Tree-valued Fleming-Viot processes: a generalization, pathwise constructions, and invariance principles

Dissertation

zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt beim Fachbereich Informatik und Mathematik
der Johann Wolfgang Goethe-Universität
in Frankfurt am Main

von
Stephan Gufler
aus Augsburg

Frankfurt 2017
(D30)

Vom Fachbereich Informatik und Mathematik der
Johann Wolfgang Goethe-Universität als Dissertation angenommen.

Dekan: Prof. Dr. Uwe Brinkschulte

Gutachter: Prof. Dr. Götz Kersting

Prof. Dr. Anton Wakolbinger

Prof. Dr. Jean-François Le Gall

Datum der Disputation: 30. Januar 2017

Abstract

We study exchangeable coalescent trees and the evolving genealogical trees in models for neutral haploid populations.

We show that every exchangeable infinite coalescent tree can be obtained as the genealogical tree of iid samples from a random marked metric measure space when the marks are added to the metric distances. We apply this representation to generalize the tree-valued Fleming-Viot process to include the case with dust in which the genealogical trees have isolated leaves.

Using the Donnelly-Kurtz lookdown approach, we describe all individuals ever alive in the population model by a random complete and separable metric space, the lookdown space, which we endow with a family of sampling measures. This yields a pathwise construction of tree-valued Fleming-Viot processes. In the case of coming down from infinity, we also read off a process whose state space is endowed with the Gromov-Hausdorff-Prohorov topology. This process has additional jumps at the extinction times of parts of the population.

In the case with only binary reproduction events, we construct the lookdown space also from the Aldous continuum random tree by removing the root and the highest leaf, and by deforming the metric in a way that corresponds to the time change that relates the Fleming-Viot process with a Dawson-Watanabe process. The sampling measures on the lookdown space are then image measures of the normalized local time measures.

We also show invariance principles for Markov chains that describe the evolving genealogy in Cannings models. For such Markov chains with values in the space of distance matrix distributions, we show convergence to tree-valued Fleming-Viot processes under the conditions of Möhle and Sagitov for the convergence of the genealogy at a fixed time to a coalescent with simultaneous multiple mergers. For the convergence of Markov chains with values in the space of marked metric measure spaces, an additional assumption is needed in the case with dust.

Acknowledgments

I would like to express my profound gratitude towards my advisor Professor Götz Kersting and Professor Anton Wakolbinger for having suggested to me studying a rich topic that opened up into many directions, for their availability, detailed reading of earlier versions of some parts of this work, helpful advice, and for having granted me complete freedom to pursue my research interests. I am also very grateful to them and to Professor Jean-François Le Gall for having acted as referees.

To the whole stochastics group in Frankfurt, I am grateful for an enjoyable atmosphere. To several fellow musicians, I am grateful for inspiring chamber music. I am deeply grateful to my parents for so much.

I gratefully acknowledge the support of the DFG Priority programme 1590 and of the Stummel-Stiftung for several conference visits.

Contents

Chapter 1

Introduction

Evolving genealogies have recently been described as stochastic processes with values in the space of isomorphy classes of metric measure spaces, or in the space of isomorphy classes of marked metric measure spaces when they are considered together with types (see e. g. [25, 46]). A metric measure space is 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. By the metric space (X, r), a population can be modeled along with genealogical distances, and the measure μ allows to draw samples. A metric measure space (X, r, μ) can be viewed as a tree if (X, r) can be embedded into a real tree. A real tree [39, 67] is a metric space (T, d) such that (i) for all $x, y \in T$, there exists an isometry from the real interval $[0, d(x, y)]$ to T, and (ii) no subspace of (T, d) is isometric to the unit circle. For example, any finite tree with edge-lengths can be seen as the real tree that is obtained by gluing together real intervals that correspond to its edges.

Metric measure spaces are called isomorphic if there exists a measure-preserving isometry between the supports of the measures. By the Gromov reconstruction theorem [48, 91], the isomorphy class of a metric measure space (X, r, μ) is determined by the distance matrix distribution, i. e. the distribution of the matrix $(r(x_i, x_j))_{i,j\in\mathbb{N}}$ where $(x_i)_{i\in\mathbb{N}}$ is a μ-iid sequence. Greven, Pfaffelhuber, and Winter [45] show that the space of isomorphy classes of metric measure spaces is Polish when endowed with the Gromov-weak topology in which convergence is defined as weak convergence of the distance matrix distributions. Passing from a metric measure space to its isomorphy class removes the labeling that is given by the elements of the metric space and that often contains no relevant information.

Greven, Pfaffelhuber, and Winter [46] introduce the neutral tree-valued Fleming-Viot process as a process with values in the space of isomorphy classes of metric measure spaces that solves a well-posed martingale problem. This process is constructed in [46] as the limit in distribution of tree-valued stochastic processes that describe the evolving genealogy in Moran models, as population size tends to infinity. In population genetics, Moran and Cannings models are classical models for the evolution of a population, we refer the reader e. g. to Etheridge [33] for more background on these models. The measure-valued Fleming-Viot process is the diffusion limit of measure-valued processes that describe the evolution of the distributions of allelic types in Moran models, see e. g. Ethier and Kurtz

[36, 37]. Donnelly and Kurtz [27, 28] couple measure-valued processes for various population sizes in their lookdown models from which they give pathwise constructions of the measure-valued Fleming-Viot process and of more general processes.

In the literature, see e. g. [25, 28, 37, 90], also other evolutionary forces than random resampling (genetic drift) are considered, such as mutation, selection, recombination, and spatial structure. In the present dissertation, we restrict ourselves to random resampling in the neutral case, and we consider genealogies apart from types.

Starting point for the research that lead to the present dissertation was G. Kersting's and A. Wakolbinger's suggestion to the author to read off a tree-valued process associated with the evolving Kingman coalescent from a Brownian excursion, and to study its path properties. They pointed out in particular the article of Evans [38], where Kingman's coalescent is studied as a random metric measure space, and the construction of Berestycki and Berestycki [4] of Kingman's coalescent from a Brownian excursion into which also a Donnelly-Kurtz [28] lookdown model is embedded.

In Chapter 3 of the present dissertation, a pathwise construction of tree-valued Fleming-Viot processes is given directly from a lookdown model with simultaneous multiple reproduction events. Here the central object is a random complete and separable metric space (Z, ρ) that describes all individuals ever alive and that we call the lookdown space. The lookdown space is closely related to the lookdown graph of Pfaffelhuber and Wakolbinger [77]. A large part of Chapter 3 is devoted to the construction of families of sampling measures on the lookdown space. In the dust-free case where the genealogical trees have no isolated leaves, the sampling measure μ_t is supported by the closure of the subset of the individuals at time $t \in \mathbb{R}_+$. We construct μ_t on an event of probability 1 that does not depend on t. We read off a tree-valued Ξ-Fleming-Viot process as the process whose state at time t is the isomorphy class of the metric measure space (Z, ρ, μ_t). From regularity of the path $t \mapsto \mu_t$ and atomicity properties of the μ_t, we deduce that the tree-valued Ξ-Fleming-Viot process has a. s. càdlàg paths in the Gromov-weak topology, with jumps at the times of large reproduction events.

As initially suggested to the author by G. Kersting and A. Wakolbinger, we also consider evolving genealogies in a stronger topology, the Gromov-Hausdorff-Prohorov-topology, which emphasizes the metric space also independently of the sampling measure. In the case of coming down from infinity, i. e. if the individuals at any fixed time have finitely many ancestors at any earlier time, we also read off a tree-valued process from the lookdown model that has jumps in the Gromov-Hausdorff-Prohorov topology when the overall structure of the evolving genealogical trees changes at the extinction times of parts of the population. We call this process a tree-valued evolving Ξ-coalescent.

Another aim in the present dissertation is to generalize tree-valued Fleming-Viot processes to the case with dust in which the genealogical trees have isolated leaves. Here the key idea is to decompose the genealogical tree into the external branches and the remaining subtree. In the pathwise construction of Chapter 3, this decomposition of the genealogical tree at time t coincides a. s. with the decomposition that is obtained by pruning this genealogical tree at the latest reproduction event on the ancestral lineage of each leaf. We call the individual on the ancestral lineage at such a reproduction time the parent of that leaf, and we extend the lookdown space so as to include also the parents of the individuals at time 0. We construct a family $(m_t, t \in \mathbb{R}_+)$ of sampling measures on the product space $\hat{Z} \times \mathbb{R}_+$ of the extended lookdown space $(\hat{Z}, \rho)$ and the mark space $\mathbb{R}_+$. The $\hat{Z}$-component of m_t is supported by the set of the parents of the individuals at time t, and the $\mathbb{R}_+$-component records the genealogical distances between the individuals at time t and their parents. Then we read off a tree-valued Ξ-Fleming-Viot process as the process whose state at time t is the isomorphy class of the marked metric measure space $(\hat{Z}, \rho, m_t)$.

A 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 $X \times \mathbb{R}_+$. Here the second factor is called the mark space which will always be $\mathbb{R}_+$ in the present dissertation. Marked metric measure spaces and the space of their isomorphy classes are introduced by Depperschmidt, Greven, and Pfaffelhuber [24, 25] to describe genealogies along with allelic types which they encode by the marks. In the present dissertation, the marks encode lengths of external branches, and the distances between the starting vertices of the external branches are given by the metric space. We define the distance matrix distribution of a marked metric measure space (X, r, m) as the distribution of the matrix $(\rho(i,j))_{i,j}$ given by

$$\rho(i,j) = (v_i + r(x_i, x_j) + v(x_j))\, \mathbf{1}\{i \neq j\},$$

where $(x_i, v_i)_{i \in \mathbb{N}}$ is an m-iid sequence in $X \times \mathbb{R}_+$.

Semi-ultrametrics, i. e. semi-metrics ρ that satisfy the strong triangle inequality

$$\max\{\rho(i,j), \rho(j,k)\} \geq \rho(i,k),$$

can be embedded into real trees as subspaces of leaves, and they are in one-to-one correspondence with realizations of coalescents. In Chapter 2, we show that every exchangeable random semi-ultrametric ρ on $\mathbb{N}$ can be represented in a two-step random experiment in whose first step we sample a marked metric measure space χ, and in whose second step we sample ρ according to the distance matrix distribution of χ. Here we also state a deterministic function ζ such that $\chi = \zeta(\rho)$ a. s. As the distance matrix distribution of a marked metric measure space is invariant and ergodic with respect to finite permutations of rows and columns, our representation decomposes the distribution of an exchangeable semi-ultrametric into its ergodic components.

The genealogical distances between the individuals at time t in the lookdown model form an exchangeable semi-ultrametric ρ_t on $\mathbb{N}$ if this holds at initial time. When we apply the deterministic function ζ from our representation result to the process $(\rho_t, t \in \mathbb{R}_+)$, the image process with values in the space of isomorphy classes of marked metric measure spaces is a tree-valued Fleming-Viot process. We also call the process of the associated distance matrix distributions a tree-valued Fleming-Viot process. A tree-valued Fleming-Viot process can thus be seen as the process of the ergodic components of the evolving genealogical distances. Other than in the pathwise construction from Chapter 3, where we use techniques specific to the lookdown model, we obtain from our representation result for exchangeable semi-ultrametrics only versions of tree-valued Fleming-Viot processes as the representation result holds on an event of probability 1 that may depend on t. We describe these versions of tree-valued Fleming-Viot processes by martingale problems, and we show the Markov property by the criterion of Rogers and Pitman [82]. We read off a tree-valued Fleming-Viot process also from the dual flow of bridges of Bertoin and Le Gall [8].

In the setting with only binary reproduction events which is associated with the Kingman coalescent, we construct in Chapter 4 the lookdown space in two-sided time and the family of sampling measures, hence also pathwise a tree-valued Fleming-Viot process, from a Brownian excursion. To this aim, we remove the root and the highest leaf from the continuum random tree that is encoded by the Brownian excursion as in Aldous [1]. Then we deform the metric in this continuum random tree according to the time change by which Shiga [87] relates the measure-valued Dawson-Watanabe process and the measure-valued Fleming-Viot process. To identify the resulting metric space as a lookdown space, we also give a lookdown representation for the binary branching forest of Neveu and Pitman [74, 75] and Le Gall [65] which is embedded into reflected Brownian motion. We obtain the sampling measures on the lookdown space as image measures of the normalized local time measures, thanks to the uniform downcrossing representation for local times of Chacon et al. [21].

