satisfies $i \in B_j(\pi)$.

We use a lookdown model as the population model. In this model, there are countably infinitely many levels which are labeled by $\mathbb{N}$, and each level is occupied by one particle at each time $t \in \mathbb{R}_+$. The particles undergo reproduction events which are encoded by a simple point measure η on $(0, \infty) \times \mathcal{P}$. A simple point measure is a purely atomic measure whose atoms all have mass 1. Let us impose a further assumption on η, namely

$$\eta((0, t] \times \mathcal{P}^n) < \infty \quad \text{for all } t \in (0, \infty) \text{ and } n \in \mathbb{N}. \tag{5.1}$$

The interpretation of a point (t, π) of η is that the following reproduction event occurs: At time $t-$, the particles on the levels $i \in \mathbb{N}$ with $i > \#\pi$ are removed. At time t, for each $i \in [\#\pi]$, the particle that was on level i at time $t-$ assumes level $\min B_i(\pi)$ and has offspring on all other levels in $B_i(\pi)$. Thus, the level of a particle is non-decreasing as time evolves. Condition (5.1) means that for each $n \in \mathbb{N}$, only finitely many particles jump away from the first n levels in bounded time intervals.

For all $0 \leq s \leq t$, each particle at time t has an ancestor at time s. We denote by $A_s(t, i)$ the level of the ancestor at time s of the particle on level i at time t such that the maps $s \mapsto A_s(t, i)$ and $t \mapsto A_s(t, i)$ are càdlàg. Then $A_s(t, i)$ is well-defined as $s \mapsto A_{t-s}(t, i)$ is non-increasing.

Remark 5.1. We will use that the trajectories of the particles are non-crossing in the following sense: For any times $s \leq t$ and particles x, y on levels $i_x \leq i_y$ at time $s \in \mathbb{R}_+$, particle x is still alive if particle y is still alive, in which case the particles x and y occupy levels $j_x \leq j_y$. In particular, if infinitely many particles at time s survive until time t, then all particles at time s survive until time t.

We are interested in the process of the genealogical distances between the particles that live at the respective times. Let $\rho_0 \in \mathbb{R}^{\mathbb{N}^2}$. (We can assume $\rho_0 \in \mathfrak{U}$ here, but differentiability will be more elementary in the larger space, as a matter of taste.) We interpret $\rho_0(i, j)$ as the genealogical distance between the particles on levels i and j at time 0. We define the genealogical distance between the particles on levels i and j at time t by

$$\rho_t(i, j) = \begin{cases} 2t - 2\sup\{s \in [0, t] : A_s(t, i) = A_s(t, j)\} & \text{if } A_0(t, i) = A_0(t, j) \\ 2t + \rho_0(A_0(t, i), A_0(t, j)) & \text{else.} \end{cases}$$

In words, the genealogical distance between two particles at a fixed time is twice the time back to their most recent common ancestor, if such an ancestor exists, else it is given by the genealogical distance between the ancestors at time zero.

Remark 5.2. If $\rho_0 \in \mathfrak{U}$, then $\rho_t \in \mathfrak{U}$ for each $t \in \mathbb{R}_+$. Indeed, a semi-metric ρ on $\mathbb{N}$ is a semi-ultrametric if and only if for each $s \in \mathbb{R}_+$, an equivalence relation $\sim$ on $\mathbb{N}$ is given by $i \sim j :\Leftrightarrow \rho(i, j) \leq s$. If this property holds for ρ_0, then the definition of ρ_t readily yields that it also holds for ρ_t.

We also describe the process $(\rho_t, t \in \mathbb{R}_+)$ in a more formal way which will be useful for the description by martingale problems in Section 5.2. With each partition $\pi \in \mathcal{P}_n$ we associate a transformation $\mathbb{R}^{n^2} \to \mathbb{R}^{n^2}$, which we also denote by π, by

$$\pi(\rho) = (\rho(\pi(i), \pi(j)))_{i,j \in [n]}. \tag{5.2}$$

Here $\pi(i)$ denotes the integer k such that i is in the k-th block, when blocks are ordered according to their minimal elements. Note that for each reproduction event encoded by a point $(s, \pi) \in \eta$, the corresponding jump of the process $(\rho_t, t \in \mathbb{R}_+)$ can be described by

$$\gamma_n(\pi)(\gamma_n(\rho_{s-})) = \gamma_n(\rho_s). \tag{5.3}$$

In particular, $\gamma_n(\pi) = \mathbf{0}_n$ if $\pi \in \mathcal{P} \setminus \mathcal{P}^n$, and $\mathbf{0}_n$ acts as the identity on $\mathbb{R}^{n^2}$. By assumption (5.1), there are only finitely many reproduction events in bounded time intervals that result in a jump of the process $(\gamma_n(\rho_t(i, j)), t \in \mathbb{R}_+)$. Between such jumps, the genealogical distances grow linearly with slope 2, that is, $\rho_t(i, j) + 2s = \rho_{t+s}(i, j)$ for distinct $i, j \in [n]$ and $t, s \in \mathbb{R}_+$ with $\eta((t, t + s] \times \mathcal{P}^n) = 0$.

Remark 5.3. Schweinsberg [48] constructs the Ξ-coalescent analogously from a point measure. The population model described in this section can be seen as the population model that underlies the dual flow of partitions in Foucart [24]. A lookdown model with a reproduction mechanism that is different in the case with simultaneous multiple reproduction events is studied by Birkner et al. [8]. In this model, a partition $\pi \in \mathcal{P}$ encodes the following reproduction event: Let $i_1 < i_2 < \ldots$ be the increasing enumeration of the integers that either form singletons or are non-minimal elements of blocks of π. For each $j \in \mathbb{N}$, the particle on level i_j moves to the level given by the j-th lowest singleton of π if π has at least j singletons, else the particle is removed. For each non-singleton block $B \in \pi$, the particle on level $\min B$ remains on its level and has one offspring on each level in $B \setminus \{\min B\}$. Here the trajectories of the particles may cross: Consider a partition $\pi \in \mathcal{P}$ such that 1 and 2 are in the same block, 4 forms a singleton, and 3 is the minimal element of a non-singleton block. If the reproduction event encoded by π occurs at time $t \in (0, \infty)$, then there exists $s \in (0, t)$ such that the particle on level 3 at time s is on level 3 also at time t, and the particle on level 2 at time s jumps to level 4 at time t. Such a crossing cannot occur in our population model by Remark 5.1.

5.2 The Ξ-lookdown model

The population model from the Subsection 5.1 will now be driven by a Poisson random measure η on $(0, \infty) \times \mathcal{P}$ as in Schweinsberg [48], Bertoin [4], and Foucart [24].

To define this Poisson random measure, we briefly recall Kingman's correspondence. For a full account, see e. g. [4, Section 2.3.2]. Kingman's correspondence is a one-to-one correspondence between the distributions of the exchangeable random partitions of $\mathbb{N}$ and the probability measures on the simplex

$$\Delta = \{x = (x_1, x_2, \ldots) : x_1 \geq x_2 \geq \ldots \geq 0, |x|_1 \leq 1\},$$

where $|x|_1 = \sum_{i \in \mathbb{N}} x_i$. Every $x \in \Delta$ can be interpreted as a partition of $[0, 1]$ into subintervals of lengths $x_1, x_2, \ldots$, and possibly another interval of length $1 - |x|_1$ which may be called the dust interval. Let $U_1, U_2, \ldots$ be iid uniform random variables with values in $[0, 1]$. The paintbox partition associated with x is the exchangeable random

partition of $\mathbb{N}$ where two different integers i and j are in the same block if and only if U_i and U_j fall into a common subinterval that is not the dust interval. This construction defines a probability kernel κ from Δ to $\mathcal{P}$. Conversely, every exchangeable random partition π in $\mathcal{P}$ has distribution $\int_\Delta \xi(dx)\kappa(x,\cdot)$ for some distribution ξ on Δ. Here x is the random vector in Δ of the asymptotic frequencies of the blocks of π.

Let Ξ be a finite measure on Δ. We decompose

$$\Xi = \Xi_0 + \Xi\{0\}\delta_0. \tag{5.4}$$

For $i, j \in \mathbb{N}$ with $i \neq j$, we denote by $K_{i,j}$ the partition in $\mathcal{P}$ that contains the block $\{i,j\}$ and apart from that only singleton blocks. We define a σ-finite measure H_Ξ on $\mathcal{P}$ by

$$H_\Xi(d\pi) = \int_\Delta \kappa(x,d\pi)\,|x|_2^{-2}\,\Xi_0(dx) + \Xi\{0\}\sum_{1\leq i<j}\delta_{K_{i,j}}(d\pi),$$

where $|x|_2 = \left(\sum_{i\in\mathbb{N}} x_i^2\right)^{1/2}$.

Let η be a Poisson random measure on $(0,\infty) \times \mathcal{P}$ with intensity $dt\, H_\Xi(d\pi)$. Note that $\kappa(x,\mathcal{P}^n) \leq \binom{n}{2}|x|_2^2$ for all $x \in \Delta$ and $n \geq 2$. This follows as in the paintbox partition associated with x, the probability that two fixed integers belong to the same block is $|x|_2^2$. The random point measure η thus satisfies condition (5.1) a. s. as

$$\mathbb{E}[\eta((0,t]\times\mathcal{P}^n)] = \int \kappa(x,\mathcal{P}^n)|x|_2^{-2}\Xi_0(dx) + \Xi\{0\}\binom{n}{2} < \infty \tag{5.5}$$

for all $t \in \mathbb{R}_+$ and $n \in \mathbb{N}$. Hence, we can and will define the population model from Subsection 5.1 from almost every realization of η and every $\rho_0 \in \mathbb{R}^{\mathbb{N}^2}$. We also let ρ_0 be a $\mathbb{R}^{\mathbb{N}^2}$-valued random variable that is independent of η. We define the $\mathbb{R}^{\mathbb{N}^2}$-valued process $(\rho_t, t \in \mathbb{R}_+)$ realization-wise from the Poisson random measure η and the random initial state ρ_0 as in the preceding subsection.

Proposition 5.4. *The process $(\rho_t, t \in \mathbb{R}_+)$ is Markov.*

Proof. The description around equation (5.3) implies that for $0 \leq t < t'$ and each $n \in \mathbb{N}$, the conditional expectation of $\gamma_n(\rho_{t'})$ given $(\rho_s, s \leq t)$ is measurable with respect to ρ_t and the restriction of η to $(t,t']\times\mathcal{P}$. The assertion follows as $n \in \mathbb{N}$ was arbitrary and as the restrictions of a Poisson random measure to disjoint subsets are independent. $\qquad\square$

For each $n \in \mathbb{N}$ and $\pi \in \mathcal{P}_n \setminus \{\mathbf{0}_n\}$, the rate at which reproduction events encoded by partitions in $\gamma_n^{-1}(\pi) = \{\pi' \in \mathcal{P} : \gamma_n(\pi') = \pi\}$ occur in the lookdown model is given by $\lambda_\pi = H_\Xi(\gamma_n^{-1}(\pi))$. The rates λ_π are calculated explicitly in (6.4) and (6.3) in Section 6.2.

Remark 5.5. The quantity λ_π is the coagulation rate q_π in Section 4.2.1 of Bertoin [4]. It is related to the quantity $\lambda_{n;k_1,\ldots,k_r;s}$ from Schweinsberg [48] by $\lambda_\pi = \lambda_{n;k_1,\ldots,k_r;s}$, where $k_1,\ldots,k_r$ denote the sizes of the non-singleton blocks of π, and $s = n - k_1 - \ldots - k_r$. This can be seen by a comparison of equations (6.4) and (6.3) with equation (11) in [48]. In particular, equation (18) in [48] implies that η satisfies a. s. condition (5.1).

In the next proposition, we state a martingale problem for the process $(\rho_t, t \in \mathbb{R}_+)$. Recall the set $\mathcal{C}$ from Section 4. For $\phi \in \mathcal{C}$ and $\rho \in \mathbb{R}^{\mathbb{N}^2}$, we write

$$\langle\nabla\phi, \underline{\underline{2}}\rangle(\rho) = 2\sum_{\substack{i,j\in\mathbb{N}\\i\neq j}}\frac{\partial}{\partial\rho(i,j)}\phi(\rho). \tag{5.6}$$

Proposition 5.6. *Define an operator $A = A_{\text{grow}} + A_{\text{repr}}$ with domain $\mathcal{C}$ by*

$$A_{\text{grow}}\phi(\rho) = \langle\nabla\phi, \underline{\underline{2}}\rangle(\rho)$$

and

$$A_{\mathrm{repr}}\phi(\rho) = \sum_{\pi \in \mathcal{P}_n \setminus \{\mathbf{0}_n\}} \lambda_\pi (\phi(\pi(\gamma_n(\rho))) - \phi(\rho))$$

for $n \in \mathbb{N}$, $\phi \in \mathcal{C}_n$, and $\rho \in \mathbb{R}^{\mathbb{N}^2}$. Then the stochastic process $(\rho_t, t \in \mathbb{R}_+)$ solves the martingale problem $(A, \mathcal{C})$.

Proposition 5.6 follows from the discussion above and the description of the process $(\gamma_n(\rho_t), t \in \mathbb{R}_+)$ around equation (5.3). As in [26], the operator A_{grow} reflects the growth of the genealogical distances between reproduction events that affects them. The operator A_{repr} stands for the jumps of the genealogical distances in reproduction events, as described by equation (5.3). We omit a formal proof of Proposition 5.6.

Remark 5.7. That the solutions of the martingale problems in Proposition 5.6 and in Proposition 6.4 below are unique can be shown by the approach from Section 11.3. We do not use this assertion in the present paper.

5.3 Properties of the genealogy at a fixed time

We consider the process $(\rho_t, t \in \mathbb{R}_+)$ from Subsection 5.2. To apply Theorem 4.1, we need exchangeability of the random variable ρ_t for each $t \in \mathbb{R}_+$.

Proposition 5.8. *Let $t \in \mathbb{R}_+$ and assume that ρ_0 is exchangeable. Then ρ_t is exchangeable.*

We prove Proposition 5.8 in Section 11.1.

Remark 5.9. For $t \in \mathbb{R}_+$, let $(\Pi_s^{(t)}, s \in [0,t])$ be the $\mathcal{P}$-valued stochastic process such that two integers $i, j \in \mathbb{N}$ are in the same block of $\Pi_s^{(t)}$ if and only if $\rho_t(i,j) \leq 2s$. Then a comparison of the Poisson process construction of the Ξ-coalescent in [48, Section 3] with the Poisson process construction from the present section shows that a Ξ-coalescent up to time t is given by the process $(\Pi_s^{(t)}, s \in [0,t))$. The distance matrix $\rho_t \wedge (2t)$ can be retrieved from $(\Pi_s^{(t)}, s \in [0,t))$ by

$$\rho_t(i,j) \wedge (2t) = 2 \inf\{s \in [0,t] : \ i \text{ and } j \text{ are in the same block of } \Pi_s^{(t)}, \text{ or } s = t\}$$

As Ξ-coalescents are exchangeable, it follows that the random variable $(\rho_t(i,j) \wedge (2t))_{i,j \in \mathbb{N}}$ is exchangeable. We remark that the collection of partitions $(\Pi_{(t-s)-}^{(t)}, 0 \leq s \leq t)$ is the dual flow of partitions from Foucart [24] in one-sided time. We also remark that preservation of exchangeability in the lookdown model is studied in e.g. [15, 16, 8].

For the application of Theorem 4.1, it is also of interest whether the states ρ_t are a.s. dust-free. Proposition 5.10 formulates the criterion from [48, Proposition 30] in our present context. We call the finite measure Ξ on Δ dust-free if

$$\Xi\{0\} > 0 \quad \text{or} \quad \int |x|_1 |x|_2^{-2} \Xi_0(dx) = \infty. \tag{5.7}$$

Proposition 5.10. *Let $t \in (0,\infty)$ and assume $\rho_0 \in \mathfrak{U}$. Then Ξ is dust-free if and only if ρ_t is a.s. dust-free.*

Proof. By Remark 5.2, $\rho_t \in \mathfrak{U}$, hence $\Upsilon(\rho)$ is well-defined. Clearly, ρ_t is dust-free if and only if the partition $\Pi_s^{(t)}$ from Remark 5.9 contains no singletons for all $s \in (0,t) \cap \mathbb{Q}$. This holds a.s. if and only if Ξ is dust-free by [48, Proposition 30]. $\square$

6 Decomposition of the genealogical distances

To apply Theorem 4.1(i) to the process $(\rho_t, t \in \mathbb{R}_+)$ from Section 5.2, we need to describe the $\hat{\mathfrak{U}}$-valued process $(\beta(\rho_t), t \in \mathbb{R}_+)$ by a martingale problem. A version of

this process that readily yields a description by a martingale problem is read off from the lookdown model in this section. We define such a process in Subsection 6.1 for a deterministic point measure η that drives the population model. In Subsection 6.2, we let η again be the Poisson random measure.