In Chapter 5, we show invariance principles for tree-valued Markov chains that we read off from Cannings models. Under the conditions of Möhle and Sagitov [72] and Sagitov [85] for convergence of the genealogy at a fixed time to a Ξ-coalescent, we show in the dust-free case that sequences of Markov chains with values in the space of metric measure spaces that describe the evolving genealogy in Cannings models converge under rescaling to tree-valued Fleming-Viot process. The Markov chains of the associated distance matrix distributions converge also in the case with dust to tree-valued Fleming-Viot processes. We include the case with dust also by considering Markov chains with values in the space of marked metric measure spaces that correspond to the decomposition of the genealogical trees in the Cannings models into the external branches and the remaining subtrees. Here an additional assumption is needed in the invariance principle as convergence of these marked metric measure spaces (in the marked Gromov-weak topology) implies convergence of the empirical distributions of the external branch lengths. To show that an additional assumption cannot be omitted, we give an example in which the limiting genealogy is given by a coalescent with dust, i. e. with positive external branch lengths. The external branch lengths in the Cannings models from the example, however, vanish due to "perturbative" reproduction events that occur at high rate but that are not visible in the limiting genealogy as each of them affects only a small part of the population.

Publications. Chapters 2, 3, and 5 are essentially the preprints [50], [49], and [51], respectively, which is work submitted for publication. Chapter 4 presents work in progress.

Chapter 2

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

We show that every exchangeable random semi-ultrametric on $\mathbb{N}$ can be obtained by sampling an iid sequence from a random marked metric measure space and adding the marks to the distances. We use this representation to define tree-valued Fleming-Viot processes from the Ξ-lookdown model. The case with dust is included for processes with values in the space of marked metric measure spaces, and for processes with values in the space of distance matrix distributions.

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 [59] 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 [79], Sagitov [84], Donnelly and Kurtz [28]) where multiple blocks are allowed to merge to a single block at the same time, and the Ξ-coalescent (Möhle and Sagitov [72], Schweinsberg [86]) where several clusters of blocks may also merge simultaneously.

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 $s \leq t$. Assuming that each pair of integers is in a common block of π_t for t sufficiently large, $(\pi_t, t \in \mathbb{R}_+)$ can equivalently be expressed as a semi-ultrametric ρ on $\mathbb{N}$

such that for all $t \in \mathbb{R}_+$ and $i, j \in \mathbb{N}$,

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

and (1.1) yields a one-to-one correspondence between these càdlàg paths and the semi-ultrametrics on $\mathbb{N}$, cf. [39, Example 3.41] and [40, p. 262]. A (semi-)ultrametric ρ can also be defined as a (semi-)metric that satisfies the strong triangle inequality

$$\max\{\rho(x,y), \rho(y,z)\} \geq \rho(x,z).$$

Evans [38] 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. [6].

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 chapter. 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 [39] 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 $d((s,i),(t,j)) = \max\{\rho(i,j) - s - t, |s-t|\}$ (cf. e. g. [40, p. 262] or [54, Section 6]). 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.

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 isometry classes. 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 $1, \dots, n$, as $n \to \infty$. Then we obtain the description by isomorphy classes of metric measure spaces of Greven, Pfaffelhuber, and Winter [45] 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 [45] 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 metric measure space is 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. An important feature is that 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, m). By the Gromov reconstruction theorem (see Theorem 4 of Vershik [91]), 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. Under an appropriate condition, Theorem 5 of [91] associates with any typical realization of a random semi-metric on $\mathbb{N}$ a metric measure space whose distance matrix distribution is the distribution of this semi-metric. In the ultrametric case, this metric measure space 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 3.9). This can also be deduced from [91, Equation (9)].

1.2 The sampling representation

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 $\mathbb{N}$. 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 [24]. 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 $X \times \mathbb{R}_+$. We call two marked metric measure spaces isomorphic if they have the same marked distance matrix distribution. We endow the space of isomorphy classes of marked metric measure spaces with the Gromov-weak topology in which marked metric measure spaces converge if and only if their marked distance matrix distributions converge weakly. This yields a Polish space, as shown in [24].

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

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

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

Theorem 1.2. *Let ρ be an exchangeable semi-ultrametric on $\mathbb{N}$. Then there exists a random variable χ with values in the space of isomorphy classes of marked metric measure spaces such that the law of ρ equals $\mathbb{E}[\nu]$, where ν denotes the distance matrix distribution of χ. That is, for a random semi-metric ρ' on $\mathbb{N}$ with conditional distribution ν given χ, the random variables ρ and ρ' have the same (unconditional) distribution. The random variable χ is unique in distribution.*

A stronger formulation of Theorem 1.2 is given in Theorem 3.18 where we state the marked metric measure space χ realizationwise as a deterministic function of ρ'. The key idea for the construction of χ is to decompose the tree that is associated with 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 χ, the marks encode the external branch lengths, and the metric space describes the remaining subtree. The external branches all have length zero in particular in the dust-free case. In this case, χ can also be replaced by the isomorphy class of a metric measure space (as in Section 3.3).

We prove the main part of Theorem 3.18 in Section 11.2. We formulate the decomposition at the external branches in Section 2 in terms of semi-ultrametrics. As we discuss in Section 3.2 and in Remark 3.15, marked metric measure spaces whose distance matrix distribution is concentrated on semi-ultrametrics correspond to real trees that are endowed with a probability measure that is concentrated on the starting vertices of the external branches. We mention that Evans, Grübel, and Wakolbinger [40] 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 [40, Section 7], a sampling representation for exchangeable ultrametrics is considered (see Remark 11.2).

We define the action of a finite permutation p of $\mathbb{N}$ on a semi-ultrametric ρ on $\mathbb{N}$ by $p(\rho) = (\rho(p(i), p(j)))_{i,j\in\mathbb{N}}$. The distance matrix distribution of a marked metric measure space (X, r, m) is invariant and ergodic with respect to the action of the group of finite permutations. Indeed (analogously to [91, Lemma 7]), a finite permutation of the distance matrix distribution corresponds to a finite permutation of the m-iid sequence $(x(i), v(i))_{i\in\mathbb{N}}$ which is invariant and ergodic with respect to finite permutations, and from which the distance matrix distribution is obtained as the distribution of $((v(i) + r(x(i), x(j)) + v(j))\,\mathbf{1}\{i \neq j\})$. Hence, Theorem 1.2 decomposes the distribution of an exchangeable semi-ultrametric ρ into ergodic components. The ergodic component, that is, the distance matrix distribution of χ, can be read off from each typical realization of ρ by Theorem 3.18 (see also Remark 3.22). The ergodic component is also characterized by the marked metric measure space χ itself. In the dust-free case, it can also be expressed as 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. The representation given by Theorem 1.2 can also be seen in the context of the more general but less explicit Aldous-Hoover-Kallenberg representation, see e. g. [56, Section 7].

1.3 Evolving genealogies

In Section 4, we lay the foundation for our study of evolving genealogies by considering a 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 to the processes of the ergodic components, expressed as isomorphy classes of metric measure spaces, marked metric measure spaces, and distance matrix distributions, respectively, as outlined in Subsection 1.2. Using the criterion of Rogers and Pitman [82, 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 [62], and Kurtz and Nappo [63].

In Sections 5 – 7, 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 [27, 28] to represent measure-valued processes along with their genealogy, see also e. g. Etheridge and Kurtz [34] and Birkner et al. [13]. 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) remains exchangeable. In [28] and in the present work, 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 [78] and Dahmer, Knobloch, and Wakolbinger [23].

In Section 8, we remove the labels from the process of the evolving genealogical trees in the infinite lookdown model by applying the result from Section 4 to the process from Sections 5 – 7. 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 [46] 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 [46, Remark 2.20], a construction of (a version of) this process from the lookdown model of Donnelly and Kurtz [27] is outlined. The aim in the present chapter 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 [25, 26] where the states are isomorphy classes of marked metric measure spaces and the marks encode allelic types. In the present work, the marks encode lengths of external branches. We consider only the neutral case, and we describe genealogies without using types.

In Section 9, 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 10, we show that tree-valued Fleming-Viot processes converge in distribution to equilibrium.

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 [42]. There are many other representation results for exchangeable structures in the literature. For instance, by the Dovbysh-Sudakov theorem, see Austin [3] 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 [77]. Kliem and Löhr [61] further study marked metric measure spaces. In their article, tree-valued Λ-Fleming-Viot processes in the dust-free case and an application of marked metric measure spaces to trait-depending branching are also mentioned. In the context of measure-valued spatial Λ-Fleming-Viot processes with dust, Véber and Wakolbinger [90] work with a skeleton structure. Functionals of coalescents like external branch lengths have also been studied, see for example [71]. Also the time evolution of such functionals has been studied for evolving coalescents, see for example [22, 58].

Bertoin and Le Gall [8–10] 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 [7]) also describe genealogies without labeling individuals. In Section 12, we construct the Fleming-Viot process with values in the space of distance matrix distributions from the dual flow of bridges. We also mention the work of Labbé [64] where relations between the lookdown model and flows of bridges are studied.

While we construct versions of tree-valued Fleming-Viot processes in the present chapter using the representation result (see Remark 8.2), a pathwise construction that uses techniques specific to the lookdown model is given in Chapter 3.

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

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

we define the space

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

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

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

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

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

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. [39]). 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 $(v(i), i)$ 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 $\varphi(i) = (v(i), i)$ 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.

3 Sampling from marked metric measure spaces

3.1 Preliminaries

Recall the space of isomorphy classes of metric measure spaces from Section 1.1. We denote this space 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 [45] showed that $\mathbb{M}$ is then a Polish space. Also recall from Section 1.2 the space of isomorphy classes of marked metric measure spaces which we denote by $\hat{\mathbb{M}}$. The distance matrix distribution of the isomorphy class of a metric measure space $\chi \in \mathbb{M}$ is denoted by ν^χ. The marked distance matrix distribution of $\chi' \in \hat{\mathbb{M}}$ is denoted by $\nu^{\chi'}$, so that $\alpha(\nu^{\chi'})$ is the distance matrix distribution of χ', in accordance with the definition in Section 1.2.

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

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

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 3.1. 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. [56, 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 3.2. The coalescents associated by (1.1) with the exchangeable semi-ultrametrics on $\mathbb{N}$ form a larger class of processes than than the so-called exchangeable coalescents defined in e. g. Section 4.2.2 of Bertoin [7]. 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

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

of ultrametric measure spaces which is a closed subspace of $\mathbb{M}$, as shown in [46, Lemma 2.3]. By the same argument, the space

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

is a closed subspace of $\hat{\mathbb{M}}$. It contains the marked metric measure spaces with ultrametric distance matrix distribution. Following e. g. [45, 46], we call the elements of $\mathbb{U}$ trees. Also the elements of $\hat{\mathbb{U}}$ may be called trees, as we discuss in this subsection.