6.1 The deterministic construction

Let η be a simple point measure on $(0, \infty) \times \mathcal{P}$ (as in Section 5.1). Let $(r_0, v_0) \in \mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}}$. We interpret (r_0, v_0) as a decomposition of genealogical distances at time 0. For $i \in \mathbb{N}$, let

$$\mathcal{P}(i) = \{\pi \in \mathcal{P} : \{i\} \notin \pi\}$$

be the set of partitions of $\mathbb{N}$ in which i does not form a singleton block. If $\eta(\{s\} \times \mathcal{P}(A_s(t, i))) > 0$ for some $s \in (0, t]$, then we set

$$v_t(i) = t - \sup\{s \in (0, t] : \eta(\{s\} \times \mathcal{P}(A_s(t, i))) > 0\},$$

else we set

$$v_t(i) = t + v_0(A_0(t, i)).$$

The quantity $v_t(i)$ is the time back until an ancestor of the particle on level i at time t is involved in a reproduction event in which it belongs to a non-singleton block, if there is such an event, else $v_t(i)$ is defined from v_0.

We let $\rho_0 = \alpha(r_0, v_0)$ and define the process $(\rho_t, t \in \mathbb{R}_+)$ from η and ρ_0 as in Subsection 5.1. We set

$$r_t(i, j) = (\rho_t(i, j) - v_t(i) - v_t(j)) \mathbb{1}\{i \neq j\}$$

for $t \in \mathbb{R}_+$ and $i, j \in \mathbb{N}$. Then (r_t, v_t) can be thought of as a decomposition of the distance matrix ρ_t in the sense of Section 2. In this decomposition, we remove from the genealogical tree at time t the part between any leaf i and the most recent reproduction event on the ancestral lineage of this leaf, and we encode the length of this part as the mark $v_t(i)$.

Remark 6.1. Consider for this remark the following change (compared to our definition from Section 5.1) in the definition of the reproduction event encoded by a point $(t, \pi) \in \eta$: For each non-singleton block $B_i(\pi)$, the reproducing particle on level i at time $t-$ dies and is replaced at time t by its offspring on all the levels in $B_i(\pi)$. Then the quantity $v_t(i)$ is the age of the particle on level i at time t if this holds for $t = 0$. Condition (6.2) below ensures that the times at which the particles on a fixed level are replaced do not accumulate.

Analogously to Section 5.1, we give another description of the process $((r_t, v_t), t \in \mathbb{R}_+)$. Let $\mathcal{S}_n$ be the set of semi-partitions of $[n]$, that is, the set of systems of nonempty disjoint subsets of $[n]$. Every partition is also a semi-partition. However, in a semi-partition, there can be missing elements, that is, elements of $[n]$ that are not contained in the union $\cup \sigma$ of the blocks of σ. By "blocks" we mean the subsets of $[n]$ that are the elements of σ. From every semi-partition $\sigma \in \mathcal{S}_n$, a partition π is obtained by inserting a singleton block for each missing element. We call π the partition associated with σ, and we define $\sigma(i) = \pi(i)$ for each $i \in [n]$, where $\pi(i)$ is defined in Section 5.1. In order that equation (6.1) below hold, we associate with each element σ of $\mathcal{S}_n$ a transformation $\mathbb{R}^{n^2} \times \mathbb{R}^n \to \mathbb{R}^{n^2} \times \mathbb{R}^n$, which we also denote by σ, by $\sigma(r, v) = (r', v')$, where

$$v'(i) = v(\sigma(i)) \mathbb{1}\{i \notin \cup \sigma\}$$

and

$$r'(i, j) = (v(\sigma(i)) \mathbb{1}\{i \in \cup \sigma\} + r(\sigma(i), \sigma(j)) + v(\sigma(j)) \mathbb{1}\{j \in \cup \sigma\}) \mathbb{1}\{i \neq j\}$$

for $i, j \in [n]$.

We define the function

$$\varsigma_n : \mathcal{P} \to \mathcal{S}_n, \quad \pi \mapsto \{B \cap [n] : B \in \pi, \#B \geq 2\} \setminus \{\emptyset\}.$$

that removes all singleton blocks from a partition of $\mathbb{N}$ and restricts the semi-partition obtained in this way to a semi-partition of $[n]$. For each reproduction event encoded by a point $(s, \pi) \in \eta$, the corresponding jump of the process $((r_t, v_t), t \in \mathbb{R}_+)$ can be described by

$$\varsigma_n(\pi)(\gamma_n(r_{s-}, v_{s-})) = \gamma_n(r_s, v_s). \tag{6.1}$$

Here we cannot use the restriction $\gamma_n(\pi)$ (of π to $[n]$) instead of $\varsigma_n(\pi)$ as we cannot read off from $\gamma_n(\pi)$ which singleton blocks in $\gamma_n(\pi)$ are also singleton blocks in π.

We define the set of partitions

$$\hat{\mathcal{P}}^n = \{\pi \in \mathcal{P} : \varsigma_n(\pi) \neq \emptyset\}.$$

We remark that $\hat{\mathcal{P}}^n$ is the set of partitions of $\mathbb{N}$ in which not all of the first n integers form singleton blocks, hence it is strictly larger than the set $\mathcal{P}^n$. Only reproduction events that are encoded by a partition in $\hat{\mathcal{P}}^n$ affect the decomposed genealogical distances on the first n levels $(\gamma_n(r_t, v_t), t \in \mathbb{R}_+)$. If η satisfies the condition

$$\eta\left((0, t] \times \hat{\mathcal{P}}^n\right) < \infty \quad \text{for all } t \in (0, \infty) \text{ and } n \in \mathbb{N}. \tag{6.2}$$

then there are only finitely many reproduction events in bounded time intervals that result in a jump of the process $(\gamma_n(r_t, v_t), t \in \mathbb{R}_+)$. Between such jumps, the matrix r_t is constant, and the entries of the vector v_t grow linearly with slope 1, that is, $v_t(i) + s = v_{t+s}(i)$ for $i \in [n]$ and $t, s \in \mathbb{R}_+$ with $\eta((t, t+s] \times \hat{\mathcal{P}}^n) = 0$.

6.2 Stochastic evolution

Now let η be the Poisson random measure from Section 5.2 whose distribution is characterized by some finite measure Ξ on Δ. Consider the population model from Subsection 6.1 driven by the Poisson random measure η. For each $n \in \mathbb{N}$ and $\sigma \in \mathcal{S}_n \setminus \{\emptyset\}$, the rate at which reproduction events encoded by a partition in $\varsigma_n^{-1}(\sigma) \in \mathcal{P}$ occur is given by

$$\begin{aligned}
\lambda_{n,\sigma} =& H_\Xi(\varsigma_n^{-1}(\sigma)) \\
=& \int_\Delta \kappa(x, \varsigma_n^{-1}(\sigma)) |x|_2^{-2} \Xi_0(dx) + \Xi\{0\} \sum_{1 \leq i < j} \mathbb{1}\{K_{i,j} \in \varsigma_n^{-1}(\sigma)\} \\
=& \int_\Delta \sum_{\substack{i_1, \ldots, i_\ell \in \mathbb{N} \\ \text{pairwise distinct}}} x_{i_1}^{k_1} \cdots x_{i_\ell}^{k_\ell} (1 - |x|_1)^{n - k_1 - \ldots - k_\ell} |x|_2^{-2} \Xi_0(dx) \\
& + \Xi\{0\} \mathbb{1}\{\ell = 1, k_1 = 2\} + \infty \mathbb{1}\{\Xi\{0\} > 0, \ell = 1, k_1 = 1\} \tag{6.3}
\end{aligned}$$

where $\ell = \#\sigma$, and $k_1, \ldots, k_\ell \geq 1$ are the sizes of the subsets in σ in arbitrary order, and Ξ_0 is defined as in (5.4). For the last equality, we consider the paintbox partition π associated with $x \in \Delta$: With the notation from the beginning of Section 5.2, integers $i, j \in [n]$ are elements of a common subset in $\varsigma_n(\pi)$ if and only if U_i and U_j fall into a common subinterval that is not the dust interval. In particular, $i \notin \cup\varsigma_n(\pi)$ if and only if U_i falls into the dust interval.

Note that the rates λ_π for $\pi \in \mathcal{P}_n \setminus \{\mathbf{0}_n\}$, which we discussed already in Remark 5.5, satisfy

$$\lambda_\pi = H_\Xi\left(\gamma_n^{-1}(\pi)\right) = H_\Xi\left(\bigcup_\sigma \{\varsigma_n^{-1}(\sigma)\}\right) = \sum_\sigma \lambda_{n,\sigma}, \tag{6.4}$$

where the union and the sum are over all semi-partitions $\sigma \in \mathcal{S}_n$ with the same non-singleton blocks as π. In (6.4), we also use the restriction map $\gamma_n : \mathcal{P} \to \mathcal{P}_n$. From equations (6.3) and (6.4), we see that $\lambda_{\{\{1,2\}\}} = \Xi(\Delta) < \infty$ and $\lambda_\pi < \infty$ for all $\pi \in \mathcal{P}_n \setminus \{\mathbf{0}_n\}$, where $\mathbf{0}_n = \{\{1\}, \ldots, \{n\}\}$. This implies $\eta((0,t] \times \mathcal{P}^n) < \infty$ a.s. for all $t \in (0, \infty)$. That is, condition (5.1) is a.s. satisfied, as stated in Section 5.2. The condition (5.7) for Ξ to be dust-free is the condition that $\lambda_{1,\{\{1\}\}} = \infty$. That is, each particle reproduces with infinite rate if and only if Ξ is dust-free. Hence, if Ξ is not dust-free, then almost every realization of η satisfies condition (6.2). Moreover, if Ξ is not dust-free, then $\lambda_{n,\sigma} < \infty$ for all $n \in \mathbb{N}$ and $\sigma \in \mathcal{S}_n \setminus \{\emptyset\}$ as a consequence of equation (6.3).

Remark 6.2. Consider the case that Ξ is concentrated on $\{(x, 0, 0, \ldots) : x \in [0,1]\} \subset \Delta$. In this case, which corresponds to the Λ-coalescent, a.s. no simultaneous multiple reproduction events occur. The measure Ξ_0 is then determined by the finite measure $\Lambda_0 = \varpi(\Xi_0)$, where $\varpi : \Delta \to [0,1]$, $x \mapsto x_1$. For $B \subset [n]$ and $k = \#B$, it then follows

$$\lambda_{n,\{B\}} = \int_{[0,1]} x^k (1-x)^{n-k} x^{-2} \Lambda_0(dx) + \Xi\{0\} \mathbb{1}\{k = 2\} + \infty \, \mathbb{1}\{\Xi\{0\} > 0, k = 1\}.$$

The rates $\lambda_{n,\sigma}$ for $\sigma \in \mathcal{S}_n$ with $\#\sigma > 1$ are equal to zero in this case.

Now we consider the $\mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}}$-valued process $((r_t, v_t), t \in \mathbb{R}_+)$ from Subsection 6.1, driven by the Poisson random measure η. The initial state is defined as a $\mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}}$-valued random variable (r_0, v_0) that is independent of η.

Proposition 6.3. *The process $((r_t, v_t), t \in \mathbb{R}_+)$ is Markov.*

Proof. This follows by the same argument as for Proposition 5.4. $\qquad \square$

Recall the set $\hat{\mathcal{C}}$ from Section 4. For $(r, v) \in \mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}}$ and $\phi \in \hat{\mathcal{C}}_n$, we write

$$\langle \nabla^v \phi, \underline{1} \rangle (r, v) = \sum_{i \in \mathbb{N}} \frac{\partial}{\partial v(i)} \phi(r, v).$$

From the discussion above and the description of the process $(\gamma_n(r_t, v_t), t \in \mathbb{R}_+)$ around equation (6.1), we deduce the next proposition.

Proposition 6.4. *Assume that Ξ is not dust-free. Define an operator $\hat{A} = \hat{A}_{\mathrm{grow}} + \hat{A}_{\mathrm{repr}}$ with domain $\hat{\mathcal{C}}$ by*

$$\hat{A}_{\mathrm{grow}} \phi(r, v) = \langle \nabla^v \phi, \underline{1} \rangle (r, v)$$

and

$$\hat{A}_{\mathrm{repr}} \phi(r, v) = \sum_{\sigma \in \mathcal{S}_n \setminus \{\emptyset\}} \lambda_{n,\sigma} (\phi(\sigma(\gamma_n(r, v))) - \phi(r, v))$$

for $n \in \mathbb{N}$, $\phi \in \hat{\mathcal{C}}_n$ and $(r, v) \in \mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}}$. Then the stochastic process $((r_t, v_t), t \in \mathbb{R}_+)$ solves the martingale problem $(\hat{A}, \hat{\mathcal{C}})$.

The operator $\hat{A}_{\mathrm{grow}}$ accounts for the growth of the marks v_t which is described in the end of Subsection 6.1. The operator $\hat{A}_{\mathrm{repr}}$ stands for the jumps of the decomposed genealogical distances in reproduction events which are described by equation (6.1). We omit a formal proof of Proposition 6.4.

Finally, we consider again the process $(\rho_t, t \in \mathbb{R}_+)$ which is defined from the Poisson random measure η and the initial state $\rho_0 = \alpha(r_0, v_0)$ as in Section 5.1. We assume $\rho_0 \in \mathfrak{U}$. Then $\rho_t \in \mathfrak{U}$ for all $t \in \mathbb{R}_+$ by Remark 5.2. Moreover, the construction in Subsection 6.1 and the definition of the map α in Section 2 yield $\rho_t = \alpha(r_t, v_t)$ and $(r_t, v_t) \in \hat{\mathfrak{U}}$ for all $t \in \mathbb{R}_+$. We further assume that

$$(r_0, v_0) = \beta(\rho_0). \tag{6.5}$$

Then by the following proposition, the decomposition (r_t, v_t) of the semi-ultrametric ρ_t is the one given by the map β from Section 2, namely the decomposition into the external branches and the remaining subtree.

Proposition 6.5. *Assumption (6.5) implies $(r_t, v_t) = \beta(\rho_t)$ a.s. for each $t \in \mathbb{R}_+$.*

We prove Proposition 6.5 in Section 11.2.

Corollary 6.6. *The process $(\beta(\rho_t), t \in \mathbb{R}_+)$ is Markov and solves the martingale problem $(\hat{A}, \hat{\mathcal{C}})$ from Proposition 6.4.*

Proof. This is immediate from Propositions 6.3, 6.4 and 6.5. The Markov property can alternatively be seen from Proposition 5.4 and Remark 4.2. □

7 Tree-valued Fleming-Viot processes

In this section, we apply Theorem 4.1 to the process $(\rho_t, t \in \mathbb{R}_+)$ from Section 5.2. By Remark 5.2, we can consider $(\rho_t, t \in \mathbb{R}_+)$ as an $\mathfrak{U}$-valued process. We call all the image processes in Theorem 4.1 tree-valued Fleming-Viot processes. To distinguish them, we also call them $\mathbb{U}$-, $\hat{\mathbb{U}}$-, and $\mathcal{U}^{\mathrm{erg}}$-valued Ξ-Fleming-Viot processes. Proposition 7.1 below states that the martingale problems for the tree-valued Fleming-Viot processes have unique solutions.