First we show in Proposition 3.3 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 3.3. *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*

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

The proof 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. 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

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

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

$$r'(x(i), x(j)) \vee |v(i) - v(j)| \leq 2\varepsilon \quad \text{a. s.} \tag{3.1}$$

By inequality (3.1) and the definition of ρ, it follows that

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

Using the definition of the map Υ, we deduce

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

For the converse inequality, we first note that

$$2v(i) \leq v(i) + v(j) + 2\varepsilon + r(i,j) = \rho(i,j) + 2\varepsilon \tag{3.2}$$

by inequality (3.1) and the definition of ρ. Moreover, for all $k \in \mathbb{N} \setminus \{i, j\}$, we obtain

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

Here we use inequality (3.2) for the first and inequality (3.1) 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. □

In Remark 3.5 below, we interpret the elements of $\hat{\mathbb{U}}$ as weighted real trees that are non-separable in general. We give a similar interpretation in Remark 3.15 in Subsection 3.4 where we have separable trees. In Remark 3.5, we use the concept of mark functions for which we refer to Depperschmidt, Greven, and Pfaffelhuber [26] and Kliem and Löhr [61]. A marked metric measure space (X, r, m) is said to admit a uniformly continuous mark function if there exists a uniformly continuous function $f : X \to \mathbb{R}_+$ such that the probability measure m on $X \times \mathbb{R}_+$ factorizes as $m(dx\, dv) = m_X(dx)\delta_{f(x)}(dv)$, where $m_X := m(\cdot \times \mathbb{R}_+)$. This is clearly a property of the isomorphy class.

Proposition 3.4. *Let $\chi \in \hat{\mathbb{U}}$. Then χ admits a uniformly continuous mark function.*

Proof. Let (r, v) be a random variable with the marked distance matrix distribution of χ, and let $\rho = \alpha(r, v)$. Proposition 3.3 yields $v(i) = \frac{1}{2} \inf_{k \in \mathbb{N}\setminus\{i\}} \rho(i, k)$ a. s. for all $i \in \mathbb{N}$. Hence,

$$r(i, j) = \rho(i, j) - v(i) - v(j) \geq 2 \max\{v(i), v(j)\} - v(i) - v(j) = |v(i) - v(j)| \quad \text{a. s.}$$

for all distinct $i, j \in \mathbb{N}$. The assertion follows by Lemma 1.13 of [61]. □

Remark 3.5. Let (X, r, m) be a marked metric measure space whose isomorphy class χ lies in $\hat{\mathbb{U}}$. We assume that the support of the measure $m(\cdot \times \mathbb{R}_+)$ is equal to X. In this remark, we extend the space (X, r) to a real tree.

By Proposition 3.4, the marked metric measure space (X, r, m) admits a uniformly continuous mark function $f : X \to \mathbb{R}_+$. For each point $x \in X$, we take infinitely many real intervals of length $f(x)$. For each of these intervals, we glue one endpoint to x. We call the other endpoint an outer endpoint and denote it by $\ell_{x,i}$, where $i \in \mathcal{I}$ and $\mathcal{I}$ is an index set for the intervals attached to x. The induced semi-metric ρ on the set of outer endpoints $\mathcal{L} = \{\ell_{x,i} : x \in X, i \in \mathcal{I}\}$ is given by

$$\rho(\ell_{x,i}, \ell_{y,j}) = f(x) + r(x, y) + f(y)$$

for $(x, i) \neq (y, j)$. By definition of $\hat{\mathbb{U}}$, our assumption on the support of m, and continuity of f, it follows that ρ is a semi-ultrametric on $\mathcal{L}$.

With $(\mathcal{L}, \rho)$, we associate as in Remark 1.1 a real tree on whose completion $\bar{T}$ a probability measure is given as the image measure of $m(\cdot \times \mathbb{R}_+)$ under the isometry $X \to \bar{T}$ that is defined analogously to the map φ in Remark 2.2. This probability measure is concentrated on the subspace $\{(x, f(x)) : x \in X\}$ of the starting vertices of the external branches. Continuity of the mark function f implies that this subspace is closed.

3.3 Dust-free semi-ultrametrics and dust-free marked metric measure spaces

We call a semi-ultrametric $\rho \in \mathfrak{U}$ dust-free if $\Upsilon(\rho) = 0$. We call a marked metric measure space (X, r, m) dust-free if the probability measure 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 $(X, r, \mu \otimes \delta_0)$.

Remark 3.6. Note that $f = 0$ in Proposition 3.4 if and only if χ is dust-free. Clearly, (the isomorphy class of) a marked metric measure space (X, r, m) in $\hat{\mathbb{U}}$ is dust-free if and only if a random variable with distribution $\nu^{(X,r,m)}$ is a. s. dust-free. In particular, in Theorem 1.2, the marked metric measure space χ is a. s. dust-free if and only if ρ is a. s. dust-free. In the ultrametric case, condition (4) in Vershik [91] therefore corresponds to the condition that the random semi-ultrametric is dust-free.

3.4 Marked metric measure spaces from marked distance matrices

In this subsection, we discuss the main concepts for the sampling representation in Subsection 3.5. Consider a two-step random experiment where we first sample a random (marked) metric measure space, conditionally on which we sample according to its (marked) distance matrix distribution in the second step. By Proposition 3.12 below, the realization of the (marked) metric measure space in the first step can be reconstructed from any typical realization of the (marked) distance matrix distribution in the second step. In the beginning of 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 also given in Remark 3.9 below. In Remark 3.22, we state their role in the context of the ergodic decomposition.

Now 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 in view of Proposition 3.10 below.

Remark 3.7. 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 3.8. *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 11.1.

Remark 3.9 (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.

When we sample according to the (distance) matrix distribution of the (marked) metric measure space that we have constructed from a realization of a suitable random (marked) distance matrix, we obtain a (marked) distance matrix with the same distribution, as we check in Proposition 3.10 below.

The following proposition can be compared with Lemma 8 of Vershik [91]. It will be needed also for the intertwining relations in Section 4.

Proposition 3.10. *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 3.11. 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 3.10, we set $(r,v) = (\rho, 0)$.

For $n \in \mathbb{N}$, we write $[n] = \{i \in \mathbb{N} : i \leq n\}$ and we denote by γ_n the restriction from $\mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}}$ to $\mathbb{R}^{n^2} \times \mathbb{R}^n$, $\gamma_n(r,v) = ((r(i,j))_{i,j\in[n]}, (v(i))_{i\in[n]})$.

Proof of Proposition 3.10. 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 $\hat{\mathfrak{D}}^*$. We have

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

The second equality follows from the definition of $\hat{\mathfrak{D}}^*$ and by dominated convergence. For the third equality, we use that summands where $\ell_1, \dots, \ell_n$ 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. □

Proposition 3.12. *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 3.13. Proposition 3.12 is essentially Vershik's proof [91, Theorem 4] of the Gromov reconstruction theorem (where metric measure spaces are considered, cf. also [24, Theorem 1] for marked metric measure spaces). The present formulation focuses on the map $\hat{\psi}$ that will be needed in the proofs of Theorems 3.18 and 4.2 below.

Remark 3.14. For $\chi \in \mathbb{M}$ and a $\mathfrak{D}$-valued random variable ρ with distribution ν^χ, Proposition 3.12 implies $\rho \in \mathfrak{D}^*$ a. s. and $\psi(\rho) = \chi$ a. s.

Proof of Proposition 3.12. Let (X', r', m') be a representative of χ. W. l. o. g. we assume that the 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. □

Remark 3.15 (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 3.12, we can associate a. s. with (r, v) a complete and separable weighted real tree $(\bar{T}, d, \mu)$ as in Remark 3.9. By the argument from Vershik [91, Lemma 7] which we recalled in the end of Section 1.2, the 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 $(\bar{T}, d, \mu)$ from Remark 3.9 is constant a. s.

3.5 The sampling representation

In Proposition 3.16, we consider a typical realization of an exchangeable semi-ultrametric ρ on $\mathbb{N}$, and its decomposition $\beta(\rho)$. Proposition 3.16 states that the sampling measure m in the definition of $\hat{\psi}(\beta(\rho))$ in Subsection 3.4 is the weak limit of the uniform probability measures therein.

Proposition 3.16. *Let ρ be an exchangeable random variable with values in $\mathfrak{U}$. Then $\beta(\rho) \in \hat{\mathfrak{D}}^*$ a. s.*

In Section 11.2, we give various proofs of this proposition. In two of them, the de Finetti theorem yields the aforementioned sampling measure m as the directing measure of an exchangeable sequence.

In the dust-free case, we need not decompose the semi-metric ρ by the map β, we can work directly with the map ψ from Subsection 3.4:

Corollary 3.17. *Let ρ be an exchangeable random variable with values in $\mathfrak{U}$ that is a. s. dust-free. Then $\rho \in \mathfrak{D}^*$ a. s.*

Proof. As ρ is a. s. dust-free, $\rho = r$ a. s. and the assertion is immediate from Proposition 3.16 and Remark 3.7. □

We obtain a stronger version of Theorem 1.2 in which the marked metric measure space χ is stated as a deterministic function of the realization of ρ.

Theorem 3.18. *Let ρ be an exchangeable $\mathfrak{U}$-valued random variable. Then a $\hat{\mathbb{U}}$-valued random variable is defined by $\chi = \hat{\psi} \circ \beta(\rho)$. Let ρ' be a $\mathfrak{U}$-valued random variable with conditional distribution $\alpha(\nu^\chi)$ given χ. Then ρ and ρ' are equal in distribution and $\chi = \hat{\psi} \circ \beta(\rho')$ a. s.*

Proof of Theorem 3.18. The random variable χ is well-defined by Proposition 3.8. That ρ and ρ' are equal in distribution follows from Propositions 3.16 and 3.10. Proposition 3.12 implies $\chi = \hat{\psi} \circ \beta(\rho')$ a. s. □

To deduce Theorem 1.2, it remains to show the uniqueness assertion, which follows from Proposition 3.19 below.

Proposition 3.19. *Let χ and χ' be $\hat{\mathbb{U}}$-valued random variables such that $\mathbb{E}[\alpha(\nu^\chi)] = \mathbb{E}[\alpha(\nu^{\chi'})]$. Then χ and χ' are equal in distribution.*

Proof. Let (r, v) be a random variable with conditional distribution ν^χ given χ, and let $\rho = \alpha(r, v)$. Propositions 3.12 and 3.3 imply $\chi = \hat{\psi} \circ \beta(\rho)$ a. s. Hence, the distribution of ρ determines the distribution of χ uniquely, which is the assertion. □

Analogously to Theorem 3.18, we obtain using Corollary 3.17 and Remarks 3.11 and 3.14:

Corollary 3.20. *Let ρ be an exchangeable $\mathfrak{U}$-valued random variable that is a. s. dust-free. Then a $\mathbb{U}$-valued random variable is defined by $\chi = \psi(\rho)$. Let ρ' be a $\mathfrak{U}$-valued random variable with conditional distribution ν^χ given χ. Then ρ and ρ' are equal in distribution and $\chi = \psi(\rho')$ a. s.*

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

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

of distance matrix distributions of marked metric measure spaces. A one-to-one correspondence between the sets $\mathcal{U}^{\mathrm{erg}}$ and $\hat{\mathbb{U}}$ is given by Proposition 3.3 and the definitions of these sets. Hence, also the elements of $\mathcal{U}^{\mathrm{erg}}$ can be seen as trees.

Remark 3.21. By Theorem 1.2, each element of $\mathcal{U}$ is a mixture of elements of $\mathcal{U}^{\mathrm{erg}}$. As the distance matrix distribution of a marked metric measure space is invariant and ergodic (with respect to the action of the group of finite permutations), and as the ergodic distributions in $\mathcal{U}$ are extreme in the convex set $\mathcal{U}$ (see e. g. [56, Lemma A1.2]), the set $\mathcal{U}^{\mathrm{erg}}$ is equal to the set of ergodic distributions in $\mathcal{U}$.

Remark 3.22. Also the more explicit Theorem 3.18 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 [56]. The function

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

plays the role of the decomposition map in Varadarajan [89] in the sense that $\zeta(\rho')$ is the ergodic component in whose support a typical realization ρ' lies. Note that this ergodic component is characterized by the marked metric measure space $\chi = \hat{\psi} \circ \beta(\rho')$, and in the dust-free case also by the metric measure space $\psi(\rho')$. Some further references on the ergodic decomposition are given e. g. in [56, p. 475].

Remark 3.23. In the context of Theorem 3.18, $(\rho, \alpha(\nu^\chi))$ and $(\rho', \alpha(\nu^\chi))$ are equal in distribution. Hence, the probability kernel Λ from $\mathcal{U}$ to $\mathfrak{U}$, given by $\Lambda(\nu, B) = \nu(B)$ for $\nu \in \mathcal{U}$, $B \subset \mathfrak{U}$ Borel, is a regular conditional distribution of ρ given $\alpha(\nu^\chi)$.

The following corollary, which will be applied in Section 12, corresponds to the uniqueness of the ergodic component of a typical realization of ρ' in the context of Remark 3.22.

Corollary 3.24. *Let $\nu \in \mathcal{U}^{\text{erg}}$ and let ρ be a random variable with distribution ν. Then the distance matrix distribution of $\hat{\psi} \circ \beta(\rho)$ equals ν a. s.*

Proof. By definition of $\mathcal{U}^{\text{erg}}$, there exists $\chi \in \hat{\mathbb{U}}$ with distance matrix distribution $\nu^\chi = \nu$. Propositions 3.3 and 3.12 imply $\hat{\psi} \circ \beta(\rho) = \chi$ a. s. □

The following corollary to Theorem 1.2 implies that $(\mathcal{U}^{\text{erg}}, d_{\text{P}})$ is Polish which will be applied in Chapter 5.

Corollary 3.25. *The subspace $\mathcal{U}^{\text{erg}}$ is closed in $\mathcal{U}$.*

Proof. Let $(\nu^n, n \in \mathbb{N}) \subset \mathcal{U}^{\text{erg}}$ be a sequence that converges to some ν in $(\mathcal{U}, d_{\text{P}})$. By Theorem 1.2, there exists a random marked metric measure space χ with $\nu = \mathbb{E}[\alpha(\nu^\chi)]$. We show that $\nu' := \alpha(\nu^\chi)$ is independent of itself. With this property, it follows that $\nu' = \nu$ a. s. and $\nu \in \mathcal{U}^{\text{erg}}$.

To show this independence, let $\Psi_1, \Psi_2 : \mathcal{U} \to \mathbb{R}$ be bounded and continuous. As also the function $\Psi_1\Psi_2$ is bounded and continuous,

$$\mathbb{E}[\Psi_1(\nu')\Psi_2(\nu')] = \lim_{n\to\infty} \Psi_1(\nu^n)\Psi_2(\nu^n) = \mathbb{E}[\Psi_1(\nu')]\mathbb{E}[\Psi_2(\nu')].$$

We conclude by [36, Theorem 3.4.6]. □

4 Application to tree-valued processes

Using the function $\hat{\psi}$ from Section 3.4, we map a Markov process 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 3.18) 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 3.20.

Using the criterion of Rogers and Pitman [82, Theorem 2], we show that also the image process is Markovian. A martingale problem for the $\mathfrak{U}$-valued process yields a martingale problem for the image process.

Remark 4.1. Theorem 4.2 below is an example for Markov mapping. That the image processes solve the martingale problems given in Theorem 4.2 can also be seen as a consequence of Lemma A.2 in Kurtz and Nappo [63]. Uniqueness of these martingale problems follows under suitable conditions from Lemma 3.5 in Kurtz [62]. In Section 8, we show uniqueness directly by duality for concrete examples.

The so-called polynomials and marked polynomials, introduced in [24, 45] have been used as domains of martingale problems in e. g. [25, 26, 46]. 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 Chapter 5. Recall the restriction γ_n from $\mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}}$ to $\mathbb{R}^{n^2} \times \mathbb{R}^n$ and the notation $[n] = \{1, \dots, n\}$ for $n \in \mathbb{N}$. We denote also by γ_n the restriction from $\mathbb{R}^{\mathbb{N}^2}$ to $\mathbb{R}^{n^2}$, $\gamma_n(r, v) = ((r_{i,j})_{i,j \in [n]}, (v_i)_{i \in [n]})$. 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. 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. 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

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

the set of marked polynomials by

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

and we define the set of test functions

$$\mathscr{C} = \{\mathcal{U}^{\text{erg}} \to \mathbb{R}, \nu \mapsto \nu\phi : \phi \in \mathcal{C}\}.$$

For a metric space E, a subset $\mathscr{D}$ of the set $M_{\text{b}}(E)$ of bounded measurable functions $E \to \mathbb{R}$, and an operator $G : \mathscr{D} \to M_{\text{b}}(E)$, 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 [36, p. 173].

Theorem 4.2. *Let $(\rho_t, t \in \mathbb{R}_+)$ be a $\mathfrak{U}$-valued 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 $(\nu_t, t \in \mathbb{R}_+) := (\alpha(\nu^{\hat{\chi}_t}), t \in \mathbb{R}_+)$. 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

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

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

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

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

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

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 $\chi \in \mathbb{U}$.

Remark 4.3. The process $(\beta(\rho_t), t \in \mathbb{R}_+)$ in Theorem 4.2 is also Markov. This follows as $\rho_t = \alpha(\beta(\rho_t))$ and as $(\rho_t, t \in \mathbb{R}_+)$ is Markov by assumption.

Remark 4.4. In Theorem 4.2, if ρ_t is dust-free for some $t \in \mathbb{R}_+$, then χ_t is the (isomorphy class of the) metric measure space associated (as in Section 3.3) with (any representative of) the dust-free marked metric measure space $\hat{\chi}_t$, and we have $\nu_t = \nu^{\chi_t}$. The process $(\chi_t, t \in \mathbb{R}_+)$ is only relevant in the dust-free case: If ρ_t is not dust-free, then $\psi(\rho_t)$ is a. s. just the arbitrary element of $\mathbb{M}$ from the definition of ψ in Section 3.4.

Remark 4.5. In particular in Sections 9 – 10 and in Chapter 5, we need convergence determining sets of test functions. As in [24, 45, 69], the sets Π and $\hat{\Pi}$ are convergence determining in $\mathbb{U}$ and $\hat{\mathbb{U}}$, respectively. The argument from [69, 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. [69, 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 [24, 45]) and by definition of $\mathcal{U}^{\mathrm{erg}}$, also $\mathscr{C}$ is an algebra. Again by [69, Theorem 2.7], it follows that $\mathscr{C}$ is convergence determining in $\mathcal{U}^{\mathrm{erg}}$.

Remark 4.6. 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 3.3 and 3.12 as in the proof of Proposition 3.19. 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 [45, Example 2.12(ii)].

The following property is central in the proof of Theorem 4.2.

Proposition 4.7. *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 4.2 hold, then*

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

and

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

If the assumptions of Theorem 4.2 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 Propositions 3.16 and 3.10, the definition of ν_t, Corollary 3.17, and Remark 3.11. □

Remark 4.8. In the context of Theorem 4.2(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 4.7 yields the intertwining relation $Q_t K = K P_t$ which is condition (b) in [82, Theorem 2]. Many papers appeared on intertwining of Markov processes, a classical one is for instance [20].

Proof of Theorem 4.2. We apply [82, 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 [82] also holds when the initial state y therein is random. Then by Proposition 4.7, condition (b) and the condition on the initial state in [82, Theorem 2] are satisfied. Condition (a) in [82, Theorem 2] follows from Proposition 3.12 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 [82, Theorem 2].

Now we prove 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,

$$\begin{aligned} 0 =& \mathbb{E}[\phi(\beta(\rho_t)) - \phi(\beta(\rho_0)) - \int_0^t \hat{A}\phi(\beta(\rho_u))du] \\ =& \mathbb{E}[\nu^{\hat{\chi}_t}\phi] - \mathbb{E}[\nu^{\hat{\chi}_0}\phi] - \int_0^t \mathbb{E}[\nu^{\hat{\chi}_u}(\hat{A}\phi)]du \\ =& \mathbb{E}[\Phi(\hat{\chi}_t) - \Phi(\hat{\chi}_0) - \int_0^t \hat{B}\Phi(\hat{\chi}_u)du] \end{aligned} \tag{4.1}$$

by Proposition 4.7, Fubini, and the definitions Φ and $\hat{B}$. By the Markov property of $(\chi_s, s \in \mathbb{R}_+)$ and equation (4.1), it now follows for all $s \in \mathbb{R}_+$ and all $(\hat{\chi}_u, u \in [0,s])$-measurable events A that

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

which shows assertion (i).

The proof of (ii) is analogous, we apply [82, 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 [82, Theorem 2] is satisfied by Propositions 3.12 and 3.3, and by definition of $\mathcal{U}^{\text{erg}}$.

Also the proof of (iii) is analogous. We apply [82, 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 4.7 and Remark 3.14. □

5 Genealogy in the lookdown model

In this and the next section, we define a Markov process $(\rho_t, t \in \mathbb{R}_+)$ to which we will later apply Theorem 4.2. We read off this process from a population model that is driven by a deterministic point measure in this section and by a Poisson random measure in the next section. We remark that for the lookdown model of Donnelly and Kurtz [27], the process of the evolving genealogical distances and its martingale problem are considered in Remark 2.20 of Greven, Pfaffelhuber, and Winter [46].

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 $[n] = \{1, \dots, n\}$. 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 $\mathbf{0}_n = \{\{1\}, \dots, \{n\}\}$ 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 $B_1(\pi), B_2(\pi), \dots$ the enumeration of the blocks of π with $\min B_1(\pi) < \min B_2(\pi) < \dots$. 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.

In the remainder of this section, we 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 6. 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 $(t, \pi) \in \eta$, the jump of the process $(\rho_t, t \in \mathbb{R}_+)$ can be described by

$$\gamma_n(\pi)(\gamma_n(\rho_{t-})) = \gamma_n(\rho_t). \tag{5.3}$$

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 [86] 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 [43]. A lookdown model with a reproduction mechanism that is different in the case with simultaneous multiple reproduction events is studied by Birkner et al. [13]. 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.

6 The Ξ-lookdown model

The population model from Section 5 will be driven by a Poisson random measure η on $(0,\infty) \times \mathcal{P}$ as in Schweinsberg [86], Bertoin [7], and Foucart [43].

To define this Poisson random measure, we briefly recall Kingman's correspondence. For a full account, see e. g. [7, 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 $|x|_p = \left(\sum_{i \in \mathbb{N}} x_i^p\right)^{1/p}$. 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 $\int_\Delta \nu(dx)\kappa(x,\cdot)$ for some distribution ν 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

$$\Xi = \Xi_0 + \Xi\{0\}\delta_0. \tag{6.1}$$

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

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

Let η be a Poisson random measure on $(0, \infty) \times \mathcal{P}$ with intensity $dt\ H_\Xi(d\pi)$. In Section 7.2, we will see that η satisfies a. s. condition (5.1), see also Remark 6.1 below. Hence, we can define the population model from Section 5 from almost every realization of η and every $\rho_0 \in \mathbb{R}^{\mathbb{N}^2}$.

From the description around equation (5.3) in Section 5 and the properties of Poisson random measures, it follows that each of the processes $(\gamma_n(\rho_t), t \in \mathbb{R}_+)$, $n \in \mathbb{N}$, is Markov, hence also the process $(\rho_t, t \in \mathbb{R}_+)$. 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 (7.4) and (7.3) in Section 7.2.

Remark 6.1. The quantity λ_π is the coagulation rate q_π in Section 4.2.1 of Bertoin [7]. It is related to the quantity $\lambda_{n;k_1,\ldots,k_r;s}$ from Schweinsberg [86] 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 (7.4) and (7.3) with equation (11) in [86]. In particular, equation (18) in [86] implies that η satisfies a. s. condition (5.1).

In the next proposition, we state a martingale problem for the process $(\rho_t, t \in \mathbb{R}_+)$ from Section 5, driven by the Poisson random measure η.

Recall the set $\mathcal{C}$ from Section 4. For $\phi \in \mathcal{C}$ and $\rho \in \mathbb{R}^{\mathbb{N}^2}$, we write

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

Proposition 6.2. *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

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

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 6.2 follows from the discussion above and the description of the process $(\gamma_n(\rho_t), t \in \mathbb{R}_+)$ in Section 5. As in [46], 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).