7.1 Processes with values in the space of metric measure spaces

In this subsection, we consider a finite measure Ξ on Δ that is dust-free. Let $\chi \in \mathbb{U}$, and let $(\rho_t, t \in \mathbb{R}_+)$ be the $\mathfrak{U}$-valued Markov process from Section 5.2 that is defined in terms of Ξ and an initial state ρ_0 with distribution ν^χ. We define a $\mathbb{U}$-valued Ξ-Fleming-Viot process $(\chi_t, t \in \mathbb{R}_+)$ with initial state $\chi \in \mathbb{U}$ by $\chi_t = \psi(\rho_t)$. As a justification for this name, we note that $\chi_0 = \psi(\rho_0) = \chi$ a.s. by Corollary 3.12(iii) and Remark 3.5. By Theorem 4.1 and Propositions 5.4, 5.6, 5.8, and 5.10, the process $(\chi_t, t \in \mathbb{R}_+)$ is Markovian and solves the martingale problem (B, Π), where the generator B is defined by $B\Phi(\chi) = \nu^\chi(A\phi)$ for $\phi \in \mathcal{C}$ with associated polynomial $\Phi \in \Pi$, and $\chi \in \mathbb{U}$. Here A is the generator defined in Proposition 5.6. The martingale problem (B, Π) is a generalization of the martingale problem in Theorem 1 of Greven, Pfaffelhuber, and Winter [26].

7.2 Processes with values in the space of marked metric measure spaces

Let Ξ be a general finite measure on the simplex Δ. Let $\hat{\chi} \in \hat{\mathbb{U}}$, let ρ_0 be a $\mathfrak{U}$-valued random variable with distribution $\alpha(\nu^{\hat{\chi}})$, and let the $\mathfrak{U}$-valued Markov process $(\rho_t, t \in \mathbb{R}_+)$ be defined, as in Section 5.2, from Ξ and the initial state ρ_0. We define a $\hat{\mathbb{U}}$-valued Ξ-Fleming-Viot process $(\hat{\chi}_t, t \in \mathbb{R}_+)$ with initial state $\chi \in \hat{\mathbb{U}}$ by $\hat{\chi}_t = \hat{\psi}(\beta(\rho_t))$ for $t \in \mathbb{R}_+$. To justify this name, we note that the initial state satisfies $\hat{\chi}_0 = \hat{\psi}(\beta(\rho_0)) = \chi$ a.s. by Theorem 3.9(iii). By Theorem 4.1 and Propositions 5.4 and 5.8, the process $(\hat{\chi}_t, t \in \mathbb{R}_+)$ is Markovian.

If Ξ is not dust-free, then by Theorem 4.1 and Corollary 6.6, the process $(\hat{\chi}_t, t \in \mathbb{R}_+)$ solves the martingale problem $(\hat{B}, \hat{\Pi})$, where the generator $\hat{B}$ is defined by $\hat{B}\Phi(\chi') = \nu^{\chi'}(\hat{A}\phi)$ for all $\phi \in \hat{\mathcal{C}}$ with associated marked polynomial Φ, and all $\chi' \in \hat{\mathbb{U}}$. Here the generator $\hat{A}$ is defined as in Proposition 6.4.

If Ξ is dust-free, then for each $t \in (0, \infty)$ by Remark 4.3 and Proposition 5.10, the marked metric measure space $\hat{\chi}_t$ is a.s. dust-free, and $\hat{\chi}_t$ is determined a.s. by the associated metric measure space χ_t.

7.3 Processes with values in the space of distance matrix distributions

Let $(\hat{\chi}_t, t \in \mathbb{R}_+)$ be the process from Section 7.2, where Ξ is a general finite measure on the simplex Δ. We define a $\mathcal{U}^{\mathrm{erg}}$-valued Ξ-Fleming-Viot process $(\xi_t, t \in \mathbb{R}_+)$ with initial state $\alpha(\nu^{\hat{\chi}_0}) \in \mathcal{U}^{\mathrm{erg}}$ by $\xi_t = \alpha(\nu^{\hat{\chi}_t})$. Again by Theorem 4.1 and Propositions 5.4, 5.6 and 5.8, it follows that $(\xi_t, t \in \mathbb{R}_+)$ is Markovian and solves the martingale problem $(C, \mathscr{C})$, where the generator C is defined by $C\Psi(\xi) = \xi(A\phi)$ for all $\xi \in \mathcal{U}^{\mathrm{erg}}$, $\phi \in \mathcal{C}$, and $\Psi \in \mathscr{C} : \xi' \mapsto \xi'\phi$. Here the generator A is defined as in Proposition 5.6.

7.4 Well-posedness of the martingale problem, connection with the flow of bridges

Proposition 7.1. *The martingale problems* (B, Π), $(\hat{B}, \hat{\Pi})$, *and* $(C, \mathscr{C})$ *are well-posed.*

That a martingale problem is well-posed means that a solution exists whose finite-dimensional distributions are uniquely determined by the initial state. A proof of Proposition 7.1 by duality is given in Section 11.3.

Remark 7.2 (Connection with the flow of bridges)**.** A random non-decreasing right-continuous function $\tilde{F} : [0,1] \to [0,1]$ with exchangeable increments and $\tilde{F}(0) = \tilde{F}(1) = 1$ is called a bridge. The dual flow of bridges of Bertoin and Le Gall [5] is a collection $F = (F_{s,t}, s < t)$ of bridges that satisfies the following properties (see [5, Section 5.1]):

(i) For every $s < t < u$, $F_{t,u} \circ F_{s,t} = F_{s,u}$ a.s.

(ii) The law of $F_{s,t}$ depends only on $t - s$. For $s_1 < s_2 < \ldots < s_n$, the bridges $F_{s_1,s_2}, F_{s_2,s_3}, \ldots, F_{s_{n-1},s_n}$ are independent.

(iii) $F_{0,0}$ is the identity function. For every $x \in [0,1]$, the random variable $F_{0,t}(x)$ converges to x in probability as t decreases to zero.

For each $s < t$, it is also assumed that $F_{s,t}$ is a.s. not the identity function.

The interpretation is that the individuals of a continuous population are represented by the elements of the interval $[0,1]$. For each $s \leq t$, the individuals in a subinterval $(x_1, x_2]$ at time s have descendants at time t that are a.s. the elements of $(F_{s,t}(x_1), F_{s,t}(x_2)]$, see [7].

In [5, Section 3], Kingman's correspondence is extended so as to represent distributions of Ξ-coalescents in terms of sampling from flows of bridges. Let F be a dual flow of bridges, and let $V = (V_i, i \in \mathbb{N})$ be an iid sequence of uniform $[0,1]$-valued random variables, independent of F. This iid sequence is interpreted as a sequence of random samples from the population at some time $t \in \mathbb{R}$. For each $s \in \mathbb{R}_+$, a partition $\tilde{\pi}_s^{(t)}$ is defined such that any integers $i, j \in \mathbb{N}$ are in the same block of $\tilde{\pi}_s^{(t)}$ if and only if $F_{t-s,t}^{-1}(V_i) = F_{t-s,t}^{-1}(V_j)$ which means that these samples have the same ancestor at time $t - s$. Here we set $f^{-1}(t) = \inf\{s \in [0,1] : f(s) > t \text{ or } s = 1\}$ for $t \in [0,1]$ and a càdlàg function $f : [0,1] \to [0,1]$. In [5, Theorem 1], it is shown that the partition-valued process $(\tilde{\pi}_s^{(t)}, s \in \mathbb{R}_+)$ obtained in this way is a version of a Ξ-coalescent of Schweinsberg [48].

For each $t \in \mathbb{R}$, there exists an event of probability 1 on which for all $s \leq s' \in \mathbb{Q}_+$, the partition $\tilde{\pi}_{s'}^{(t)}$ can be obtained by merging blocks of the partition $\tilde{\pi}_s^{(t)}$. We can thus define a.s. an ultrametric $\tilde{\rho}_t$ by

$$\tilde{\rho}_t(i, j) = 2 \inf \left\{ s \in \mathbb{Q}_+ : i \text{ and } j \text{ are in the same block of } \tilde{\pi}_s^{(t)} \right\}.$$

Moreover, we define a.s. a random variable $\tilde{\nu}_t$ with values in the space $(\mathcal{U}, d_{\mathrm{P}})$ of exchangeable distributions on $\mathfrak{U}$ such that $\tilde{\nu}_t$ is a regular conditional distribution of $\tilde{\rho}_t$ given the collection of bridges $(F_{t-s,t}, s \in \mathbb{Q}_+)$. For the existence of this regular conditional distribution, see e.g. [31, Theorem 6.3].

In Section 12 of the arXiv version [27] of this article, it is shown that there exists a finite measure Ξ on Δ such that the process $(\tilde{\xi}_t, t \in \mathbb{R})$ is a stationary $\mathcal{U}^{\mathrm{erg}}$-valued Ξ-Fleming-Viot process in two-sided time (as in Section 9). In the proof, the conditional distribution of $\tilde{\rho}_{t+u}$ given $(\tilde{\xi}_s, s \le t)$ is compared for each $t \in \mathbb{R}$, $u \in \mathbb{R}_+$ to a corresponding object in the lookdown construction. From this, it is deduced that conditional distribution of $\tilde{\xi}_{t+u}$ given $(\tilde{\xi}_s, s \le t)$ is as desired.

8 Some semigroup properties

In this section, we state Feller continuity of tree-valued Ξ-Fleming-Viot processes, and that the domains of the martingale problems for them are cores. We consider $\hat{\mathbb{U}}$-valued Ξ-Fleming-Viot processes in detail, analogous results hold for the other processes from Section 7.

Let Ξ be a finite measure on the simplex Δ. For $\chi \in \hat{\mathbb{U}}$, let $(\hat{\chi}_t, t \in \mathbb{R}_+)$ under the probability measure $\mathbb{P}_\chi$ with associated expectation $\mathbb{E}_\chi$ be the $\hat{\mathbb{U}}$-valued Ξ-Fleming-Viot process from Section 7.2 with initial state χ. We denote by $C_{\mathrm{b}}(E)$ the set of bounded continuous $\mathbb{R}$-valued functions on a metric space E. We endow $C_{\mathrm{b}}(E)$ with the supremum norm.

The results in this section rely on the following lemma which we prove in Section 11.4 using the lookdown construction.

Lemma 8.1. *For each $t \in \mathbb{R}_+$ and $\Phi \in \hat{\Pi}$, the function $\hat{\mathbb{U}} \to \mathbb{R}$, $\chi \mapsto \mathbb{E}_\chi[\Phi(\hat{\chi}_t)]$ is an element of $\hat{\Pi}$.*

As a corollary, we obtain the Feller continuity of a $\hat{\mathbb{U}}$-valued Ξ-Fleming-Viot process, namely that its semigroup preserves the set of bounded continuous functions.

Corollary 8.2. *For each $t \in \mathbb{R}_+$ and $f \in C_{\mathrm{b}}(\hat{\mathbb{U}})$, the map $\hat{\mathbb{U}} \to \mathbb{R}$, $\chi \mapsto \mathbb{E}_\chi[f(\hat{\chi}_t)]$ is continuous.*

Proof. This follows from Lemma 8.1 as the set $\hat{\Pi}$ of marked polynomials is convergence determining, we use the definition of convergence in distribution in $\hat{\mathbb{U}}$. $\qquad \square$

Let $\hat{L}$ denote the closure of $\hat{\Pi}$ in $C_{\mathrm{b}}(\hat{\mathbb{U}})$ with respect to the supremum norm. For application in [28], we note two more corollaries of Lemma 8.1. The first of them states that the semigroup of a $\hat{\mathbb{U}}$-valued Ξ-Fleming-Viot process can be restricted to a semigroup on $\hat{L}$ that is strongly continuous.

Corollary 8.3. *Let $f \in \hat{L}$. Then for each $t \in \mathbb{R}_+$, the function $\hat{\mathbb{U}} \to \mathbb{R}$, $\chi \mapsto \mathbb{E}_\chi[f(\hat{\chi}_t)]$ is an element of $\hat{L}$. Moreover,*

$$\lim_{t \downarrow 0} \sup_{\chi \in \hat{\mathbb{U}}} |\mathbb{E}_\chi[f(\hat{\chi}_t)] - \mathbb{E}_\chi[f(\hat{\chi}_0)]| = 0.$$

Proof. The first assertion follows from Lemma 8.1 and the definition of $\hat{L}$. As $(\hat{\chi}_t, t \in \mathbb{R}_+)$ solves the martingale problem $(\hat{B}, \hat{\Pi})$ from Section 7.2,

$$\mathbb{E}_\chi[\Phi(\hat{\chi}_t)] - \mathbb{E}_\chi[\Phi(\hat{\chi}_0)] = \mathbb{E}_\chi \left[\int_0^t \hat{B}\Phi(\hat{\chi}_s)ds \right]$$

for all $t \in \mathbb{R}_+$ and $\Phi \in \hat{\Pi}$. The second assertion follows as $\hat{B}\Phi$ is bounded and by definition of $\hat{L}$. $\qquad \square$

The next corollary says that the semigroup on $\hat{L}$ of a $\hat{\mathbb{U}}$-valued Ξ-Fleming-Viot process is generated by the closure of the operator $\hat{B}$ with domain $\hat{\Pi}$, see [19, Chapter 1] for the definitions.

Corollary 8.4. *The subspace $\hat{\Pi} \subset C_{\mathrm{b}}(\hat{\mathbb{U}})$ is a core for the generator of the semigroup on $\hat{L}$ of a $\hat{\mathbb{U}}$-valued Ξ-Fleming-Viot process.*

Proof. We note that $\hat{B}$ is the restriction of the generator of the semigroup to $\hat{\Pi}$ and apply Proposition 1.3.3 and Corollary 1.1.6 of [19], using Lemma 8.1 and Corollary 8.3. □

Let L be the closure of Π in $C_{\mathrm{b}}(\mathbb{U})$ and let L' be the closure of $\mathscr{C}$ in $C_{\mathrm{b}}(\mathcal{U}^{\mathrm{erg}})$, with respect to the supremum norm. In the same way as above, it can be shown: The semigroup on L' of a $\mathcal{U}^{\mathrm{erg}}$-valued Ξ-Fleming-Viot process is strongly continuous and generated by the closure of the operator C with domain $\mathscr{C}$ from Section 7.3. If Ξ is dust-free, then the semigroup on L of a $\mathbb{U}$-valued Ξ-Fleming-Viot process is strongly continuous and generated by the closure of the operator B with domain Π from Section 7.1. Continuity properties analogous to Proposition 8.2 also hold.

9 Convergence to equilibrium

Let Ξ be a finite measure on the simplex Δ with $\Xi(\Delta) > 0$. We show convergence to equilibrium for the $\hat{\mathfrak{U}}$-valued process $(\beta(\rho_t), t \in \mathbb{R}_+)$ from Section 6.2. From this, we deduce in Proposition 9.1 that also the tree-valued Ξ-Fleming-Viot process from Section 7.2 converges to equilibrium. In the same way, it can be shown that the other processes from Section 7 converge to equilibrium.

We define stationary processes and use a coupling argument. Analogously to Section 5.2, let $\bar{\eta}$ be a Poisson random measure on $\mathbb{R} \times \mathcal{P}$ with intensity $dt\, H_\Xi(d\pi)$. This Poisson random measure drives a population model in two-sided time (with time axis $\mathbb{R}$) where the reproduction events and the ancestral levels $\bar{A}_s(t, i)$ are defined as in Section 5.1. Then we define the stationary $\mathfrak{U}$-valued process $(\bar{\rho}_t, t \in \mathbb{R})$ of the genealogical distances by

$$\bar{\rho}_t(i, j) = 2t - 2\sup\{s \in (-\infty, t] : \bar{A}_s(t, i) = \bar{A}_s(t, j)\}$$

for $t \in \mathbb{R}$, $i, j \in \mathbb{N}$. On an event of probability 1, all these distances are finite. This follows from the assumption that $\Xi(\Delta) > 0$. That $\bar{\rho}_t$ is indeed a semi-ultrametric for each $t \in \mathbb{R}$ can be seen as in Remark 5.2. Clearly, ρ_t is exchangeable, which follows from exchangeability of the Ξ-coalescent as in Remark 5.9 or can be shown as in the proof of Proposition 5.8.