6.1 Properties of the genealogy at a fixed time

We assume that the process $(\rho_t, t \in \mathbb{R}_+)$ is obtained from the Poisson random measure η and a $\mathbb{R}^{\mathbb{N}^2}$-valued random variable ρ_0. To apply Theorem 4.2, we need exchangeability of the random variable ρ_t for each $t \in \mathbb{R}_+$.

Proposition 6.3. *Let $t \in \mathbb{R}_+$ and assume that ρ_0 is exchangeable. Then ρ_t is exchangeable.*

Remark 6.4. 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 [86, Section 3] with the Poisson process construction from Sections 5 and 6 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 [43] in one-sided time. We also remark that preservation of exchangeability in the lookdown model is studied in e. g. [13, 27, 28].

To prove Proposition 6.3, we show in Lemma 6.5 below that exchangeability is preserved in single reproduction events. Then we construct the state ρ_t, restricted to the first $n \in \mathbb{N}$ particles, from the 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 from the end of Section 5.

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

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

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 the space $\mathcal{P}_n$ of partitions of $[n]$ or in $\mathbb{R}^{n^2}$ is called exchangeable if its distribution is invariant under the action of S_n.

Lemma 6.5. *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 6.5 can be seen as a generalization of Lemma 4.3 of Bertoin [7].

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

$$\pi'(i) = f(\pi')\,(p(\pi')(p(i))) \tag{6.2}$$

for all $\pi' \in \mathcal{P}_n$ and $i \in [n]$. This equation holds as $\pi'(i)$, by its definition in Section 5, 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 the definition (5.2) of the transformation on $\mathbb{R}^{n^2}$ associated with each element of $\mathcal{P}_n$, equation (6.2) implies

$$\pi'(\rho') = p\,(p(\pi')\,(f(\pi')(\rho'))) \tag{6.3}$$

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

By assumption, $f(\pi)$ and π are equal in distribution. As π and ρ are independent, $f(\pi)(\rho)$ and ρ 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. This implies that $\pi(\rho)$ and $p(\pi)\,(f(\pi)(\rho))$ are equal in distribution. The assertion follows from equation (6.3). □

Proof of Proposition 6.3. Let $n \in \mathbb{N}$. For $s \in \mathbb{R}_+$, we define the map

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

where $\underline{\underline{2}}_n = 2(\mathbf{1}\{i \neq j\})_{i,j\in[n]}$. 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 in Section 5, 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 6.5 implies that $\gamma_n(\rho_t)$ is exchangeable. The assertion follows as the distribution of ρ_t is determined by the finite dimensional distributions. □

For the application of Theorem 4.2, it is of interest whether the states ρ_t are a. s. dust-free. Proposition 6.6 formulates the criterion from [86, Proposition 30] in our present context. We call the finite measure Ξ on Δ dust-free if

$$\Xi\{0\} > 0 \quad \text{or} \quad \int |x|_1\, |x|_2^{-2}\, \Xi_0(dx) = \infty. \tag{6.4}$$

Proposition 6.6. *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 6.4 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 [86, Proposition 30]. □

7 Decomposition of the genealogical distances

To apply Theorem 4.2(i), we need to describe also 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 this process in Subsection 7.1 for a deterministic point measure η that drives the population model. In Subsection 7.2, we let η again be the Poisson random measure.

7.1 The deterministic construction

Let η be a simple point measure on $(0, \infty) \times \mathcal{P}$ as in Section 5. 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 Section 5. We set

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

for $t \in \mathbb{R}_+$ and $i, j \in \mathbb{N}$. Then (r_0, v_0) can be thought of as a decomposition of the distance matrix ρ_t in the sense of Section 2.

Remark 7.1. Consider the following change (compared to the beginning of Section 5) 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 (7.2) below ensures that the times at which the particles on a fixed level are replaced do not accumulate.

Analogously to Section 5, 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 the beginning of Section 5. In order that equation (7.1) 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

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

and

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

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 $[n]$. The semi-partition $\varsigma_n(\pi)$ captures the effect on $\gamma_n(r_{t-}, v_{t-})$ of a reproduction event $(t, \pi) \in \eta$:

$$\varsigma_n(\pi)(\gamma_n(r_{t-}, v_{t-})) = \gamma_n(r_t, v_t) \tag{7.1}$$

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

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

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

$$\eta((0,t] \times \hat{\mathcal{P}}^n) < \infty \quad \text{for all } t \in (0, \infty) \text{ and } n \in \mathbb{N}. \tag{7.2}$$

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 $\eta((t, t+s] \times \hat{\mathcal{P}}^n) = 0$.

7.2 Stochastic evolution

Now let η be defined as the Poisson random measure from Section 6 whose distribution is characterized by some finite measure Ξ on Δ. Consider the population model from Section 5 driven by the Poisson random measure η, with the initial state 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. 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

$$\begin{aligned}\lambda_{n,\sigma} =& H_\Xi(\varsigma_n^{-1}(\sigma)) \\ =& \int_\Delta \kappa(x, \varsigma_n^{-1}(\sigma))\, |x|_2^{-2}\, \Xi_0(dx) + \Xi\{0\} \sum_{1\le i<j} \mathbf{1}\{K_{i,j} \in \varsigma_n^{-1}(\sigma)\} \\ =& \int_\Delta \sum_{\substack{i_1,\ldots,i_\ell \in \mathbb{N} \\ \text{pairwise distinct}}} x_{i_1}^{k_1} \cdots x_{i_\ell}^{k_\ell} (1-|x|_1)^{n-k_1-\ldots-k_\ell} |x|_2^{-2} \Xi_0(dx) \\ &+ \Xi\{0\}\, \mathbf{1}\{\ell = 1, k_1 = 2\} + \infty\, \mathbf{1}\{\Xi\{0\} > 0, \ell = 1, k_1 = 1\} \end{aligned} \tag{7.3}$$

where $\ell = \#\sigma$, and $k_1, \ldots, k_\ell \ge 1$ are the sizes of the subsets in σ in arbitrary order, and Ξ_0 is defined as in (6.1). For the last equality, we consider the paintbox partition π associated with $x \in \Delta$: With the notation from the beginning of Section 6, 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 6.1, satisfy

$$\lambda_\pi = H_\Xi(\gamma_n^{-1}(\pi)) = H_\Xi(\bigcup_\sigma \{\varsigma_n^{-1}(\sigma)\}) = \sum_\sigma \lambda_{n,\sigma}, \tag{7.4}$$

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 (7.4), we also use the restriction map $\gamma_n : \mathcal{P} \to \mathcal{P}_n$. From equations (7.3) and (7.4), 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 asserted in Section 6. The condition (6.4) 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 (7.2). 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 (7.3).

Remark 7.2. 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

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

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 process $((r_t, v_t), t \in \mathbb{R}_+)$ from Subsection 7.1, 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. 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 $(\gamma_n(r_t, v_t), t \in \mathbb{R}_+)$ in Section 7.1, we deduce the next proposition.

Proposition 7.3. *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 7.1. The operator $\hat{A}_{\mathrm{repr}}$ stands for the jumps of the decomposed genealogical distances in reproduction events which are described by equation (7.1).

Let now $(\rho_t, t \in \mathbb{R}_+)$ be the process defined from η and some $\mathfrak{U}$-valued random variable ρ_0 that is independent of η. Then $\rho_t \in \mathfrak{U}$ for all $t \in \mathbb{R}_+$ by Remark 5.2. Assume

$$(r_0, v_0) = \beta(\rho_0). \tag{7.5}$$

Then the construction in Section 7.1 and the definition of the map α in Section 2 yield $(r_t, v_t) \in \hat{\mathfrak{U}}$ and $\rho_t = \alpha(r_t, v_t)$ for all $t \in \mathbb{R}_+$. The following proposition states that the decomposition (r_t, v_t) of the semi-ultrametric ρ_t is the one given by the map β from Section 2.

Proposition 7.4. *Assumption (7.5) implies $(r_t, v_t) = \beta(\rho_t)$ a. s. for each $t \in \mathbb{R}_+$.*

Corollary 7.5. *The process $(\beta(\rho_t), t \in \mathbb{R}_+)$ solves the martingale problem $(\hat{A}, \hat{\mathcal{C}})$.*

Proof. This is immediate from Propositions 7.3 and 7.4. □

To prove Proposition 7.4, we use the following lemma.

Lemma 7.6. *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 6.4. We work on the event of probability 1 on which $v_t(i) > 0$, condition (7.2) is satisfied, and for each $s \in (0, t) \cap \mathbb{Q}$, the partition $\Pi_s^{(t)}$ contains infinitely many blocks if it contains singletons. This is indeed an event of probability 1 as $\lambda_{1,\{\{1\}\}} < \infty$ and by Kingman's correspondence.

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 7.1 and as $\eta((0,t] \times \hat{\mathcal{P}}^i) < \infty$ by condition (7.2) which means that the reproduction events in which particles on levels not larger than i reproduce do not accumulate.

Moreover, by condition (7.2), 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 that satisfies $\frac{1}{2}\rho_t(i,k) = v_t(i)$. □

Proof of Proposition 7.4. Let $t \in (0,\infty)$. From the definitions of the reproduction events in Section 5 and of the quantity $v_t(i)$ in Section 7.1, 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 $A_s(t,i)$ a time s. The definitions Υ in Section 2 and of ρ_t in Section 5 imply $0 \leq v_t(i) \wedge t \leq \Upsilon(\rho_t)(i) \wedge t$ for all $i \in \mathbb{N}$.

In the case that Ξ is dust-free, the assertion follows as $\Upsilon(\rho_t) = 0$ a. s. by Proposition 6.6.

Now we assume that Ξ is not dust-free. Let $i \in \mathbb{N}$. Lemma 7.6 yields $\Upsilon(\rho_t)(i) \leq v_t(i)$ a. s. on the event $\{v_t(i) < t\}$. On $\{v_t(i) \geq t\}$, the exchangeable partition $\Pi_t^{(t)}$, defined in Remark 6.4, contains the singleton block $\{A_0(t,i)\}$, and it follows by Kingman's correspondence that it has infinitely many blocks a. s. on $\{v_t(i) \geq t\}$. A. s. on $\{v_t(i) \geq t\}$, all particles at time zero survive until time t by Remark 5.1. Hence, as $\Upsilon(\rho_0) = v_0$ by assumption (7.5),

$$\begin{aligned} v_t(i) &= t + v_0(A_0(t,i)) = t + \tfrac{1}{2} \inf_{j \in \mathbb{N} \setminus \{A_0(t,i)\}} \rho_0(A_0(t,i), j) \\ &= \tfrac{1}{2} \inf_{j \in \mathbb{N} \setminus \{i\}} \rho_t(i,j) = \Upsilon(\rho_t)(i) \quad \text{a. s. on } \{v_t(i) \geq t\}. \end{aligned}$$

□

8 Tree-valued Fleming-Viot processes

In this section, we apply Theorem 4.2 to the process $(\rho_t, t \in \mathbb{R}_+)$ from Section 6. We call all the image processes in Theorem 4.2 tree-valued Fleming-Viot processes. To distinguish them, we also call them $\mathbb{U}$-, $\hat{\mathbb{U}}$, and $\mathcal{U}^{\text{erg}}$-valued Ξ-Fleming-Viot processes. In this section, we also show uniqueness of the martingale problems for tree-valued Fleming-Viot processes.

8.1 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 6 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 $\chi_t = \psi(\rho_t)$. Remark 3.14 ensures that $\psi(\rho_0) = \chi$ a. s. By Theorem 4.2 and Propositions 6.2, 6.3, and 6.6, 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 6.2. The martingale problem (B, Π) is a generalization of the martingale problem in Theorem 1 of Greven, Pfaffelhuber, and Winter [46].

Proposition 8.1. *The martingale problem (B, Π) is unique.*

Here uniqueness means that the finite-dimensional distributions of is solutions are uniquely specified when the initial state is given.

Proof. We use a function-valued dual process. This method is applied in the context of tree-valued Fleming-Viot processes in [25], another dual process is used in [46]. 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

$$\pi(\phi)(\rho) = \phi(\pi(\rho)).$$

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 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 $\underline{\underline{2}}_n = 2(\mathbf{1}\{i \neq j\})_{i,j \in [n]}$.