Let η denote the restriction of $\bar{\eta}$ to $(0, \infty) \times \mathcal{P}$. Let $\chi \in \hat{\mathbb{U}}$ be arbitrary, and let ρ_0 be a $\mathfrak{U}$-valued random variable with distribution $\alpha(\nu^\chi)$, independent of η. Let the process $(\rho_t, t \in \mathbb{R}_+)$ be defined from ρ_0 and η as in Section 5.1. For $n \geq 2$, on the event $\{\max_{i,j \in [n]} \bar{\rho}_t(i, j) < 2t\}$, the marked distance matrix $\gamma_n(\beta(\rho_t))$ does not depend on ρ_0. This follows from the construction in Section 5.1 and the definition of the map β in Section 2. As $\bar{\rho}_t$ can also be obtained from $\bar{\rho}_0$ and η as in Section 5.1, it follows that $\gamma_n(\beta(\rho_t)) = \gamma_n(\beta(\bar{\rho}_t))$ on the event $\{\max_{i,j \in [n]} \bar{\rho}_t(i, j) < 2t\}$. By stationarity of $(\bar{\rho}_t, t \in \mathbb{R})$, it follows that

$$|\mathbb{E}[\phi(\beta(\rho_t))] - \mathbb{E}[\phi(\beta(\bar{\rho}_0))]| \leq 2\sup|\phi|\, \mathbb{P}\left(\max_{i,j \in [n]} \bar{\rho}_0(i, j) \geq 2t\right) \to 0 \quad (t \to \infty) \qquad (9.1)$$

for all $\phi \in \hat{\mathcal{C}}_n$.

We call a $\hat{\mathbb{U}}$-valued random variable that is distributed as $\bar{\chi}_0 := \hat{\psi}(\beta(\bar{\rho}_0))$ a Ξ-coalescent measure tree, generalizing the Λ-coalescent measure tree from [26]. A $\hat{\mathbb{U}}$-valued Ξ-Fleming-Viot process $(\chi_t, t \in \mathbb{R}_+)$ with initial state χ is given by $\chi_t = \hat{\psi}(\beta(\rho_t))$, as in Section 7.2.

Proposition 9.1. *The $\hat{\mathbb{U}}$-valued random variable χ_t converges in distribution to a Ξ-coalescent measure tree as $t \to \infty$.*

Proof. As in Proposition 10.10 below, we obtain $\mathbb{E}[\Phi(\bar{\chi}_0)] = \mathbb{E}[\phi(\bar{\rho}_0)]$ and $\mathbb{E}[\Phi(\chi_t)] = \mathbb{E}[\phi(\rho_t)]$. The convergence (9.1) then yields that $\mathbb{E}[\Phi(\chi_t)]$ converges to $\mathbb{E}[\Phi(\bar{\chi}_0)]$ as $t \to \infty$ for all marked polynomials $\Phi \in \hat{\Pi}$. The assertion follows as the set $\hat{\Pi}$ is convergence determining in $\hat{\mathbb{U}}$. $\qquad\square$

A stationary $\hat{\mathbb{U}}$-valued Ξ-Fleming-Viot process can be defined by $(\hat{\psi}(\beta(\bar{\rho}_t)), t \in \mathbb{R})$. In [26, Theorem 3], duality is used to show that the tree-valued Fleming-Viot process converges to an equilibrium. In [16, Theorem 4.1], convergence to stationarity of measure-valued Fleming-Viot processes is also proved by a coupling argument.

10 Proofs of the general results

In Subsection 10.1, we prove Proposition 3.4 which is needed for the proof of the uniqueness result (Proposition 3.11) in Subsection 10.3. We prove the sampling representation (Theorem 3.9) in Subsections 10.2 – 10.4. Theorem 4.1 gives the application to tree-valued processes and is proved in Subsection 10.5.

10.1 Proof of Proposition 3.4

The proof of this result from Section 3.2 relies on the fact that in a separable metric space, an iid sequence with respect to a probability measure on the Borel sigma algebra has no isolated elements.

Proof of Proposition 3.4. Let $((x(i), v(i)), i \in \mathbb{N})$ be an m-iid sequence in $X \times \mathbb{R}_+$. We may assume

$$r = (r(i,j))_{i,j \in \mathbb{N}} = (r'(x(i), x(j)))_{i,j \in \mathbb{N}}.$$

We write $\rho = \alpha(r, v)$. We show that $v = \Upsilon(\rho)$ a.s. from which the assertion follows by definition of the map β.

Let $\varepsilon > 0$ and $i \in \mathbb{N}$. By separability, $X \times \mathbb{R}_+$ can be covered by countably many balls of diameter ε. This implies

$$m\{(x', v') \in X \times \mathbb{R}_+ : r'(x(i), x') \vee |v(i) - v'| \le 2\varepsilon\} > 0 \quad \text{a.s.},$$

and that there exists a random $j \in \mathbb{N} \setminus \{i\}$ with

$$r'(x(i), x(j)) \vee |v(i) - v(j)| \le 2\varepsilon \quad \text{a.s.} \tag{10.1}$$

By inequality (10.1) and the definition of ρ, it follows that

$$2v(i) + 4\varepsilon \ge v(i) + v(j) + r(i,j) = \rho(i,j) \quad \text{a.s.}$$

Using the definition of the map Υ, we deduce

$$v(i) + 2\varepsilon \ge \tfrac{1}{2}\rho(i,j) \ge \Upsilon(\rho)(i) \quad \text{a.s.}$$

For the converse inequality, we first note that

$$2v(i) \le v(i) + v(j) + 2\varepsilon + r(i,j) = \rho(i,j) + 2\varepsilon \tag{10.2}$$

by inequality (10.1) and the definition of ρ. Moreover, for all $k \in \mathbb{N} \setminus \{i, j\}$, we obtain

$$
\begin{aligned}
2v(i) - 2\varepsilon &\le \rho(i,j) \le \rho(i,k) \vee \rho(k,j) \\
&\le v(k) + r(i,k) \vee r(k,j) + v(i) \vee v(j) \\
&\le v(k) + r(i,k) + v(i) + |r(k,j) - r(i,k)| + |v(j) - v(i)| \\
&\le \rho(i,k) + r(i,j) + 2\varepsilon \le \rho(i,k) + 4\varepsilon \quad \text{a.s.}
\end{aligned}
$$

Here we use inequality (10.2) for the first and inequality (10.1) for the fifth and sixth step, the definition of ρ for the third and fifth step, and ultrametricity for the second step. By definition of the map Υ, we obtain

$$\Upsilon(\rho)(i) = \tfrac{1}{2} \inf_{k \in \mathbb{N} \setminus \{i\}} \rho(i,k) \geq v(i) - 3\varepsilon \quad \text{a. s.}$$

As $\varepsilon > 0$ and $i \in \mathbb{N}$ were arbitrary, it follows that $\Upsilon(\rho) = v$ a. s. $\qquad\square$

10.2 Measurability of the construction of (marked) metric measure spaces

In this subsection, we show Proposition 3.7 from Section 3.3. We only discuss measurability of the map $\hat{\psi} : \mathfrak{D} \times \mathbb{R}_+^{\mathbb{N}} \to \hat{\mathbb{M}}$ therein. Measurability of the map $\psi : \mathfrak{D} \times \mathbb{M}$ follows along the same lines.

Recall that the Prohorov distance between two probability measures μ and μ' on the Borel sigma algebra on a metric space (Z, d^Z) is given by

$$d_{\mathrm{P}}^Z(\mu, \mu') = \inf\{\varepsilon > 0 : \mu(F) \leq \mu'(F^\varepsilon) + \varepsilon \text{ for all closed } F \subset Z\}, \qquad (10.3)$$

where $F^\varepsilon = \{z \in Z : d^Z(z, F) < \varepsilon\}$. If (Z, d^Z) is separable, then the coupling characterization of the Prohorov distance holds, which can be found e. g. in [19, Theorem 3.1.2]:

$$d_{\mathrm{P}}^Z(\mu, \mu') = \inf_{\xi} \inf\{\varepsilon > 0 : \xi\{(x,y) \in Z^2 : d^Z(x,y) > \varepsilon\} < \varepsilon\}, \qquad (10.4)$$

where the first infimum is over all couplings ξ of the probability measures μ and μ'.

We also use the marked Gromov-Prohorov distance d_{mGP} which metrizes the marked Gromov-weak topology on $\hat{\mathbb{M}}$, see [12]. It is defined by

$$d_{\mathrm{mGP}}((X, r, m), (X', r', m')) = \inf_{Z, \varphi, \varphi'} d_{\mathrm{P}}^Z(\hat{\varphi}(m), \hat{\varphi}'(m'))$$

for marked metric measure spaces (X, r, m) and (X', r', m'). Here the infimum is over all isometric embeddings $\varphi : X \to Z$ and $\varphi' : X' \to Z$ into complete and separable metric spaces (Z, d^Z). The space $Z \times \mathbb{R}_+$ is endowed with the product metric $d^{Z \times \mathbb{R}_+}((z, v), (z', v')) = d^Z(z, z') \vee |v - v'|$, analogously for $X \times \mathbb{R}_+$ and $X' \times \mathbb{R}_+$. The maps $\hat{\varphi} : X \times \mathbb{R}_+ \to Z \times \mathbb{R}_+$ and $\hat{\varphi}' : X' \times \mathbb{R}_+ \to Z \times \mathbb{R}_+$ are defined by $\hat{\varphi}(x, v) = (\varphi(x), v)$, $(x, v) \in X \times \mathbb{R}_+$ and $\hat{\varphi}'(x', v) = (\varphi'(x'), v)$, $(x', v) \in X' \times \mathbb{R}_+$.

We write $\hat{\mathfrak{D}} = \mathfrak{D} \times \mathbb{R}_+^{\mathbb{N}}$. For $n \in \mathbb{N}$, we denote by

$$\hat{\mathfrak{D}}_n = \{(r, v) \in \mathbb{R}_+^{n^2} \times \mathbb{R}_+^n : r(i,i) = 0, r(i,j) = r(j,i),$$
$$r(i,j) + r(j,k) \geq r(i,k) \text{ for all } i, j, k \in [n]\}$$

the space of decomposed semimetrics on $[n]$ which we view as a subspace of $\mathbb{R}_+^{n^2} \times \mathbb{R}^n$. We denote by $\hat{\psi}_n : \hat{\mathfrak{D}}_n \to \hat{\mathbb{M}}$ the function that maps $(r, v) \in \hat{\mathfrak{D}}_n$ to the isomorphy class of the marked metric measure space $([n], r, n^{-1} \sum_{i=1}^n \delta_{(i, v(i))})$, here we also identify the elements of the semi-metric space $([n], r)$ with distance zero.

Lemma 10.1. *The map $\hat{\psi}_n : \mathfrak{D}_n \to \hat{\mathbb{M}}$ is continuous.*

Proof. W. l. o. g. we can assume that $\hat{\mathfrak{D}}$ is endowed with the metric d that is given by

$$d((r, v), (r', v')) = \sup_{k \in \mathbb{N}} \left(\left(\max_{i, j \in [k]} |r(i,j) - r'(i,j)| \vee \max_{i \in [k]} |v(i) - v'(i)| \right) \wedge (2^{-k}) \right)$$

for all $(r, v), (r', v') \in \hat{\mathfrak{D}}$. For $(r, v), (r', v') \in \hat{\mathfrak{D}}_n$, we define a probability measure ξ on $(\hat{\mathfrak{D}})^2$ as the distribution of $((r(x_i, x_j))_{i,j \in \mathbb{N}}, (\tilde{v}_i)_{i \in \mathbb{N}}, (r'(x_i', x_j'))_{i,j \in \mathbb{N}}, (\tilde{v}_i')_{i \in \mathbb{N}})$, where

$(x_i, \tilde{v}_i, x'_i, \tilde{v}'_i)_{i \in \mathbb{N}}$ is an iid sequence with distribution $n^{-1} \sum_{k=1}^{n} \delta_{(k,v(k),k,v'(k))}$. Then $\xi(\cdot \times \hat{\mathfrak{D}}) = \nu^{\hat{\psi}_n(r,v)}$ and $\xi(\hat{\mathfrak{D}} \times \cdot) = \nu^{\hat{\psi}_n(r',v')}$. For

$$c := \max_{i,j \in [n]} |r(i,j) - r'(i,j)| \vee \max_{i \in [n]} |v(i) - v'(i)|,$$

the coupling characterization (10.4) implies

$$d_{\mathrm{P}}\left(\nu^{\hat{\psi}_n(r,v)}, \nu^{\hat{\psi}_n(r',v')}\right) \le c + \xi\left\{(y,y') \in \hat{\mathfrak{D}}^2 : d(y,y') > c\right\} = c.$$

Continuity of $\hat{\psi}_n$ follows by definition of the marked Gromov-weak topology. $\qquad \square$

Proof of Proposition 3.7. Let $(r,v) \in \hat{\mathfrak{D}}^*$ and let (X,r) be the metric completion of $(\mathbb{N}, r)$. We endow the product space $X \times \mathbb{R}_+$ with the metric $d^{X \times \mathbb{R}_+}((x,v),(x',v')) = r(x,v) \vee |v - v'|$. The definition of $\hat{\mathfrak{D}}^*$ yields $\lim_{n \to \infty} d_{\mathrm{P}}^{X \times \mathbb{R}_+}(n^{-1} \sum_{i=1}^{n} \delta_{(i,v(i))}, m) = 0$ for a probability measure m on $X \times \mathbb{R}_+$. As $\hat{\psi}(r,v)$ equals the isomorphy class of (X,r,m), and as $\hat{\psi}_n(\gamma_n(r,v))$ equals the isomorphy class of $(X,r,n^{-1} \sum_{i=1}^{n} \delta_{(i,v(i))})$ for each $n \in \mathbb{N}$, the definition of the marked Gromov-Prohorov metric implies that $\lim_{n \to \infty} d_{\mathrm{mGP}}(\hat{\psi}(r,v), \hat{\psi}_n(\gamma_n(r,v))) = 0$.

For $(r,v) \in \hat{\mathfrak{D}} \setminus \hat{\mathfrak{D}}^*$, the image $\hat{\psi}(r,v)$ is constant by definition. Using Lemma 10.1 and Lemma 10.2 below, we deduce measurability of $\hat{\psi}$. $\qquad \square$

Lemma 10.2. *The subsets $\hat{\mathfrak{D}}^* \subset \hat{\mathfrak{D}}$ and $\mathfrak{D}^* \subset \mathfrak{D}$ are measurable.*

Proof. We represent $\hat{\mathfrak{D}}^*$ by countable unions and intersections of measurable sets. The assertion on $\mathfrak{D}$ follows along the same lines by removing the marks v.

For $(r,v) \in \hat{\mathfrak{D}}$, let (X,r) be the metric completion of (X,r). We endow the product space $X \times \mathbb{R}_+$ with the metric $d^{X \times \mathbb{R}_+}((x,v),(x',v')) = r(x,v) \vee |v - v'|$ and define for $n \in \mathbb{N}$ the probability measures $m^n = n^{-1} \sum_{i=1}^{n} \delta_{(i,v(i))}$ on $X \times \mathbb{R}_+$. The assertion $(r,v) \in \hat{\mathfrak{D}}^*$ is equivalent to the assertion that $(m^n, n \in \mathbb{N})$ is a Cauchy sequence with respect to the Prohorov metric on $X \times \mathbb{R}_+$. Hence,

$$\hat{\mathfrak{D}}^* = \bigcap_{\varepsilon \in \mathbb{Q} \cap (0,\infty)} \bigcup_{k \in \mathbb{N}} \bigcap_{\ell \ge n \ge k} \hat{\mathfrak{D}}_{\varepsilon,\ell,n}$$

with

$$\hat{\mathfrak{D}}_{\varepsilon,\ell,n} := \left\{(r,v) \in \hat{\mathfrak{D}} : d_{\mathrm{P}}(m^\ell, m^n) < \varepsilon\right\},$$

where m^ℓ and m^n are considered as probability measures on the (finite) support of m^ℓ in $X \times \mathbb{R}_+$. Using the definition (10.3) of the Prohorov metric, we can write

$$\hat{\mathfrak{D}}_{\varepsilon,\ell,n} = \bigcap_{F \subset \{1,\dots,\ell\}} \left\{\frac{1}{\ell} \sum_{i=1}^{\ell} \mathbb{1}\{\exists j \in F \text{ with } r(i,j) \vee |v(i) - v(j)| = 0\}\right.$$

$$\left. < \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}\{\exists j \in F \text{ with } r(i,j) \vee |v(i) - v(j)| < \varepsilon\} + \varepsilon\right\}. \qquad \square$$

10.3 Resampling from marked metric measure spaces

We will use the statements from this section to prove assertions (ii) and (iii) of Theorem 3.9. In the end of this section, we also prove Proposition 3.11 from Section 3.4.

The following proposition can be compared with Lemma 8 of Vershik. We construct a marked metric measure space from a marked distance matrix. When we sample according to its marked distance matrix distribution, the assertion is that we arrive at a random variable that has the same distribution as the marked distance matrix with which we started. Recall the functions $\hat{\psi}$, ψ and the sets $\hat{\mathfrak{D}}^*$, $\mathfrak{D}$ from Section 3.3.

Proposition 10.3. *Let (r, v) be an exchangeable random variable with values in $\hat{\mathfrak{D}}^*$. Let (r', v') be a random variable with values in $\mathfrak{D} \times \mathbb{R}_+^{\mathbb{N}}$ and conditional distribution $\nu^{\hat{\psi}(r,v)}$ given $\hat{\psi}(r,v)$. Then (r', v') and (r, v) are equal in distribution.*

Remark 10.4. For an exchangeable random variable with values in $\mathfrak{D}^*$ and a random variable ρ' with conditional distribution $\nu^{\psi(\rho)}$ given $\psi(\rho)$, the random variables ρ and ρ' are equal in distribution. This follows from Proposition 10.3, we set $(r, v) = (\rho, 0)$.

Proof of Proposition 10.3. Let $n \in \mathbb{N}$ and let $\phi : \mathbb{R}_+^{n^2} \times \mathbb{R}_+^n \to \mathbb{R}$ be bounded and continuous. Let (X, r, m) be the representative of $\hat{\psi}(r, v)$ as in the definition of $\hat{\psi}$. We have

$$
\begin{aligned}
\mathbb{E}\left[\phi \circ \gamma_n(r', v')\right] &= \mathbb{E}\left[\int m^{\otimes n}(dx\, dv'')\phi((r(x(i), x(j)))_{i,j\in[n]}, (v''(i))_{i\in[n]})\right] \\
&= \lim_{k\to\infty} \frac{1}{k}\sum_{\ell_1=1}^k \cdots \frac{1}{k}\sum_{\ell_n=1}^k \mathbb{E}\left[\phi((r(\ell_i, \ell_j))_{i,j\in[n]}, (v(\ell_i))_{i\in[n]})\right] \\
&= \mathbb{E}\left[\phi \circ \gamma_n(r, v)\right].
\end{aligned}
$$

Here the assumption $(r, v) \in \hat{\mathfrak{D}}^*$ ensures that m is the weak limit of the uniform probability measures $\frac{1}{k}\sum_{\ell=1}^k \delta_{(\ell, v(\ell))}$ on $X \times \mathbb{R}_+$. This yields the second equality by dominated convergence. For the third equality, we use that summands where $\ell_1, \ldots, \ell_n$ are not pairwise distinct vanish in the limit, and that for all other summands, the expectation in the second line equals by exchangeability the expectation in the third line. $\qquad \square$

In the next proposition, we start with a marked metric measure space and sample (r, v) according to its marked distance matrix distribution. The marked metric measure space that we construct from any typical realization of (r, v) turns out to be isomorphic to the marked metric measure space with which we started.

Proposition 10.5. *Let $\chi \in \hat{\mathbb{M}}$ and let (r, v) be a $\mathfrak{D} \times \mathbb{R}_+^{\mathbb{N}}$-valued random variable with distribution ν^χ. Then $(r, v) \in \hat{\mathfrak{D}}^*$ a. s. and $\hat{\psi}(r, v) = \chi$ a. s.*

Remark 10.6. Proposition 10.5 is essentially Vershik's proof [51, Theorem 4] of the Gromov reconstruction theorem (where metric measure spaces are considered, cf. also [12, Theorem 1] for marked metric measure spaces). The present formulation focuses on the map $\hat{\psi}$ that will be used in the proofs of Theorems 3.9(iii) and 4.1 below.

Remark 10.7. For $\chi \in \mathbb{M}$ and a $\mathfrak{D}$-valued random variable ρ with distribution ν^χ, Proposition 10.5 implies $\rho \in \mathfrak{D}^*$ a. s. and $\psi(\rho) = \chi$ a. s.

Proof of Proposition 10.5. Let (X', r', m') be a representative of χ. W. l. o. g. we assume that the closed support of the probability measure $m'(\cdot \times \mathbb{R}_+)$ is the whole space X', and that $(r, v) = ((r'(x(i), x(j)))_{i,j\in\mathbb{N}}, v)$ for an m'-iid sequence (x, v). We denote by (X, r) the completion of $(\mathbb{N}, r)$. We endow $X' \times \mathbb{R}_+$ with the product metric $d^{X'\times\mathbb{R}_+}((x_1', v_1'), (x_2', v_2')) = r'(x_1', x_2') \vee |v_1' - v_2'|$, and analogously $X \times \mathbb{R}_+$. As the sequence $(x(i))_{i\in\mathbb{N}}$ is a. s. dense in X', the isometry that maps $x(i)$ to i for all $i \in \mathbb{N}$ can a. s. be extended to a (surjective) isometry φ from X' to X. An isometry $\hat{\varphi}$ from $X' \times \mathbb{R}_+$ to $X \times \mathbb{R}_+$ is a. s. given by $(x, v') \mapsto (\varphi(x), v')$. By the Glivenko-Cantelli theorem, the probability measures $m'^n := n^{-1}\sum_{i=1}^n \delta_{(x(i), v(i))}$ on $X' \times \mathbb{R}_+$ converge weakly to m' a. s. As $\hat{\varphi}$ is continuous, the probability measures $m^n := n^{-1}\sum_{i=1}^n \delta_{(i, v(i))} = \hat{\varphi}(m'^n)$ on $X \times \mathbb{R}_+$ converge weakly to $m := \hat{\varphi}(m')$ a. s. This implies $(r, v) \in \mathfrak{D}^*$ a. s. and that $\hat{\psi}(r, v)$ equals the isomorphy class of (X, r, m) a. s. The second assertion follows as $\hat{\varphi}$ is a. s. a measure-preserving isometry from $X' \times \mathbb{R}_+$ to $X \times \mathbb{R}_+$, which implies that (X', r', m') and (X, r, m) have a. s. the same marked distance matrix distribution. $\qquad \square$

Remark 10.8 (Marked metric measure spaces and weighted real trees)**.** Let $\chi \in \hat{\mathbb{U}}$, and let (r, v) be a $\hat{\mathfrak{U}}$-valued random variable with the marked distance matrix distribution of χ. By Proposition 10.5, we have $(r, v) \in \hat{\mathfrak{D}}^*$ a. s., hence we can associate with any typical realization of (r, v) a complete and separable weighted real tree $(\bar{T}, d, \mu)$ as in Remark 3.8. As in Proposition 3.14, the random marked distance matrix (r, v) is ergodic with respect to the action of the group of finite permutations. This yields that the measure-preserving isometry class of the weighted real tree $(\bar{T}, d, \mu)$ is an a. s. constant random variable. Its typical realization can be associated with χ.

Proof of Proposition 3.11. Let (r, v) be a random variable with conditional distribution ν^χ given χ. Then we can assume $\rho = \alpha(r, v)$. Propositions 3.4 and 10.5 imply $\chi = \hat{\psi} \circ \beta(\rho)$ a. s. Hence, the distribution of ρ determines the distribution of χ uniquely, which is the "only if" assertion. The other direction clearly holds as the distribution of χ determines the distribution of ρ uniquely. $\qquad\square$

10.4 Proof of the sampling representation

We give two proofs of Theorem 3.9(i) from Section 3.4 that build on a common part, namely statement (10.7) below. The plan for the first proof is the following: We partition the completion of the tree (T, d) associated with the semi-ultrametric ρ (as in Remark 1.1) into small subsets. Into each of these subsets, we lay an atom whose mass is given by the asymptotic frequency of those integers that label the leaves of T that are the endpoints of the external branches that begin in this subset. By exchangeability, these asymptotic frequencies exist, and (10.7) yields that they add up to one. We obtain an atomic probability measure on the product space of the metric completion of the tree and the mark space $\mathbb{R}_+$ by defining the $\mathbb{R}_+$-component as the distance to the top of the coalescent tree. Using the coupling characterization (10.4) of the Prohorov metric, we show that this probability measure converges as the subsets become infinitely small, and that the limit measure coincides with the limit of the uniform measures in the definition of $\hat{\mathfrak{D}}^*$.

As a slight difference to the description in the preceding paragraph, we will work with the space (X, r) that corresponds to the completion of the space only of the starting vertices of the external branches, but we will occasionally recall the relation to the whole tree. We will use definitions also from Section 2.

Proof of Theorem 3.9. Let $(r, v) = \beta(\rho)$. Then $v = \Upsilon(\rho)$ by definition of the map β. Let (X, r) be the metric completion of the semi-metric space $(\mathbb{N}, r)$.

Let $\varepsilon > 0$. As the distribution of the random variable $v(i)$ has at most countably many atoms, there exists a deterministic sequence $0 < h_1^{(\varepsilon)} < h_2^{(\varepsilon)} < \dots$ that increases to infinity and that satisfies
$$h_1^{(\varepsilon)} < \varepsilon, \quad h_{n+1}^{(\varepsilon)} - h_n^{(\varepsilon)} < \varepsilon,$$
and
$$\mathbb{P}\left(v(i) = h_n^{(\varepsilon)}\right) = 0 \tag{10.5}$$
for all $i, j, n \in \mathbb{N}$. We set $h_0^{(\varepsilon)} = 0$ and we write $I_n^\varepsilon = [h_{n-1}^{(\varepsilon)}, h_n^{(\varepsilon)})$ for $n \in \mathbb{N}$.

We define an equivalence relation $\sim^\varepsilon$ on $\mathbb{N}$ such that two distinct integers i, j are equivalent if and only if there exists $n \in \mathbb{N}$ with
$$v(i), v(j), \tfrac{1}{2}\rho(i, j) \in I_n^\varepsilon.$$

To show transitivity, we consider $i, j, k \in \mathbb{N}$ with $i \neq k$, $i \sim^\varepsilon j$, and $j \sim^\varepsilon k$. Then there exists $n \in \mathbb{N}$ with $v(i), v(j), v(k), \rho(i, j)/2, \rho(j, k)/2 \in I_n^\varepsilon$. As
$$v(i) \le \rho(i, k)/2 \le (\rho(i, j) \vee \rho(j, k))/2$$
by definition of Υ and ultrametricity, it follows that $i \sim^\varepsilon k$.

Note that the definitions in Section 2 imply

$$r(i,j) = \left(\tfrac{1}{2}\rho(i,j) - v(i) + \tfrac{1}{2}\rho(i,j) - v(j) \right) \mathbb{1}\{i \neq j\} < 2\varepsilon \tag{10.6}$$

for $i \sim^\varepsilon j$. (That is, in the context of Remark 2.2, the starting points of external branches that end in leaves $(0,i), (0,j)$ of T with $i \sim^\varepsilon j$ have distance smaller than 2ε.)

In the next two paragraphs, we prove the following claim:

$$\text{A. s., the partition of } \mathbb{N} \text{ given by } \sim^\varepsilon \text{ contains no singleton blocks.} \tag{10.7}$$

For each $i, n \in \mathbb{N}$ the sequence $(\mathbb{1}\{v(j) \in I_n^\varepsilon, \rho(i,j)/2 \in I_n^\varepsilon\}, j \in \mathbb{N} \setminus \{i\})$ is exchangeable. By the de Finetti theorem, it is conditionally iid. Hence, on the event that there exists $j \in \mathbb{N} \setminus \{i\}$ with $v(j) \in I_n^\varepsilon$ and $\rho(i,j)/2 \in I_n^\varepsilon$, there exists a. s. another (in fact, infinitely many) such j in $\mathbb{N} \setminus \{i\}$.

For $j \in \mathbb{N}$, the definition of Υ and condition (10.5) imply the existence of (random) $n \in \mathbb{N}$ and $i \in \mathbb{N} \setminus \{j\}$ such that $v(j) \in I_n^\varepsilon$ and $\rho(i,j)/2 \in I_n^\varepsilon$ a. s. As shown in the preceding paragraph, there exists a. s. an integer $k \in \mathbb{N} \setminus \{i,j\}$ with $v(k) \in I_n^\varepsilon$ and $\rho(i,k)/2 \in I_n^\varepsilon$. From

$$v(k) \leq \rho(j,k)/2 \leq (\rho(i,j) \vee \rho(i,k))/2,$$

it follows that $\rho(j,k)/2 \in I_n^\varepsilon$ a. s. This proves (10.7).

Now we show that the asymptotic frequencies exist and add up to one. For $A \subset \mathbb{N}$ and $k \in \mathbb{N}$, we denote the relative frequency by $|A|_k = k^{-1}\#(A \cap [k])$ and the asymptotic frequency by $|A| = \lim_{k \to \infty} |A|_k$, provided the limit exists. As the random partition given by $\sim^\varepsilon$ is exchangeable, the asymptotic frequencies of its blocks exist a. s. by Kingman's correspondence. Let $B^\varepsilon(i)$ denote the equivalence class of $i \in \mathbb{N}$ with respect to $\sim^\varepsilon$, and let

$$M^\varepsilon = \{j \in \mathbb{N} : j = \min B^\varepsilon(i) \text{ for some } i \in \mathbb{N}\}$$

be the set of minimal elements of the equivalence classes of $\sim^\varepsilon$. As the exchangeable partition given by $\sim^\varepsilon$ has no singleton blocks a. s., it has proper frequencies by Kingman's correspondence, that is,

$$\sum_{i \in M^\varepsilon} |B^\varepsilon(i)| = 1 \quad \text{a. s.}$$

Consequently, on an event of probability 1, a probability measure m^ε on the product sigma algebra on $X \times \mathbb{R}_+$ is given by

$$m^\varepsilon = \sum_{i \in M^\varepsilon} |B^\varepsilon(i)| \delta_{(i,v(i))}. \tag{10.8}$$

(Into each of the subsets of (X,r) given by $\sim^\varepsilon$, the first component of the measure m^ε lays an atom with mass given by the asymptotic frequency of the integers that label the corresponding leaves in T.)

Let $\varepsilon_1 > \varepsilon_2 > \ldots > 0$ with $\lim_{\ell \to \infty} \varepsilon_\ell = 0$. For each $\ell \in \mathbb{N}$, we replace ε with ε_ℓ everywhere in this proof until now, and we use the notations introduced so far. We also assume that for $k \leq \ell$, the sequence $(h_n^{(\varepsilon_k)}, n \in \mathbb{N})$ is contained in $(h_n^{(\varepsilon_\ell)}, n \in \mathbb{N})$. That is, the partitions $\{I_n^{\varepsilon_k}, n \in \mathbb{N}\}$ of $\mathbb{R}_+$ are nested.

For $k \leq \ell$ and $i \in M^{\varepsilon_k}$, let $i_1, i_2, \ldots$ be an enumeration of $M^{\varepsilon_\ell} \cap B^{\varepsilon_k}(i)$. Then

$$B^{\varepsilon_k}(i) = B^{\varepsilon_\ell}(i_1) \uplus B^{\varepsilon_\ell}(i_2) \uplus \ldots.$$

By Fatou's lemma and as a. s., the partition given by $\sim^{\varepsilon_\ell}$ has proper frequencies, it follows that

$$|B^{\varepsilon_k}(i)| = |B^{\varepsilon_\ell}(i_1)| + |B^{\varepsilon_\ell}(i_2)| + \ldots \quad \text{a. s.}$$

Using equation (10.8), we deduce

$$m^{\varepsilon_k}\{(i, v(i))\} = \sum_{j \in B^{\varepsilon_k}(i)} m^{\varepsilon_\ell}\{(j, v(j))\} \quad \text{a.s.} \tag{10.9}$$

A.s., a coupling of m^{ε_k} and m^{ε_ℓ} is given by the probability measure

$$K = \sum_{(i,j)} m^{\varepsilon_\ell}\{(j, v(j))\} \delta_{((i,v(i)),(j,v(j)))} \tag{10.10}$$

on $(X \times \mathbb{R}_+)^2$, where the sum is over all pairs (i, j) with $i \in M^{\varepsilon_k}$ and $j \in M^{\varepsilon_\ell} \cap B^{\varepsilon_k}(i)$. Indeed, as equation (10.9) implies

$$K(\{(i, v(i))\} \times (X \times \mathbb{R}_+)) = \sum_{j \in B^{\varepsilon_k}(i)} m^{\varepsilon_\ell}\{(j, v(j))\} = m^{\varepsilon_k}\{(i, v(i))\} \quad \text{a.s.},$$

K is a.s. a coupling of m^{ε_k} and m^{ε_ℓ}.