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

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

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 $\chi' \in \mathbb{U}$.

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 [36] 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 [36]. □

8.2 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 6, 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 $t \in \mathbb{R}_+$. By Propositions 3.3 and 3.12, the initial state satisfies $\hat{\chi}_0 = \chi$ a. s.

If Ξ is not dust-free, then by Theorem 4.2 and Propositions 6.3 and 7.3, the process $(\hat{\chi}_t, t \in \mathbb{R}_+)$ is Markovian and 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 7.3. Also the martingale problem $(\hat{B}, \hat{\Pi})$ is unique, the proof is analogous to Proposition 8.1.

If Ξ is dust-free, then for each $t \in \mathbb{R}_+$, the marked metric measure space $\hat{\chi}_t$ is a. s. dust-free, hence $\hat{\chi}_t$ is given a. s. by the associated metric measure space χ_t, cf. Remark 4.4. In this case, the process $(\hat{\chi}_t, t \in \mathbb{R}_+)$ solves a martingale problem that is similar to (B, Π).

8.3 Processes with values in the space of distance matrix distributions

Let $(\hat{\chi}_t, t \in \mathbb{R}_+)$ be the process from Section 8.2, where Ξ is a general finite measure on the simplex Δ. We define a $\mathcal{U}^{\text{erg}}$-valued Ξ-Fleming-Viot process $(\nu_t, t \in \mathbb{R}_+)$ with initial state $\alpha(\nu^{\hat{\chi}_0}) \in \mathcal{U}^{\text{erg}}$ by $\nu_t = \alpha(\nu^{\hat{\chi}_t})$. Again by Theorem 4.2 and Propositions 6.2 and 6.3, it follows that $(\nu_t, t \in \mathbb{R}_+)$ is Markovian and solves the martingale problem $(C, \mathscr{C})$, where the generator C is defined by $C\Psi(\nu) = \nu(A\phi)$ for all $\nu \in \mathcal{U}^{\text{erg}}$, $\phi \in \mathcal{C}$, and $\Psi \in \mathscr{C} : \nu' \mapsto \nu'\phi$. Here the generator A is defined as in Proposition 6.2. Uniqueness of the martingale problem follows analogously to Proposition 8.1, we use the same $\mathcal{C}_n$-valued dual process.

Remark 8.2. The martingale problems characterize only versions of the processes in Subsections 8.1 – 8.3. Moreover, Proposition 3.16 shows $\beta(\rho_t) \in \hat{\mathfrak{D}}^*$ (and in the dust-free case also $\rho_t \in \mathfrak{D}^*$ by Corollary 3.17) only for a fixed t (or countably many t) on an event of probability one. At the other times, we do not exclude in the present chapter that $\hat{\psi}(\beta(\rho_t))$ is just the arbitrary state in the definition of $\hat{\psi}$. By techniques specific to the lookdown model, it is shown in Chapter 3 that the aforementioned assertions on ρ_t also hold simultaneously for all $t \in \mathbb{R}_+$ on an event of probability one (see Theorems 3.1(i) and 3.9(i), and Remarks 4.4 and 4.13 in Chapter 3).

9 Some semigroup properties

In this section, we use the lookdown construction to study Feller continuity of tree-valued Ξ-Fleming-Viot processes, and to show 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 8.

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 8.2 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.

Proposition 9.1 states the Feller continuity of a $\hat{\mathbb{U}}$-valued Ξ-Fleming-Viot process, namely that its semigroup preserves the set of bounded continuous functions.

Proposition 9.1. *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. We fix $t \in \mathbb{R}_+$, and we denote by $\mathcal{N}$ the space of simple point measures on $(0,\infty) \times \mathcal{P}$. Recall the deterministic construction from Section 7.1 and denote by $g : \mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}} \times \mathcal{N} \to \mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}}$ the function that maps the initial state (r_0, v_0) and the simple point measure η to (r_t, v_t). Clearly, the function g is continuous in $\mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}}$. Let $\phi \in \hat{\mathcal{C}}_n$. By definition of the marked Gromov-weak topology and dominated convergence, also the function

$$h : \hat{\mathbb{U}} \to \mathbb{R}, \quad \chi \mapsto \int \nu^\chi(d(r,v)) \int \mathbb{P}(\eta \in d\eta')\phi \circ g((r,v),\eta')$$

is continuous, where η is the Poisson random measure from Section 6.

For $\chi \in \hat{\mathbb{U}}$, under $\mathbb{P}_\chi$, let (r_0, v_0) be a $\hat{\mathfrak{U}}$-valued random variable with distribution ν^χ. Let (r_t, v_t) be the $\hat{\mathfrak{U}}$-valued random variable defined from (r_0, v_0) and Ξ as in Section 7.2. By construction, $h(\chi) = \mathbb{E}_\chi[\phi(r_t, v_t)]$. Let $\Phi \in \hat{\Pi}$ be the marked polynomial associated with ϕ. By Propositions 7.4 and 4.7, and as we may assume $\hat{\psi}(r_t, v_t) = \hat{\chi}_t$ a. s., it follows that $h(\chi) = E_\chi[\Phi(\hat{\chi}_t)]$, hence $\chi \mapsto \mathbb{E}_\chi[\Phi(\hat{\chi}_t)]$ is continuous for each marked polynomial $\Phi \in \hat{\Pi}$. The assertion follows as the set $\hat{\Pi}$ of marked polynomials is convergence determining and by definition of convergence in distribution in $\hat{\mathbb{U}}$. □

Using the lookdown construction, we will also show the following lemma.

Lemma 9.2. *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 $\hat{\Pi}$.*

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 Chapter 5, we note two corollaries of Lemma 9.2. The first corollary 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 9.3. *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 9.2 and the definition of $\hat{L}$. As $(\hat{\chi}_t, t \in \mathbb{R}_+)$ solves the martingale problem $(\hat{B}, \hat{\Pi})$ from Section 8.2,

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

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 $\hat{L}$. □

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 [36, Chapter 1] for the definitions.

Corollary 9.4. *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 [36], using Lemma 9.2 and Corollary 9.3. □

Proof of Lemma 9.2. Let the space $\mathcal{N}$ be defined as in the proof of Proposition 9.1. Let $t \in \mathbb{R}_+$ and $n \in \mathbb{N}$. Note that in the construction in Sections 5 and 7.1, 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 the Poisson random measure from Section 6. 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. As in the proof of Proposition 9.1, Propositions 7.4 and 4.7 imply $\mathbb{E}_\chi[\Phi(\hat{\chi}_t)] = \nu^\chi f$ for all $\chi \in \hat{\mathbb{U}}$. Hence, $\chi \mapsto \mathbb{E}_\chi[\Phi(\hat{\chi}_t)]$ is in $\hat{\Pi}$. □

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 8.3. 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 8.1. Continuity properties analogous to Proposition 9.1 also hold.

10 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 7.2. From this, we deduce that also the tree-valued Ξ-Fleming-Viot process from Section 8.2 converges to equilibrium. In the same way, it can be shown that the other processes from Section 8 converge to equilibrium.

We define stationary processes and use a coupling argument. Analogously to Section 6, 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 in two-sided time (indexed by $\mathbb{R}$) where the reproduction events and the ancestral levels $\bar{A}_s(t, i)$ are defined as in Section 5. 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 6.4, or from Proposition 6.3.

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. By the construction in Section 5, 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. As $\bar{\rho}_t$ can also be obtained from $\bar{\rho}_0$ and η as in Section 5, 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, it follows that

$$|\mathbb{E}[\phi(\beta(\rho_t))] - \mathbb{E}[\phi(\beta(\bar{\rho}_0))]| \leq 2 \sup |\phi|\, \mathbb{P}(\max_{i,j \in [n]} \bar{\rho}_0(i, j) \geq 2t) \to 0 \quad (t \to \infty) \tag{10.1}$$

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 [46]. 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 8.2. As in Proposition 4.7, 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 (10.1) now yields that $\mathbb{E}[\Phi(\chi_t)]$ converges to $\mathbb{E}[\Phi(\bar{\chi}_0)]$ as $t \to \infty$ for all marked polynomials $\Phi \in \hat{\Pi}$. Using that the set $\hat{\Pi}$ is convergence determining in $\hat{\mathbb{U}}$, we deduce the following proposition.

Proposition 10.1. *The $\hat{\mathbb{U}}$-valued random variable χ_t converges in distribution to a Ξ-coalescent measure tree as $t \to \infty$.*

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 [46, Theorem 3], duality is used to show that the tree-valued Fleming-Viot process converges to an equilibrium. In [28, Theorem 4.1], convergence to stationarity of measure-valued Fleming-Viot processes is also proved by a coupling argument.

11 Proofs

In Section 11.1, we prove that the maps ψ and $\hat{\psi}$ from Section 3.4 are measurable. In Section 11.2, we prove Proposition 3.16 which is central for the proof of Theorem 1.2. We use the coupling characterization of the Prohorov metric, see e. g. [36, Theorem 3.1.2]: The Prohorov distance between two probability measures μ and μ' on the Borel sigma algebra on a separable metric space (Z, d^Z) is given by

$$d_{\mathrm{P}}^Z(\mu,\mu') = \inf_{\nu} \inf\{\varepsilon > 0 : \nu\{(x,y) \in Z^2 : d^Z(x,y) > \varepsilon\} < \varepsilon\}, \tag{11.1}$$

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

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

We show measurability of the map $\hat{\psi}$. Measurability of ψ follows along the same lines.

We use the marked Gromov-Prohorov metric d_{mGP} which metrizes the marked Gromov-weak topology on $\hat{\mathbb{M}}$, see [24]. It is defined by

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

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 $d^{Z\times\mathbb{R}_+}((z,v),(z',v')) = d^Z(z,z') \vee |v-v'|$. 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

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

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.

Remark 11.1. Clearly, the map $\hat{\psi}_n$ is continuous. For convenience, we state a proof. W. l. o. g. we can assume that $\hat{\mathfrak{D}}$ is endowed with the metric d that is given by

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

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 ν 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 $\nu(\cdot \times \hat{\mathfrak{D}}) = \nu^{\hat{\psi}_n(r,v)}$ and $\nu(\hat{\mathfrak{D}} \times \cdot) = \nu^{\hat{\psi}_n(r',v')}$. 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 (11.1) implies

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

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

Proof of Proposition 3.8. 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 $\lim_{n\to\infty} d_{\mathrm{mGP}}(\hat{\psi}(r,v), \hat{\psi}_n(\gamma_n(r,v))) = 0$. Using Remark 11.1 and that $\hat{\mathfrak{D}}^*$ is a measurable subset of $\hat{\mathfrak{D}}$, we deduce measurability of $\hat{\psi}$. □

11.2 Construction of the marked metric measure space in the sampling representation

We give three proofs of Proposition 3.16 that build on a common part, namely statement (11.4) 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 (11.4) 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 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 from Section 2.

Proof of Proposition 3.16. 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 $0 < h_1^{(\varepsilon)} < h_2^{(\varepsilon)} < \ldots$ that increases to infinity and that satisfies

$$h_1^{(\varepsilon)} < \varepsilon, \quad h_{n+1}^{(\varepsilon)} - h_n^{(\varepsilon)} < \varepsilon,$$

and

$$\mathbb{P}(v(i) = h_n^{(\varepsilon)}) = 0 \tag{11.2}$$

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

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

by definition of Υ and ultrametricity, it follows that $i \sim^\varepsilon k$.

Note that the definitions in Section 2 imply

$$r(i,j) = (\tfrac{1}{2}\rho(i,j) - v(i) + \tfrac{1}{2}\rho(i,j) - v(j))\, \mathbf{1}\{i \neq j\} < 2\varepsilon \tag{11.3}$$

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:

$$\text{A. s., the partition of } \mathbb{N} \text{ given by } \sim^\varepsilon \text{ contains no singleton blocks.} \tag{11.4}$$

For each $i, n \in \mathbb{N}$ the sequence $(\mathbf{1}\{v(j) \in I_n^\varepsilon, \rho(i,j)/2 \in I_n^\varepsilon\}, j \in \mathbb{N} \setminus \{i\})$ 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 (11.2) 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 (11.4).