   In words, the probability measure m^{ε_ℓ} can be obtained by splitting each atom of m^{ε_k} into fragments. Let us sample a point $(j, v(j))$ according to m^{ε_ℓ}, and let $(i, v(i))$ be the point such that the atom of m^{ε_ℓ} at $(j, v(j))$ is one of the fragments of the atom of m^{ε_k} at $(i, v(i))$. Then the pair $((i, v(i)), (j, v(j)))$ has distribution K.

   For every pair (i, j) that appears in the sum in equation (10.10), we have $i \sim^{\varepsilon_k} j$, hence $|v(i) - v(j)| < \varepsilon_k$ and $r(i, j) < 2\varepsilon_k$. Hence, the coupling characterization of the Prohorov metric (10.4) yields

$$d_{\mathrm{P}}^{X \times \mathbb{R}_+}(m^{\varepsilon_k}, m^{\varepsilon_\ell}) \leq 2\varepsilon_k \tag{10.11}$$

a.s. for all $k \leq \ell$, when $X \times \mathbb{R}_+$ is endowed with the product metric $d^{X \times \mathbb{R}_+}$ that is given by $d^{X \times \mathbb{R}_+}((x, v), (x', v')) = r(x, x') \vee |v - v'|$. As a consequence, on an event of probability 1, the sequence $(m^{\varepsilon_\ell}, \ell \in \mathbb{N})$ in the space of probability measures on the complete space $X \times \mathbb{R}_+$ is Cauchy, we denote its limit by m.

   Consider for $n, \ell \in \mathbb{N}$ also the probability measure $m_n^{\varepsilon_\ell}$ on $X \times \mathbb{R}_+$, given by

$$m_n^{\varepsilon_\ell} = \sum_{i \in M^{\varepsilon_\ell}} |B^{\varepsilon_\ell}(i)|_n \delta_{(i,v(i))} \quad \text{a.s.}$$

As there exists a.s. a coupling K' of the probability measures $m_n^{\varepsilon_\ell}$ and m^{ε_ℓ} with

$$K'\{(y, y)\} = m_n^{\varepsilon_\ell}\{y\} \wedge m^{\varepsilon_\ell}\{y\}$$

for all $y \in X \times \mathbb{R}_+$, the coupling characterization of the Prohorov metric (10.4) implies for each $k \in \mathbb{N}$

$$d_{\mathrm{P}}^{X \times \mathbb{R}_+}(m_n^{\varepsilon_\ell}, m^{\varepsilon_\ell})$$
$$\leq K'\{(y, y') \in (X \times \mathbb{R}_+)^2 : y \neq y'\}$$
$$\leq m^{\varepsilon_\ell}\{(j, v(j)) : j \in M^{\varepsilon_\ell}, j > k\} + K'\{((i, v(i)), (j, v(j))) : i, j \in M^{\varepsilon_\ell}, i \neq j, j \leq k\}$$
$$= \sum_{\substack{j \in M^{\varepsilon_\ell} \\ j > k}} |B^{\varepsilon_\ell}(j)| + \sum_{\substack{j \in M^{\varepsilon_\ell} \\ j \leq k}} \left| |B^{\varepsilon_\ell}(j)|_n - |B^{\varepsilon_\ell}(j)| \right| \quad \text{a.s.}$$

Letting first n and then k tend to infinity, we deduce

$$\lim_{n \to \infty} d_{\mathrm{P}}^{X \times \mathbb{R}_+}(m_n^{\varepsilon_\ell}, m^{\varepsilon_\ell}) = 0 \quad \text{a.s.} \tag{10.12}$$

   Moreover, we define for each $n \in \mathbb{N}$ the probability measure

$$m_n = n^{-1} \sum_{i=1}^{n} \delta_{(i,v(i))}$$

on $X \times \mathbb{R}_+$. (The first component corresponds to a probability measure on the starting vertices of the external branches that end in one of the first n leaves of T. These starting

vertices are weighted according to the number of the corresponding leaves i, where we count leaves i, j with $\rho(i, j) = 0$ as separate leaves.) By (10.11),

$$d_{\mathrm{P}}^{X \times \mathbb{R}_+}(m_n^{\varepsilon_\ell}, m_n) \leq 4\varepsilon_\ell \quad \text{a.s.} \tag{10.13}$$

for all $n, \ell \in \mathbb{N}$. From (10.11), (10.12), and (10.13), we obtain

$$m = \text{w-}\lim_{n \to \infty} m_n \quad \text{a.s.} \tag{10.14}$$

This shows assertion (i).

Assertion (ii) follows from assertion (i) and Proposition 10.3. Proposition 10.5 implies assertion (iii). □

The idea for the second proof of Theorem 3.9(i) is to construct directly by the de Finetti theorem a sampling measure on a subspace of the metric completion of the coalescent tree associated with ρ. To this aim, we fix by conditioning the closure of the subspace of the starting vertices of the external branches that end in the leaves labeled by the odd integers. By (10.7), this subspace contains a.s. the sequence of the starting vertices of the external branches associated with the even integers, and this sequence is exchangeable. For a related result, see also Forman, Haulk, and Pitman [23], where trees are embedded into ℓ_1.

Remark 10.9. The second proof given below goes in a direction that is similar to the argument in Section 7 of [22] for the construction of the sampling measure μ on the real tree $\mathbf{S} = \Gamma(\mathbf{T})$. That the equality $\Gamma(\mathbf{T}) = \Gamma(\mathbf{T}^-) = \Gamma(\mathbf{T}^+)$ on p. 268 in [22] holds for the embedding of $\Gamma(\mathbf{T}^-)$ and $\Gamma(\mathbf{T}^+)$ into $\Gamma(\mathbf{T})$ can be seen from (10.7) as in the proof below as $\Gamma(\mathbf{T})$, $\Gamma(\mathbf{T}^-)$, and $\Gamma(\mathbf{T}^+)$ then correspond to X, X_1, and X_2 therein. The real tree $\Gamma(\mathbf{T}^-)$ can then be endowed with a measure like X_1 is endowed with μ^1. Note that the starting vertices of the external branches and the subtree spanned by them are called the points of attachment and the core, respectively, in [22].

We remark that the second last paragraph of the proof below shows that the isomorphy class of the weighted real tree $(\mathbf{S}, \mu)$ is a.s. equal to $\psi(r)$ where $(r, v) = \beta(d)$ and d is the exchangeable ultrametric on $\mathbb{N}$ from [22, Section 7], which corresponds to ρ below. This equality can also be deduced from Theorem 3.9, Remark 3.6, as $\psi(r)$ is a.s. constant by the ergodicity assumption in [22], and from the Gromov reconstruction theorem.

Second proof of Theorem 3.9(i). Let $(r, v) = \beta(\rho)$. We construct the first component of the sampling measure, showing $r \in \mathfrak{D}^*$ a.s.

We denote by $\mathbb{N}_1$ the odd, and by $\mathbb{N}_2$ the even integers. Let (X, r) denote the metric completion of $(\mathbb{N}, r)$. A.s. by (10.6) and (10.7), there exists for each $i \in \mathbb{N}_2$ an integer $j \in \mathbb{N}_1$ with $r(i, j) < 2\varepsilon$. As ε can be chosen arbitrarily small, it follows that i is a.s. contained in the closure X_1 of the subset $\mathbb{N}_1$ of (X, r) a.s., hence $X_1 = X$ a.s. (Recall from Remark 2.2 that $\mathbb{N}$ corresponds here to the set of starting vertices of the external branches in the coalescent tree (T, d) associated with ρ.)

For $i \in \mathbb{N}_1$, let

$$v^1(i) = \tfrac{1}{2} \inf_{j \in \mathbb{N}_1 \setminus \{i\}} \rho(i, j).$$

(This is the length of the external branch that ends in the leaf $(0, i)$ in the subtree spanned by the leaves with labels in $\mathbb{N}_1$.) By exchangeability of the sequence $(\rho(i, j) : j \in \mathbb{N} \setminus \{i\})$ and by definition of $v = \Upsilon(\rho)$, it follows that $v^1(i) = v(i)$ a.s. Let $\rho^1 = (\rho(i, j))_{i,j \in \mathbb{N}_1}$ be the restriction of ρ to $\mathbb{N}_1$. We define the random variable $r^1 = (r^1(i, j))_{i,j \in \mathbb{N}_1}$ by

$$r^1(i, j) = \left(\rho^1(i, j) - v^1(i) - v^1(j) \right) \mathbb{1}\{i \neq j\}.$$

By definition of r in Section 2, it follows that $r^1 = (r(i, j))_{i,j \in \mathbb{N}_1}$ a.s.

Let Λ be a regular conditional distribution of ρ given ρ^1. Then for a. a. ρ^1, under $\Lambda(\rho^1, \cdot)$, the complete and separable metric space (X_1, r) is a. s. constant as r^1 is ρ^1-measurable.

Moreover, the sequence $2, 4, 6, \ldots$ of the even integers, viewed as a sequence in (X_1, r), is exchangeable under $\Lambda(\rho^1, \cdot)$ for a. a. ρ^1. To see this, we use that the Borel sigma algebra on (X_1, r) is generated by the balls around the elements of $\mathbb{N}_1 \subset X_1$. Let $n \in \mathbb{N}$, and let $B_2, \ldots, B_{2n}$ be some finite intersections of such balls. Note that $\{2 \in B_2, \ldots, 2n \in B_{2n}\}$ can be written as an intersection of events of the form $\{\rho(i, j) < c\}$, where $i \in \mathbb{N}_2$, $j \in \mathbb{N}_1$ and $c \in (0, \infty)$. Using this, the uniqueness lemma, and the elementary fact that the conditional distribution of ρ given its restriction ρ^1 is invariant under permutations that leave $\mathbb{N}_1$ fixed, we obtain the claimed exchangeability.

For this exchangeable sequence, the de Finetti theorem yields, $\Lambda(\rho^1, \cdot)$-a. s. for a. a. ρ^1, a sampling measure μ^1 on (X_1, r) that is the weak limit of the probability measures $\mu_n^1 := n^{-1} \sum_{i=1}^{n} \delta_{2i}$ on (X_1, r). By the same argument as above, also the closure X_2 of the subset $\mathbb{N}_2$ in (X, r) equals X a. s. On the event of probability 1 on which $\mathbb{N}_2$ is a dense subset of $X_2 = X = X_1$, an isometry $\varphi : X_1 \to X_2$ is given by $\varphi(i) = i$ for $i \in \mathbb{N}_2$. As also the weak limit of the image measures $\varphi(\mu_n^1)$ on (X_2, r) exists a. s., we have shown $(r(2i, 2j))_{i,j \in \mathbb{N}} \in \mathfrak{D}^*$ a. s. This implies $r \in \mathfrak{D}^*$ a. s. as r and $(r(2i, 2j))_{i,j \in \mathbb{N}}$ are equal in distribution by exchangeability of r.

That $(r, v) \in \hat{\mathfrak{D}}^*$ can be shown analogously by considering the sequence $(i, v(i))_{i \in \mathbb{N}_2}$ in the space $X_1 \times \mathbb{R}_+$ which we endow with the metric $d^{X_1 \times \mathbb{R}_+}((x', v'), (x'', v'')) = r(x', x'') \vee |v' - v''|$. □

10.5 Proof of Theorem 4.1

The following property is central in the proof of Theorem 4.1.

Proposition 10.10. *Let $t \in \mathbb{R}_+$, and let $f : \mathfrak{U} \to \mathbb{R}$, $g : \hat{\mathfrak{U}} \to \mathbb{R}$ be bounded measurable functions. If the assumptions of Theorem 4.1 hold, then*

$$\mathbb{E}[g(\beta(\rho_t))] = \mathbb{E}[\nu^{\hat{\chi}_t} g]$$

and

$$\mathbb{E}[f(\rho_t)] = \mathbb{E}[\xi_t f].$$

If the assumptions of Theorem 4.1 hold and ρ_t is a. s. dust-free, then

$$\mathbb{E}[f(\rho_t)] = \mathbb{E}[\nu^{\chi_t} f].$$

Proof. This is immediate from Theorem 3.9, the definition of ξ_t, and Corollary 3.12. □

Remark 10.11. In the context of Theorem 4.1(i), let $(P_t, t \in \mathbb{R}_+)$ denote the semigroup on $M_b(\hat{\mathfrak{U}})$ of the Markov process $(\beta(\rho_t), t \in \mathbb{R}_+)$, and let $(Q_t, t \in \mathbb{R}_+)$ denote the semigroup on $M_b(\hat{\mathbb{U}})$ of the Markov process $(\hat{\chi}_t, t \in \mathbb{R}_+)$. Let K denote the probability kernel from $\hat{\mathbb{U}}$ to $\hat{\mathfrak{U}}$, given by $K(\chi, \cdot) = \nu^\chi$ for $\chi \in \hat{\mathbb{U}}$. Then Proposition 10.10 yields the intertwining relation $Q_t K = K P_t$ which is condition (b) in [46, Theorem 2]. Many papers appeared on intertwining of Markov processes, a classical one is for instance [9].

Proof of Theorem 4.1. We apply [46, Theorem 2] to the semigroup of the Markov process $(\beta(\rho_t), t \in \mathbb{R}_+)$, the measurable map $\hat{\psi} : \hat{\mathfrak{U}} \to \hat{\mathbb{U}}$, and the kernel K from $\hat{\mathbb{U}}$ to $\hat{\mathfrak{U}}$ given by $K(\chi, \cdot) = \nu^\chi$. Clearly, Theorem 2 in [46] also holds when the initial state y therein is random. Then by Proposition 10.10, condition (b) and the condition on the initial state in [46, Theorem 2] are satisfied. Condition (a) in [46, Theorem 2] follows from Proposition 10.5 as $f(\chi) = \nu^\chi(f \circ \hat{\psi})$ for all $\chi \in \hat{\mathbb{U}}$ and all bounded measurable $f : \hat{\mathbb{U}} \to \mathbb{R}$. The Markov property of $(\hat{\chi}_t, t \in \mathbb{R}_+)$ now follows from [46, Theorem 2].

Now we give a proof that $(\hat{\chi}_t, t \in \mathbb{R}_+)$ solves the martingale problem $(\hat{B}, \hat{\Pi})$. If $(\beta(\rho_t), t \in \mathbb{R}_+)$ solves the martingale problem $(\hat{A}, \hat{\mathcal{C}})$ in (i), then for all $\phi \in \hat{\mathcal{C}}$ with associated marked polynomial Φ,

$$
\begin{aligned}
0 =&\, \mathbb{E}\left[\phi(\beta(\rho_t)) - \phi(\beta(\rho_0)) - \int_0^t \hat{A}\phi(\beta(\rho_u))du\right] \\
=&\, \mathbb{E}[\nu^{\hat{\chi}_t}\phi] - \mathbb{E}[\nu^{\hat{\chi}_0}\phi] - \int_0^t \mathbb{E}[\nu^{\hat{\chi}_u}(\hat{A}\phi)]du \\
=&\, \mathbb{E}\left[\Phi(\hat{\chi}_t) - \Phi(\hat{\chi}_0) - \int_0^t \hat{B}\Phi(\hat{\chi}_u)du\right]
\end{aligned}
\tag{10.15}
$$

by Proposition 10.10, Fubini, and the definitions Φ and $\hat{B}$. By the Markov property of $(\chi_s, s \in \mathbb{R}_+)$ and equation (10.15), it now follows for all $s \in \mathbb{R}_+$ and all $(\hat{\chi}_u, u \in [0, s])$-measurable events A that

$$
\mathbb{E}\left[\Phi(\hat{\chi}_{s+t}) - \Phi(\hat{\chi}_s) - \int_s^{s+t} B\Phi(\hat{\chi}_u)du; A\right] = 0
$$

which shows assertion (i).