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

$$m^\varepsilon = \sum_{i \in M^\varepsilon} |B^\varepsilon(i)| \delta_{(i,v(i))}. \tag{11.5}$$

(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 $k \le \ell$, 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 each $k \le \ell$ and $i \in M^{\varepsilon_k}$, there exist a. s. $i_1, i_2, \ldots \in M^{\varepsilon_\ell}$ with

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

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

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

Hence, a coupling K of the probability measures m^{ε_k} and m^{ε_ℓ} on $X \times \mathbb{R}_+$ is given a. s. by

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

for $i \in M^{\varepsilon_k}$ and $j \in M^{\varepsilon_\ell}$ with $j \in B^{\varepsilon_k}(i)$. By construction, $i \sim^{\varepsilon_k} j$, hence $|v(i) - v(j)| < \varepsilon_k$ and $r(i,j) < 2\varepsilon_k$. By definition of m^ε and the coupling characterization of the Prohorov metric (11.1), this implies

$$d_{\mathrm{P}}^{X\times\mathbb{R}_+}(m^{\varepsilon_k}, m^{\varepsilon_\ell}) \le 2\varepsilon_k \tag{11.6}$$

a. s. for all $k \le \ell$, 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 (11.1) implies for each $k \in \mathbb{N}$

$$\begin{aligned}
&d_{\mathrm{P}}^{X\times\mathbb{R}_+}(m_n^{\varepsilon_\ell}, m^{\varepsilon_\ell}) \\
&\le K'\{(y,y') \in (Z \times \mathbb{R}_+)^2 : y \ne y'\} \\
&\le m^{\varepsilon_\ell}\{(j, v(j)) : j \in M^{\varepsilon_\ell}, j > k\} + K'\{((i,v(i)),(j,v(j))) : i,j \in M^{\varepsilon_\ell}, i \ne j, j \le k\} \\
&= \sum_{\substack{j \in M^{\varepsilon_\ell} \\ j > k}} |B^{\varepsilon_\ell}(j)| + \sum_{\substack{j \in M^{\varepsilon_\ell} \\ j \le k}} \big||B^{\varepsilon_\ell}(j)|_n - |B^{\varepsilon_\ell}(j)|\big| \quad \text{a. s.}
\end{aligned}$$

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

$$\lim_{n\to\infty} d_{\mathrm{P}}^{X\times\mathbb{R}_+}(m_n^{\varepsilon_\ell}, m^{\varepsilon_\ell}) = 0 \quad \text{a. s.} \tag{11.7}$$

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 $X \times \mathbb{R}_+$. (Its first component corresponds to the measure on the starting points of the external branches that end in one of the first n leaves of T such that leaves are weighted according to the multiplicity given by the semi-metric ρ.) By (11.6),

$$d_{\mathrm{P}}^{X\times\mathbb{R}_+}(m_n^{\varepsilon_\ell}, m_n) \le 4\varepsilon_\ell \quad \text{a. s.} \tag{11.8}$$

for all $n, \ell \in \mathbb{N}$. From (11.6), (11.7), and (11.8), we obtain

$$m = \text{w-}\lim_{n\to\infty} m_n \quad \text{a. s.} \tag{11.9}$$

This shows the assertion. □

In the first proof of Proposition 3.16, we constructed the sampling measure "by hand". The idea for the second proof is to consider, in the metric completion of the coalescent tree associated with ρ, the closure of the subspace of the starting vertices of the external branches that end in the leaves labeled by the odd integers, and to show that this complete subspace contains the sequence of the starting vertices of the external branches associated with the even integers. To this aim, we use (11.4) from the first proof. The de Finetti theorem then yields the sampling measure for this exchangeable sequence. For a related result, see also Forman, Haulk, and Pitman [42], where trees are embedded into ℓ_1.

Remark 11.2. The second proof given below goes in a direction that is similar to the argument in Section 7 of [40] 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 [40] holds for the embedding of $\Gamma(\mathbf{T}^-)$ and $\Gamma(\mathbf{T}^+)$ into $\Gamma(\mathbf{T})$ can be seen 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 [40]. Another variant of the construction of the sampling measure is given in Remark 11.3 below.

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 [40, Section 7], which corresponds to ρ below. This equality can also be deduced from Theorem 3.18, Remark 3.7, as $\psi(r)$ is a. s. constant by the ergodicity assumption in [40], and from the Gromov reconstruction theorem.

Second proof of Proposition 3.16. 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 (11.3) and (11.4), 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

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

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 $2,4,6,\dots$ 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 $B_2,\dots,B_{2n}$ be some finite intersections of such balls. Note that $\{2\in B_2,\dots,2n\in B_{2n}\}$ 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 $\mu^1_n:=n^{-1}\sum_{i=1}^n\delta_{2i}$ 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 $\varphi(\mu^1_n)$ 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 $d^{X_1\times\mathbb{R}_+}((x',v'),(x'',v''))=r(x',x'')\vee|v'-v''|$. □

In the third proof, we use (11.4) to obtain on an event of probability 1 a labeling of the elements of the completion of $(\mathbb{N},r)$ that is determined by the equivalence class $[\![r]\!]$ of r with respect to finite permutations.

Remark 11.3. The sampling measure μ on the real tree $\mathbf{S} = \Gamma(\mathbf{T})$ from Remark 11.2 can also be obtained as in the third proof given below. Indeed, $i \in \mathbb{N}$ corresponds to the point of attachment $\Pi(i)$, and (Y, r) to the closure of $\{\Pi(i) : i \in \mathbb{N}\}$ in $\Gamma(\mathbf{T})$ in [40, Section 7]. When we label the elements of the closure of $\{\Pi(i) : i \in \mathbb{N}\}$ as in the proof below, we obtain a complete and separable metric space Y' that is a. s. constant by the ergodicity assumption in [40]. The de Finetti theorem, applied to the exchangeable sequence $\Pi(1), \Pi(2), \ldots$ in Y', then yields the sampling measure μ.

Third proof of Proposition 3.16. Let $(r, v) = \beta(\rho)$. We say that sequences $(y_j^1)_{j\in\mathbb{N}}$ and $(y_j^2)_{j\in\mathbb{N}}$ in $(\mathbb{N}, r)$ are equivalent if $\lim_{j\to\infty} r(y_j^1, y_j^2) = 0$. Let Y be the set of the equivalence classes of those Cauchy sequences $(y_j)_{j\in\mathbb{N}}$ in $(\mathbb{N}, r)$ that satisfy $y_{j+1} > y_j$ for all $j \in \mathbb{N}$. We endow Y with the metric induced by r, namely $r(y^1, y^2) = \lim_{j\to\infty} r(y_j^1, y_j^2)$ for $y^1, y^2 \in Y$ with any representatives $(y_j^1)_{j\in\mathbb{N}}$ and $(y_j^2)_{j\in\mathbb{N}}$.

By (11.3) and (11.4), exchangeability of $\sim^\varepsilon$ therein, and Kingman's correspondence, there exists a. s. for each $i \in \mathbb{N}$ and $\varepsilon > 0$ an arbitrarily large integer j with $r(i, j) < \varepsilon$. For each Cauchy sequence $(y_j^1)_{j\in\mathbb{N}}$ in $(\mathbb{N}, r)$, we thus find an equivalent Cauchy sequence $(y_j^2)_{j\in\mathbb{N}}$ in $(\mathbb{N}, r)$ with $r(y_j^1, y_j^2) < 1/j$ and $y_{j+1}^2 > y_j^2$ for all $j \in \mathbb{N}$. Hence, we can identify each element i of $\mathbb{N}$ with the equivalence class (of Cauchy sequences) in Y whose representatives are equivalent to the Cauchy sequence $(i, i, \ldots)$ in $(\mathbb{N}, r)$. Then (Y, r) is a metric completion of $(\mathbb{N}, r)$.

Let $[\![r]\!]$ denote the equivalence class (of distance matrices) of r with respect to finite permutations, and let Λ be a regular conditional distribution of r given $[\![r]\!]$. Note that finite permutations of r preserve the equivalence classes that are the elements of Y, and also the distances between these elements of Y. That is, the complete and separable metric space (Y, r) is determined by $[\![r]\!]$, hence a. s. constant under $\Lambda([\![r]\!], \cdot)$ for a. a. $[\![r]\!]$. We claim that for a. a. $[\![r]\!]$ under $\Lambda([\![r]\!], \cdot)$, the sequence $1, 2, 3, \ldots$ is exchangeable in (Y, r). Then, the de Finetti theorem yields the a. s. existence of the weak limit of the probability measures $n^{-1}\sum_{i=1}^n \delta_i$ on (Y, r), which means that $r \in \mathfrak{D}^*$ a. s. Analogously, it can be seen that $(r, v) \in \hat{\mathfrak{D}}^*$ a. s. by considering the sequence $(i, v(i))_{i\in\mathbb{N}}$ in the space $Y \times \mathbb{R}_+$ and a regular conditional distribution of (r, v) given the equivalence class of (r, v) with respect to finite permutations.

Now we carry out the proof of the claimed exchangeability of $1, 2, 3, \ldots$. We fix $[\![r]\!]$ and hence (Y, r). Let $n, \ell \in \mathbb{N}$. For $i \in [n]$, $k \in [\ell]$, let $y^{i,k}$ be an element of Y. Let the Cauchy sequences $(y_j^{i,k})_{j\in\mathbb{N}}$ in $(\mathbb{N}, r)$ be representatives of the $y^{i,k}$. Here we can assume w. l. o. g. that $y_j^{i,k} > n$ for all $j \in \mathbb{N}$. Note that each $y^{i,k}$ has also the representative $(y_j^{i,k})_{j\in\mathbb{N}}$ in any $(\mathbb{N}, r')$ with $[\![r']\!] = [\![r]\!]$. Hence, as we fix $[\![r]\!]$, we can also fix the representatives $(y_j^{i,k})_{j\in\mathbb{N}}$ of the $y^{i,k}$. Let $(f^{i,k})$ be a collection of bounded continuous functions, and let p be a permutation of $[n]$. By exchangeability of r,

$$\int \Lambda([\![r]\!], dr') \prod_{i\in[n],k\in[\ell]} f^{i,k} \circ r'(i, y_j^{i,k}) = \int \Lambda([\![r]\!], dr') \prod_{i\in[n],k\in[\ell]} f^{i,k} \circ r'(p(i), y_j^{i,k}) \quad (11.10)$$

for a. a. $[\![r]\!]$ and $j \in \mathbb{N}$. As in the first part of the proof, we can a. s. associate with each $i \in [n]$ a Cauchy sequence $(z_j^i)_{j\in\mathbb{N}}$ in $(\mathbb{N}, r')$ that satisfies $\lim_{j\to\infty} r'(i, z_j^i) = 0$ and

$z^i_{j+1} > z^i_j$ for all $j \in \mathbb{N}$. Then,

$$\lim_{j\to\infty} r'(i, y^{i,k}_j) = \lim_{j\to\infty} r'(z^i_j, y^{i,k}_j),$$

and similarly for the integrand on the right-hand side of equation (11.10). As $i \in [n]$ is identified with the equivalence class of $(z^i_j)_{j\in\mathbb{N}}$ in Y, and as each semi-metric r' on $\mathbb{N}$ with $[\![r']\!] = [\![r]\!]$ induces the same metric $r' = r$ on Y, letting j tend to infinity on both sides of equation (11.10) yields

$$\int \Lambda([\![r]\!], dr') \prod_{i\in[n],k\in[\ell]} f^{i,k} \circ r'(i, y^{i,k}) = \int \Lambda([\![r]\!], dr') \prod_{i\in[n],k\in[\ell]} f^{i,k} \circ r'(p(i), y^{i,k})$$

by dominated convergence. By approximating indicator variables with the $f^{i,k}$, we deduce that

$$\int \Lambda([\![r]\!], dr') \prod_{i\in[n],k\in[\ell]} \mathbf{1}\{r'(i, y^{i,k}) < c_{i,k}\} = \int \Lambda([\![r]\!], dr') \prod_{i\in[n],k\in[\ell]} \mathbf{1}\{r'(p(i), y^{i,k}) < c_{i,k}\}$$

for any numbers $c_{i,k} > 0$. The assertion follows by the uniqueness lemma as the balls of the form $\{y \in Y : r(y, y^{i,k}) < c_{i,k}\}$ generate the Borel sigma algebra on (Y, r). □

12 Construction from the flow of bridges

A random non-decreasing right-continuous function $\tilde{F} : [0,1] \to [0,1]$ with exchangeable increments and $\tilde{F}(0) = 0$, $\tilde{F}(1) = 1$ is called a bridge. We view a bridge as a random variable with values in the space of càdlàg paths $[0,1] \to [0,1]$ which we endow with the Skorohod metric. The dual flow of bridges of Bertoin and Le Gall [8] is a collection $F = (F_{s,t}, s < t)$ of bridges that satisfies the following properties (see [8, 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 $s \le t$, 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 [10].