The proof of (ii) is analogous, we apply [46, Theorem 2] to the Markov process $(\rho_t, t \in \mathbb{R}_+)$, the measurable map $\mathfrak{U} \to \mathcal{U}^{\text{erg}}$, $\rho \mapsto \alpha(\nu^{\hat{\psi}(\beta(\rho))})$, and the probability kernel from $\mathcal{U}^{\text{erg}}$ to $\mathfrak{U}$ given by $(\nu, B) \mapsto \nu(B)$. In particular, condition (a) in [46, Theorem 2] is satisfied by Propositions 10.5 and 3.4, and by definition of $\mathcal{U}^{\text{erg}}$.

Also the proof of (iii) is analogous. We apply [46, Theorem 2] to the process $(\rho_t, t \in \mathbb{R}_+)$, the measurable map $\psi : \mathfrak{U} \to \mathbb{U}$, and the probability kernel from $\mathbb{U}$ to $\mathfrak{U}$ given by $(\chi, B) \mapsto \nu^\chi(B)$. We use the assumption that ρ_t is a.s. dust-free in the application of Proposition 10.10 and Remark 10.7. $\qquad\square$

11 Proofs related to the lookdown model

This section contains the remaining proofs of the results from Sections 5 – 8.

11.1 Exchangeability in the lookdown model

To prove Proposition 5.8, we show in Lemma 11.1 below that exchangeability of the genealogical distances is preserved in single reproduction events. Then we construct the genealogical distance matrix ρ_t at time t, restricted to the first $n \in \mathbb{N}$ particles, from the initial state ρ_0 and the reproduction events before time t that affect the genealogical distances between the first n individuals. Here we use the description of the process $(\gamma_n(\rho_t), t \in \mathbb{R}_+)$ by its jumps and the evolution between the jumps from the end of Section 5.1.

For $n \in \mathbb{N}$, we define the action of the group S_n of permutations of $[n]$ on the set $\mathcal{P}_n$ of partitions of $[n]$, and on $\mathbb{R}^{n^2}$, respectively, by

$$
p(\pi) = \{p(B) : B \in \pi\} \quad \text{and} \quad p(\rho) = (\rho(p(i), p(j)))_{i,j\in[n]}
\tag{11.1}
$$

for each $p \in S_n$, $\pi \in \mathcal{P}_n$, $\rho \in \mathbb{R}^{n^2}$. A random variable with values for instance in $\mathcal{P}_n$ or in $\mathbb{R}^{n^2}$ is called exchangeable if its distribution is invariant under the action of S_n.

Lemma 11.1. *Let $n \in \mathbb{N}$, let π be an exchangeable random partition of $[n]$, and let ρ be an exchangeable random variable with values in $\mathbb{R}^{n^2}$. Assume that π and ρ are independent. Then the random variable $\pi(\rho)$ is exchangeable.*

Lemma 11.1 can be seen as a generalization of Lemma 4.3 of Bertoin [4].

Proof. Let $p \in S_n$. For each partition $\pi' \in \mathcal{P}_n$, the blocks of π' are in one-to-one correspondence with the blocks of $p(\pi')$ via the bijection that maps a block $B \in \pi'$ to the block $p(B) \in p(\pi')$. Also, the blocks of π' are in one-to-one correspondence with the integers in $[n]$ that are the minimal elements of the blocks of π'. The same holds for the blocks of $p(\pi')$ and their minimal elements. It follows that the minimal elements of the blocks of π' are in one-to-one correspondence with the minimal elements of the blocks of $p(\pi')$. We extend this one-to-one correspondence arbitrarily to a bijection from $[n]$ to itself which we denote by $f(\pi')$. This defines a map $f : \mathcal{P}_n \to S_n$ which satisfies

$$\pi'(i) = f(\pi')\left(p(\pi')(p(i))\right) \tag{11.2}$$

for all $\pi' \in \mathcal{P}_n$ and $i \in [n]$. This equation holds as $\pi'(i)$, by its definition in Section 5.1, is a minimal element of a block of π' and as $p(\pi')(p(i))$ is the minimal element of the corresponding block of π'. By equation (11.1) and the definition (5.2) of the transformation on $\mathbb{R}^{n^2}$ associated with each element of $\mathcal{P}_n$, equation (11.2) implies

$$\pi'(\rho') = p\left(p(\pi')\left(f(\pi')(\rho')\right)\right) \tag{11.3}$$

for all $\pi' \in \mathcal{P}_n$ and $\rho' \in \mathbb{R}^{n^2}$.

By assumption, $p(\pi)$ and π are equal in distribution. As the distribution of $f(\pi')(\rho)$ is the same for all $\pi' \in \mathcal{P}_n$, namely equal to the distribution of ρ, it follows that $f(\pi)(\rho)$ and π are independent, and that $f(\pi)(\rho)$ is equal in distribution to ρ. This implies that $\pi(\rho)$ and $p(\pi)\left(f(\pi)(\rho)\right)$ are equal in distribution as also ρ and π are independent by assumption. By equation (11.3), it follows that $\pi(\rho)$ and $p^{-1}(\pi(\rho))$ are equal in distribution, which yields the assertion. $\qquad \square$

Proof of Proposition 5.8. Let $n \in \mathbb{N}$. For $s \in \mathbb{R}_+$, we define the map

$$\lambda_s : \mathbb{R}^{n^2} \to \mathbb{R}^{n^2}, \quad \rho' \mapsto \rho' + \underline{\underline{2}}_n s,$$

where $\underline{\underline{2}}_n = 2(\mathbb{1}\{i \neq j\})_{i,j \in [n]}$. We will use the map λ_s to account for the linear growth of the genealogical distances between reproduction events.

On an event of probability 1, let $(t_1, \pi_1), (t_2, \pi_2), \ldots$ be the points of η in $(0,t] \times \mathcal{P}^n$. Let $L = \eta((0,t] \times \mathcal{P}^n)$. Conditionally given $(t_1, \ldots, t_L)$, the partitions $\pi_1, \ldots, \pi_L$ are independent and for each $k \in \mathbb{N}$, the restriction $\gamma_n(\pi_k)$ is exchangeable. This follows from the properties of Poisson random measures and the definition of η. From the description around equation (5.3), we have

$$\gamma_n(\rho_t) = \lambda_{t-t_L} \circ \gamma_n(\pi_L) \circ \lambda_{t_L - t_{L-1}} \circ \ldots \circ \gamma_n(\pi_1) \circ \lambda_{t_1}(\gamma_n(\rho_0)) \quad \text{a.s.}$$

on the event $\{L \geq 1\}$, and $\gamma_n(\rho_t) = \lambda_t(\gamma_n(\rho_0))$ a.s. on $\{L = 0\}$. By assumption, $\gamma_n(\rho_0)$ is exchangeable, and Lemma 11.1 implies that $\gamma_n(\rho_t)$ is exchangeable. The assertion follows as $n \in \mathbb{N}$ was arbitrary and as the distribution of ρ_t is determined by the distributions of the restrictions $\gamma_n(\rho_t)$. $\qquad \square$

11.2 Equality of decompositions

To prove Proposition 6.5, we use the following lemma. Its meaning is that if the ancestral lineage of an individual i at time t can be traced back until a most recent reproduction event on that lineage, then there exists a.s. another individual k at time t that descends from this reproduction event.

Lemma 11.2. *Assume that Ξ is not dust-free. Let $t \in (0, \infty)$ and $i \in \mathbb{N}$. Then a.s. on the event $\{v_t(i) < t\}$, there exists an integer $k \in \mathbb{N} \setminus \{i\}$ with $v_t(i) = \frac{1}{2}\rho_t(i,k)$.*

Proof. Recall the process $(\Pi_s^{(t)}, s \in \mathbb{R}_+)$ from Remark 5.9. We work on the intersection of $\{v_t(i) < t\}$ with the event of probability 1 on which condition (6.2) is satisfied, $v_t(i) > 0$, and for each $s \in (0, t) \cap \mathbb{Q}$, the partition $\Pi_s^{(t)}$ contains infinitely many blocks if it contains singletons. The latter event indeed has probability 1 by Kingman's correspondence and as t is a. s. not the time of a reproduction event.

At time $t - v_i(t)$, a reproduction event occurs that is encoded by a partition in which the block that contains $A_{t-v_t(i)}(t, i)$ contains some other element j. This follows from the definition of $v_t(i)$ in Section 6.1 and as $\eta((0, t] \times \hat{\mathcal{P}}^i) < \infty$ by condition (6.2) which means that the reproduction events in which particles on levels not larger than i reproduce do not accumulate.

Moreover, by condition (6.2), there exists a time $s \in (t - v_i(t), t) \cap \mathbb{Q}$ with $\eta((t - v_t(i), s] \times \hat{\mathcal{P}}_j) = 0$, which implies that the particle on level j at time $t - v_t(i)$ is still on level j at time s.

By definition of $v_t(i)$, the partition $\Pi_s^{(t)}$ contains the singleton block $\{A_s(t, i)\}$, hence $\Pi_s^{(t)}$ has infinitely many blocks. This means that infinitely many particles at time s survive until time t. Remark 5.1 implies that all particles at time s survive until time t. Therefore, the particle that was on level j at the times $t - v_t(i)$ and s is on some level k at time t. The most recent common ancestor of the particles on levels i and k at time t lives at time $t - v_t(i)$, hence $\frac{1}{2}\rho_t(i, k) = v_t(i)$. $\square$

Proof of Proposition 6.5. Let $t \in (0, \infty)$ and $i \in \mathbb{N}$. We have to show that $v_t(i) = \Upsilon(\rho_t)(i)$ a. s.

From the definitions of the reproduction events in Section 5.1 and of the quantity $v_t(i)$ in Section 6.1, it follows that for each $s \in (t - v_t(i) \wedge t, t]$, only the particle on level i at time t descends from the particle on level $A_s(t, i)$ at time s. The definitions of Υ in Section 2 and of ρ_t in Section 5.1 imply $0 \le v_t(i) \wedge t \le \Upsilon(\rho_t)(i) \wedge t$.

In the case that Ξ is dust-free, we have $\Upsilon(\rho_t) = 0$ a. s. by Proposition 5.10, hence also $v_t(i) = 0$ a. s.

Now we assume that Ξ is not dust-free. Lemma 11.2 yields $\Upsilon(\rho_t)(i) \le v_t(i)$ a. s. on the event $\{v_t(i) < t\}$.

We claim that on the event $\{v_t(i) \ge t\}$, all individuals at time 0 have descendants at time t. This can be seen as follows: For each $s \in (0, t)$, the exchangeable partition $\Pi_s^{(t)}$, defined in Remark 5.9, contains the singleton block $\{A_s(t, i)\}$ on the event $\{v_t(i) \ge t\}$. By Kingman's correspondence, it follows that $\Pi_s^{(t)}$ has infinitely many blocks a. s. on $\{v_t(i) \ge t\}$. Using Remark 5.1, we deduce that a. s. on $\{v_t(i) \ge t\}$, all particles at any time $s \in (0, t)$ survive until time t. As condition (5.1) is a. s. satisfied, each individual at time 0 retains its level for a positive time a. s., whence all individuals at time 0 survive until time t a. s.

Hence, as $\Upsilon(\rho_0) = v_0$ by assumption (6.5),

$$v_t(i) = t + v_0(A_0(t, i)) = t + \frac{1}{2} \inf_{j \in \mathbb{N} \setminus \{A_0(t,i)\}} \rho_0(A_0(t, i), j)$$

$$= \frac{1}{2} \inf_{j \in \mathbb{N} \setminus \{i\}} \rho_t(i, j) = \Upsilon(\rho_t)(i) \quad \text{a. s. on } \{v_t(i) \ge t\}. \qquad \square$$

11.3 Uniqueness for the martingale problems for tree-valued Fleming-Viot processes

Proof of Proposition 7.1. We consider the martingale problem (B, Π), the proofs for the other martingale problems are analogous. It remains to show uniqueness of the solution discussed in Section 7. We use a function-valued dual process. This method is applied in the context of tree-valued Fleming-Viot processes in [13], another dual process is used in

[26]. We fix $n \in \mathbb{N}$ and work with a dual process with state space $\mathcal{C}_n$. With each element π of $\mathcal{P}_n$, we also associate a transformation $\mathcal{C}_n \to \mathcal{C}_n$, which we also denote by π, by

$$\pi(\phi)(\rho) = \phi(\pi(\rho)), \quad \rho \in \mathbb{R}^{n^2}, \phi \in \mathcal{C}_n.$$

Here $\pi(\rho)$ is defined in equation (5.2). We define an independent process $(\phi_t, t \in \mathbb{R}_+)$ as the Markov process with càdlàg paths in $\mathcal{C}_n$ such that

- for each $\pi \in \mathcal{P}_n \setminus \{\mathbf{0}_n\}$ at rate λ_π, the process jumps from ϕ to $\pi(\phi)$,

- and between these jumps, the process evolves deterministically according to

$$\phi_{t+s}(\rho) = \phi_t(\rho + \underline{\underline{2}}_n s)$$

for $s, t \in \mathbb{R}_+$ and $\rho \in \mathbb{R}^{n^2}$, where $\underline{\underline{2}}_n = 2(\mathbb{1}\{i \neq j\})_{i,j \in [n]}$.

The process $(\phi_t, t \in \mathbb{R}_+)$ solves the martingale problem $(B^\downarrow, \mathscr{D})$, where

$$\mathscr{D} = \left\{ \mathcal{C}_n \to \mathbb{R}, \phi \mapsto \nu^{\chi'} \phi : \chi' \in \mathbb{U} \right\}$$

and an operator $B^\downarrow$ with domain $\mathscr{D}$ is defined by $B^\downarrow = B^\downarrow_{\text{coal}} + B^\downarrow_{\text{shrink}}$,

$$B^\downarrow_{\text{coal}} \nu^{\chi'}(\phi) = \sum_{\pi \in \mathcal{P}_n \setminus \{\mathbf{0}_n\}} \lambda_\pi \left(\nu^{\chi'}(\pi(\phi)) - \nu^{\chi'} \phi \right)$$

and

$$B^\downarrow_{\text{shrink}} \nu^{\chi'}(\phi) = \nu^{\chi'} \langle \nabla \phi, \underline{\underline{2}} \rangle$$

for $\phi \in \mathcal{C}_n$ and $\chi' \in \mathbb{U}$. Here we use the notation $\langle \nabla \phi, \underline{\underline{2}} \rangle$ from equation (5.6).

From this definition, we have $B(\nu \cdot \phi)(\chi') = B^\downarrow \nu^{\chi'}(\phi)$ for all $\phi \in \mathcal{C}_n$ and $\chi' \in \mathbb{U}$, where $\nu \cdot \phi$ is the polynomial associated with ϕ. For all $t \in \mathbb{R}_+$ and all polynomials $\Phi \in \Pi$ of degree at most n, it follows from Theorem 4.4.11 in [19] that $E[\Phi(\tilde{\chi}_t)]$ is equal for all solutions $((\tilde{\chi}_t, t \in \mathbb{R}_+); P)$ of the martingale problem (B, Π) with initial state χ_0. As $n \in \mathbb{N}$ was arbitrary and the space Π of polynomials is separating, the uniqueness assertion follows from Theorem 4.4.2 in [19]. $\qquad \square$

11.4 Proof of Lemma 8.1

Using the lookdown construction, we show that the semigroup of an $\hat{\mathbb{U}}$-valued Ξ-Fleming-Viot process preserves the set of marked polynomials.

Proof of Lemma 8.1. Let $\mathcal{N}$ denote the space of simple point measures on $(0, \infty) \times \mathcal{P}$. Let $t \in \mathbb{R}_+$ and $n \in \mathbb{N}$. Note that in the construction in Sections 5.1 and 6.1, the restriction $\gamma_n(r_t, v_t)$ depends only on the simple point measure η and the restriction $\gamma_n(r_0, v_0)$ of the initial state. We may thus define the function $g_n : \mathbb{R}^{n^2} \times \mathbb{R}^n \times \mathcal{N} \to \mathbb{R}^{n^2} \times \mathbb{R}^n$ that maps the restriction $\gamma_n(r_0, v_0)$ of the initial state and the point measure η to $\gamma_n(r_t, v_t)$. Note that when the simple point measure is fixed, g_n is a differentiable function on $\mathbb{R}^{n^2} \times \mathbb{R}^n$ with bounded uniformly continuous derivative.

Let $\phi \in \hat{\mathcal{C}}_n$. We define the function

$$f : \mathbb{R}^{n^2} \times \mathbb{R}^n \to \mathbb{R}, \quad (r, v) \mapsto \int \mathbb{P}(\eta \in d\eta') \phi \circ g_n((r, v), \eta'),$$

where η is now the Poisson random measure from Section 5.2. By dominated convergence and the mean value theorem, also the function f is differentiable with bounded uniformly continuous derivative, and we obtain that $f \in \hat{\mathcal{C}}_n$.

Let Φ be the marked polynomial associated with ϕ. For $\chi \in \hat{\mathbb{U}}$, let (r_0, v_0) be a random variable with the marked distance matrix distribution of χ, and let (r_t, v_t) be defined from (r_0, v_0) and the independent Poisson random measure η as in Section 6.2. From Propositions 6.5 and 10.10, and as we may assume that the $\hat{\mathbb{U}}$-valued Ξ-Fleming-Viot process $(\hat{\chi}_s, s \in \mathbb{R}_+)$ from Section 8 satisfies $\hat{\chi}_t = \hat{\psi}(r_t, v_t)$ a.s., we obtain that $\mathbb{E}_\chi[\Phi(\hat{\chi}_t)] = \nu^\chi f$ for all $\chi \in \hat{\mathbb{U}}$. Hence, $\chi \mapsto \mathbb{E}_\chi[\Phi(\hat{\chi}_t)]$ is in $\hat{\Pi}$. □

List of notation

Here we collect notation that is used globally in the article.

Miscellaneous
$\mathbb{R}_+ = [0, \infty)$, $\mathbb{Q}_+ = \mathbb{R}_+ \cap \mathbb{Q}$, $\mathbb{N} = \{1, 2, 3, \ldots\}$, $[n] = \{1, \ldots, n\}$ for $n \in \mathbb{N}$, $[0] = \emptyset$
γ_n: restriction map in various contexts, (p. 12/l. -18, p. 15/ l. 10)
$\xi f = \int \xi(dx) f(x)$ for a measure ξ and a function f (p. 12/l. -12)
$M_{\mathrm{b}}(E)$: set of bounded measurable functions $E \to \mathbb{R}$
$\varphi(\mu) = \mu \circ \varphi^{-1}$: pushforward measure under a measurable function φ (p. 8/l. 1)

(Marked) distance matrices
$\mathfrak{U}$: space of semi-ultrametrics on $\mathbb{N}$, (p. 6/l. -11)
$\hat{\mathfrak{U}}$: space of decomposed semi-ultrametrics on $\mathbb{N}$, (p. 6/l. -4)
$\mathfrak{D}, \hat{\mathfrak{D}}$: spaces of (decomposed) semimetrics on $\mathbb{N}$ (p. 6/l. -10, p. 27/l. -12)
α: map that retrieves the semi-ultrametric from a decomposed semi-ultrametric (p. 6/l. -6)
$\beta : \mathfrak{U} \to \hat{\mathfrak{U}}$: decomposition map into the external branches and the remaining subtree (p. 7/l. 3)
$\Upsilon(\rho)$: vector of the lengths of the external branches in the coalescent tree associated with ρ (p. 7/l. 2)

(Marked) metric measure spaces
$\mathbb{M}$: space of isomorphy classes of metric measure spaces (p. 7/l. -12)
$\mathbb{U}$: space of isomorphy classes of ultrametric measure spaces (p. 8/l. -18)
$\hat{\mathbb{M}}, \hat{\mathbb{U}}$: spaces of isomorphy classes of marked metric measure spaces (p. 7/l. -7, p. 8/l. -15)
ν^χ: distance matrix distribution of $\chi \in \mathbb{M}$ (p. 3/l. -4, p. 7/l. -2) or marked distance matrix distribution of $\chi \in \hat{\mathbb{M}}$ (p. 4/l. 14, p. 7/l. -2)
$\mathcal{U}^{\mathrm{erg}}$: space of distance matrix distributions (p. 11/l. 7)
$\psi : \mathfrak{D} \to \mathbb{M}, \hat{\psi} : \mathfrak{D} \times \mathbb{R}_+^{\mathbb{N}} \to \hat{\mathbb{M}}$: construction of (marked) metric measure spaces (p. 9/l. 9, p. 9/l. 15)
$\mathfrak{D}^*, \hat{\mathfrak{D}}^*$: sets of (marked) distance matrices with a good sampling measure (p. 9/l. 13, p. 9/l. 19)
$\mathcal{C}_n, \mathcal{C}, \hat{\mathcal{C}}_n, \hat{\mathcal{C}}$: sets of bounded differentiable functions with bounded uniformly continuous derivative (p. 12)
Π: set of polynomials on $\mathbb{U}$ (p. 12/l. -5)
$\hat{\Pi}$: set of marked polynomials on $\hat{\mathbb{U}}$ (p. 12/l. -3)
$\mathscr{C}$: a set of test functions on $\mathcal{U}^{\mathrm{erg}}$ (p. 12/l. -1)

Partitions and semi-partitions
$\mathcal{P}$: Set of partitions of $\mathbb{N}$
$B_i(\pi)$: i-th block of a partition π (p. 15/l. 14)
$\#\pi$: number of blocks of a partition π
$K_{i,j}$: partition of $\mathbb{N}$ that contains only $\{i, j\}$ and singleton blocks (p. 17/l. 8)
$\mathcal{P}_n$: Set of partitions of $[n]$, associated transformations (equation (5.2))
$\mathbf{0}_n = \{\{1\}, \ldots, \{n\}\} \in \mathcal{P}_n$
$\mathcal{P}^n$: Set of partitions of $\mathbb{N}$ in which the first n integers are not all in different blocks (p. 15/l. 13)
$\hat{\mathcal{P}}^n$: Set of partitions of $\mathbb{N}$ in which the first n integers are not all in singleton blocks (p. 20/ l. 11)
$\mathcal{S}_n$ set of semi-partitions of $[n]$, associated transformations (p. 19/l. -13, p. 19/l. -6)
$\Delta = \{x = (x_1, x_2, \ldots) : x_1 \geq x_2 \geq \ldots 0, |x|_1 \leq 1\}$
$|x|_p = (\sum_i x_i^p)^{1/p}$ for $x \in \Delta$
$\kappa(x, \cdot)$: paintbox distribution associated with $x \in \Delta$ (p. 17/l. 3)

Genealogy in the lookdown model

η: point measure on $(0, \infty) \times \mathcal{P}$ that encodes the reproduction events, (p. 15/l. 20, p. 17/l. 12)

$A_s(t, i)$: level of the ancestor at time s of the particle on level i at time t (p. 15/l. -17)

$\rho_t(i, j)$: genealogical distance (p. 15/ l. -4)

(r_t, v_t): decomposed genealogical distance (p. 19)

$\Xi = \Xi_0 + \Xi\{0\}\delta_0$, equation (5.4)

H_Ξ: characteristic measure of η (p. 17/l. 10)

References

[1] D. Aldous, *The continuum random tree. III*, Ann. Probab. **21** (1993), no. 1, 248–289. MR-1207226

[2] T. Austin, *Exchangeable random measures*, Ann. Inst. Henri Poincaré Probab. Stat. **51** (2015), no. 3, 842–861. MR-3365963

[3] J. Berestycki, N. Berestycki, and J. Schweinsberg, *Small-time behavior of beta coalescents*, Ann. Inst. Henri Poincaré Probab. Stat. **44** (2008), no. 2, 214–238. MR-2446321

[4] J. Bertoin, *Random fragmentation and coagulation processes*, Cambridge University Press, Cambridge, 2006. MR-2253162

[5] J. Bertoin and J.-F. Le Gall, *Stochastic flows associated to coalescent processes*, Probab. Theory Related Fields **126** (2003), no. 2, 261–288. MR-1990057

[6] J. Bertoin and J.-F. Le Gall, *Stochastic flows associated to coalescent processes. II. Stochastic differential equations*, Ann. Inst. H. Poincaré Probab. Statist. **41** (2005), no. 3, 307–333. MR-2139022

[7] J. Bertoin and J.-F. Le Gall, *Stochastic flows associated to coalescent processes. III. Limit theorems*, Illinois J. Math. **50** (2006), no. 1–4, 147–181 (electronic). MR-2247827

[8] M. Birkner, J. Blath, M. Möhle, M. Steinrücken, and J. Tams, *A modified lookdown construction for the Xi-Fleming-Viot process with mutation and populations with recurrent bottlenecks*, ALEA Lat. Am. J. Probab. Math. Stat. **6** (2009), 25–61. MR-2485878

[9] P. Carmona, F. Petit, and M. Yor, *Beta-gamma random variables and intertwining relations between certain Markov processes*, Rev. Mat. Iberoamericana **14** (1998), no. 2, 311–367. MR-1654531

[10] I. Dahmer and G. Kersting, *The total external length of the evolving Kingman coalescent*, Probab. Theory Related Fields **167** (2017), 1165–1214. MR-3627436

[11] I. Dahmer, R. Knobloch, and A. Wakolbinger, *The Kingman tree length process has infinite quadratic variation*, Electron. Commun. Probab. **19** (2014), no. 87. MR-3298276

[12] A. Depperschmidt, A. Greven, and P. Pfaffelhuber, *Marked metric measure spaces*, Electron. Commun. Probab. **16** (2011), 174–188. MR-2783338

[13] A. Depperschmidt, A. Greven, and P. Pfaffelhuber, *Tree-valued Fleming-Viot dynamics with mutation and selection*, Ann. Appl. Probab. **22** (2012), no. 6, 2560–2615. MR-3024977

[14] A. Depperschmidt, A. Greven, and P. Pfaffelhuber, *Path-properties of the tree-valued Fleming-Viot process*, Electron. J. Probab. **18** (2013), no. 84. MR-3109623

[15] P. Donnelly and T. G. Kurtz, *A countable representation of the Fleming-Viot measure-valued diffusion*, Ann. Probab. **24** (1996), no. 2, 698–742. MR-1404525

[16] P. Donnelly and T. G. Kurtz, *Particle representations for measure-valued population models*, Ann. Probab. **27** (1999), no. 1, 166–205. MR-1681126

[17] E. B. Dynkin, *Markov processes. Vol. I*, Springer-Verlag, 1965. MR-0193671

[18] A. Etheridge and T. G. Kurtz, *Genealogical constructions of population models* (2017). arXiv:1402.6724v3. To appear in Ann. Probab.

[19] S. N. Ethier and T. G. Kurtz, *Markov processes*, Wiley, New York, 1986. MR-838085

[20] S. N. Evans, *Kingman's coalescent as a random metric space*, Stochastic models (Ottawa, ON, 1998), CMS Conf. Proc., vol. 26, Amer. Math. Soc., Providence, RI, 2000, pp. 105–114. MR-1765005

[21] S. N. Evans, *Probability and real trees*, Lecture Notes in Mathematics, vol. 1920, Springer, Berlin, 2008, Lectures from the 35th Summer School on Probability Theory held in Saint-Flour, July 6–23, 2005. MR-2351587

[22] S. N. Evans, R. Grübel, and A. Wakolbinger, *Doob–Martin boundary of Rémy's tree growth chain*, Ann. Probab. **45** (2017), no. 1, 225–277. MR-3601650

[23] N. Forman, C. Haulk, and J. Pitman, *A representation of exchangeable hierarchies by sampling from real trees*, Probab. Theory Relat. Fields (2017). https://doi.org/10.1007/s00440-017-0799-4

[24] C. Foucart, *Generalized Fleming-Viot processes with immigration via stochastic flows of partitions*, ALEA Lat. Am. J. Probab. Math. Stat. **9** (2012), no. 2, 451–472. MR-3069373

[25] A. Greven, P. Pfaffelhuber, and A. Winter, *Convergence in distribution of random metric measure spaces (Λ-coalescent measure trees)*, Probab. Theory Related Fields **145** (2009), no. 1–2, 285–322. MR-2520129

[26] A. Greven, P. Pfaffelhuber, and A. Winter, *Tree-valued resampling dynamics Martingale problems and applications*, Probab. Theory Related Fields **155** (2013), no. 3–4, 789–838. MR-3034793

[27] S. Gufler, *A representation for exchangeable coalescent trees and generalized tree-valued Fleming-Viot processes* (2017). arXiv:1608.08074v3

[28] S. Gufler, *Invariance principles for tree-valued Cannings chains* (2017). arXiv:1608.08203v2

[29] S. Gufler, *Pathwise construction of tree-valued Fleming-Viot processes*, Electron. J. Probab. **23** (2018).

[30] B. Hughes, *Trees and ultrametric spaces: a categorical equivalence*, Adv. Math. **189** (2004), no. 1, 148–191. MR-2093482

[31] O. Kallenberg, *Foundations of modern probability*, Second edition, Springer, New York, 2002. MR-1876169

[32] O. Kallenberg, *Probabilistic symmetries and invariance principles*, Springer, New York, 2005. MR-2161313

[33] G. Kersting, J. Schweinsberg, and A. Wakolbinger, *The evolving beta coalescent*, Electron. J. Probab. **19** (2014), no. 64. MR-3238784

[34] J. F. C. Kingman, *The coalescent*, Stochastic Process. Appl. **13** (1982), no. 3, 235–248. MR-671034

[35] S. Kliem and W. Löhr, *Existence of mark functions in marked metric measure spaces*, Electron. J. Probab. **20** (2015), no. 73. MR-3371432

[36] S. Kliem and A. Winter, *Evolving phylogenies of trait-depending branching with mutation and competition. Part I: Existence* (2017). arXiv:1705.03277

[37] T. G. Kurtz, *Martingale problems for conditional distributions of Markov processes*, Electron. J. Probab. **3** (1998), no. 9, 29 pp. MR-1637085

[38] T. G. Kurtz and G. Nappo, *The filtered martingale problem*, The Oxford handbook of nonlinear filtering, Oxford Univ. Press, 2011, pp. 129–165. MR-2884595

[39] C. Labbé, *From flows of Λ-Fleming-Viot processes to lookdown processes via flows of partitions*, Electron. J. Probab. **19** (2014), no. 55. MR-3227064

[40] W. Löhr, *Equivalence of Gromov-Prohorov- and Gromov's $\square_\lambda$-metric on the space of metric measure spaces*, Electron. Commun. Probab. **18** (2013), no. 17. MR-3037215

[41] M. Möhle, *Asymptotic results for coalescent processes without proper frequencies and applications to the two-parameter Poisson-Dirichlet coalescent*, Stochastic Process. Appl. **120** (2010), no. 11, 2159–2173. MR-2684740

[42] M. Möhle and S. Sagitov, *A classification of coalescent processes for haploid exchangeable population models*, Ann. Probab. **29** (2001), no. 4, 1547–1562. MR-1880231

[43] P. Pfaffelhuber and A. Wakolbinger, *The process of most recent common ancestors in an evolving coalescent*, Stochastic Process. Appl. **116** (2006), no. 12, 1836–1859. MR-2307061

[44] P. Pfaffelhuber, A. Wakolbinger, and H. Weisshaupt, *The tree length of an evolving coalescent*, Probab. Theory Related Fields **151** (2011), no. 3–4, 529–557. MR-2851692

[45] J. Pitman, *Coalescents with multiple collisions*, Ann. Probab. **27** (1999), no. 4, 1870–1902. MR-1742892

[46] L. C. G. Rogers and J. W. Pitman, *Markov functions*, Ann. Probab. **9** (1981), no. 4, 573–582. MR-624684

[47] S. Sagitov, *The general coalescent with asynchronous mergers of ancestral lines*, J. Appl. Probab. **36** (1999), no. 4, 1116–1125. MR-1742154

[48] J. Schweinsberg, *Coalescents with simultaneous multiple collisions*, Electron. J. Probab. **5** (2000), no. 4. MR-1781024

[49] V. S. Varadarajan, *Groups of automorphisms of Borel spaces*, Trans. Amer. Math. Soc. **109** (1963), 191–220. MR-0159923

[50] A. Véber and A. Wakolbinger, *The spatial Lambda-Fleming-Viot process: an event-based construction and a lookdown representation*, Ann. Inst. Henri Poincaré Probab. Stat. **51** (2015), no. 2, 570–598. MR-3335017

[51] A. M. Vershik, *Random and universal metric spaces*, Dynamics and randomness II, Kluwer Acad. Publ., Dordrecht, vol. 10, 2004, pp. 199–228. MR-2153330