In [8, 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 [8, 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 [86].

For each $t \in \mathbb{R}$, there exists an event of probability 1 on which for all $s \le s' \in \mathbb{Q}_+$, 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

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

The assumption that for each $r < s$, the bridge $F_{r,s}$ is a. s. not the identity function implies that the infimum in the definition of $\tilde{\rho}_t(i,j)$ is a. s. not over the empty set.

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. [55, Theorem 6.3].

Analogously to Sections 8.2 and 8.3, for a finite measure Ξ on Δ, a stationary $\mathcal{U}^{\mathrm{erg}}$-valued Ξ-Fleming-Viot process $(\nu_t, t \in \mathbb{R})$ is given by $\nu_t = \alpha(\nu^{\hat{\psi} \circ \beta(\bar{\rho}_t)})$, where $(\bar{\rho}_t, t \in \mathbb{R})$ is defined as in Section 10. We note that a stationary $\mathcal{U}^{\mathrm{erg}}$-valued Ξ-Fleming-Viot process can be read off from the dual flow of bridges:

Theorem 12.1. *There exists a finite measure Ξ on Δ such that the process $(\tilde{\nu}_t, t \in \mathbb{R})$ is a version of a stationary $\mathcal{U}^{\mathrm{erg}}$-valued Ξ-Fleming-Viot process.*

For the proof of Theorem 12.1, we show that $(\tilde{\nu}_t, t \in \mathbb{R})$ is a Markov process and has the transition kernel of a $\mathcal{U}^{\mathrm{erg}}$-valued Ξ-Fleming-Viot process. In the following, we fix $u \in \mathbb{R}_+$.

First, we define for each finite measure Ξ on Δ a probability kernel Λ_Ξ from $\mathcal{U}$ to $\mathfrak{U}$ such that for each $\nu \in \mathcal{U}$, the distribution $\Lambda_\Xi(\nu, \cdot)$ is the distribution of a random variable ρ which we define as follows. Let ρ' be a random variable with distribution ν. Let ρ'' be an independent $\mathfrak{U}$-valued random variable that is distributed as the random ultrametric associated with a Ξ-coalescent. That is, ρ'' shall be distributed as the random variable $\bar{\rho}_u$ mentioned above, cf. Remark 6.4. We define a partition π of $\mathbb{N}$ such that i and j are in the same block of π if and only if $\rho''(i,j) < 2u$. Let $B_1(\pi), B_2(\pi), \ldots$ be the blocks of π, ordered increasingly according to their smallest element. For $i \in \mathbb{N}$, let $A(i)$ be the integer j such that $i \in B_j(\pi)$.

Then we set for $i, j \in \mathbb{N}$

$$\rho(i,j) = \begin{cases} \rho''(i,j) \wedge (2u) & \text{if } \rho''(i,j) < 2u \\ 2u + \rho'(A(i), A(j)) & \text{else.} \end{cases}$$

In the following, we also fix $t \in \mathbb{R}$.

Remark 12.2. Let $(\bar{\rho}_s, s \in \mathbb{R})$ be defined as in Section 10 from a measure Ξ, and let $\nu_t = \alpha(\nu^{\hat{\psi}\circ\beta(\bar{\rho}_t)})$. Note that Λ_Ξ is a regular conditional distribution of $\bar{\rho}_{t+u}$ given ν_t. This follows as $\bar{\rho}_{t+u} \wedge (2u)$ is independent of ν_t and $\bar{\rho}_t$, as $\bar{\rho}_{t+u}(i,j) = 2u + \bar{\rho}_t(A_t(t+u,i), A_t(t+u,j))$ for $i,j \in \mathbb{N}$ with $\bar{\rho}_{t+u}(i,j) \geq 2u$, as $A_t(t+u,i)$ can be read off from $\bar{\rho}_{t+u} \wedge (2u)$ like $A(i)$ can be read off from ρ'' in the definition of Λ_Ξ, and as ν_t is a regular conditional distribution of $\bar{\rho}_t$ given ν_t by Remark 3.23.

Lemma 12.3. *There exists a finite measure Ξ on Δ such that Λ_Ξ is a regular conditional distribution of $\tilde{\rho}_{t+u}$ given $\tilde{\nu}_t$. Moreover, $\tilde{\rho}_{t+u}$ is conditionally independent of $(\tilde{\nu}_s, s \leq t)$ given $\tilde{\nu}_t$.*

Proof. We claim that given the collection of bridges $(F_{r,s} : r < s \leq t)$, the random variable $\tilde{\rho}_{t+u}$ has conditional distribution $\Lambda_\Xi(\tilde{\nu}_t, \cdot)$ for some finite measure Ξ on Δ. By construction of $\tilde{\nu}_s$, this claim implies both assertions of the lemma.

We assume that the coalescent process $(\tilde{\pi}_s^{(t+u)}, s \in \mathbb{R}_+)$ and the associated ultrametric $\tilde{\rho}_{t+u}$ are constructed as above from F and a sequence $(V_i, i \in \mathbb{N})$ of independent uniformly distributed $[0,1]$-valued random variables that is independent of F. By [8, Theorem 1], there exists a finite measure Ξ on Δ such that $(\tilde{\pi}_s^{(t+u)}, s \in \mathbb{R}_+)$ is a version of a Ξ-coalescent.

Let $B_1(\tilde{\pi}_u^{(t+u)}), B_2(\tilde{\pi}_u^{(t+u)}), \ldots$ be the blocks of $\tilde{\pi}_u^{(t+u)}$ in increasing order according to their respective smallest element. For $i \in \mathbb{N}$, we define $\tilde{A}(i) = j$ where j is the integer such that $i \in B_j(\tilde{\pi}_u^{(t+u)})$. We define a sequence $V' = (V'_i, i \in \mathbb{N})$ analogously to equation (3) of [8]: For $i \in \mathbb{N}$ with $i \leq \#\tilde{\pi}_u^{(t+u)}$, we set $V'_i = F_{t,t+u}^{-1}(V_j)$, where j is any element of $B_i(\tilde{\pi}_u^{(t+u)})$. If the number of blocks $\#\tilde{\pi}_u^{(t+u)}$ is finite, we extend the sequence $(V'_i, i \leq \#\tilde{\pi}_u^{(t+u)})$ to $(V'_i, i \in \mathbb{N})$ using an independent sequence of independent uniform random variables on $[0,1]$.

Let $\ell \in \mathbb{N}$ and $0 \leq u_1 \leq u_2 \leq \ldots \leq u_\ell = u$. Repeated application of [8, Lemma 2] to the bridges $F_{t+u-u_1,t+u}, \ldots, F_{t+u-u_\ell,t+u-u_{\ell-1}}$ (similarly to [8, Corollary 1]) yields that V' is a sequence of independent $[0,1]$-valued uniformly distributed random variables that is also independent of $\tilde{\pi}_{u_1}^{(t+u)}, \ldots, \tilde{\pi}_{u_\ell}^{(t+u)}$. By construction and property (ii) of the dual flow of bridges, V' and $\tilde{\pi}_{u_1}^{(t+u)}, \ldots, \tilde{\pi}_{u_\ell}^{(t+u)}$ are also independent of $(F_{r,s} : r < s \leq t)$.

We define $\tilde{\rho}_t$ from F and the sequence V'. Then $\tilde{\rho}_t$ is conditionally independent of $\tilde{\rho}_{t+u} \wedge (2u)$ given the collection of bridges $(F_{r,s} : r < s \leq t)$. This follows from the above by the uniqueness lemma as for $i,j \in \mathbb{N}$ and $m = 1, \ldots, \ell$, $\{\tilde{\rho}_{t+u}(i,j) \leq u_m\}$ is, up to null events, the event that i and j are in the same block of $\tilde{\pi}_{u_m}^{(t+u)}$.

By construction, $\tilde{\nu}_t$ is a conditional distribution of $\tilde{\rho}_t$ given $(F_{r,s} : r < s \leq t)$. We also define the coalescent process $(\tilde{\pi}_s^{(t)}, s \in \mathbb{R}_+)$ from V' and F. Then $\tilde{\rho}_t$ is the associated ultrametric. For $i,j \in \mathbb{N}$ and $s \in \mathbb{R}_+$, the following events are equal up to null events:

$$\begin{aligned}&\{\tilde{\rho}_{t+u}(i,j) \leq 2(u+s)\} = \{i,j \text{ are in the same block of } \tilde{\pi}_{u+s}^{(t+u)}\}\\ &= \{\tilde{A}(i), \tilde{A}(j) \text{ are in the same block of } \tilde{\pi}_s^{(t)}\} = \{\tilde{\rho}_t(\tilde{A}(i), \tilde{A}(j)) \leq 2s\}.\end{aligned}$$

For the equality up to null events of the second and the third event, we use the definition of V' and property (i) of the dual flow of bridges. It follows that a. s.,

$$\tilde{\rho}_{t+u}(i,j) = \begin{cases} \tilde{\rho}_{t+u}(i,j) \wedge (2u) & \text{if } \tilde{\rho}_{t+u}(i,j) < 2u \\ 2u + \tilde{\rho}_t(\tilde{A}(i), \tilde{A}(j)) & \text{else.} \end{cases}$$

The claim follows as $\tilde{A}(i)$ can a. s. be read off from $\tilde{\rho}_{t+u} \wedge (2u)$ in the same way as $A(i)$ is read off from ρ'' in the definition of Λ_Ξ. □

To deduce Theorem 12.1, we use that $\tilde{\nu}_{t+u} \in \mathcal{U}^{\text{erg}}$ a. s.

Proof of Theorem 12.1. Let $t \in \mathbb{R}$ and $u > 0$. By Remark 3.21 and as the sequence V in the definition of $\tilde{\rho}_t$ is iid, $\tilde{\nu}_t \in \mathcal{U}^{\text{erg}}$ a. s. That $\tilde{\nu}_t$ is concentrated on the ergodic distributions can be seen directly or by an application of e. g. [56, Lemma 7.35]. By construction, $\tilde{\nu}_{t+u}$ is a regular conditional distribution of $\tilde{\rho}_{t+u}$ given $\tilde{\nu}_{t+u}$. By Corollary 3.24, $\tilde{\nu}_{t+u} = \zeta(\tilde{\rho}_{t+u})$ a. s., where $\zeta : \mathfrak{U} \to \mathcal{U}^{\text{erg}}$, $\rho \mapsto \nu^{\hat{\psi} \circ \beta(\rho)}$. Hence, by Lemma 12.3, there exists a finite measure Ξ on Δ such that $\Lambda_\Xi(\cdot, \zeta^{-1}(\cdot))$ is a regular conditional distribution of $\tilde{\nu}_{t+u}$ given $\tilde{\nu}_t$, and $\tilde{\nu}_{t+u}$ is conditionally independent of $(\tilde{\nu}_s, s \leq t)$ given $\tilde{\nu}_t$. The latter property is the Markov property of $(\tilde{\nu}_s, s \in \mathbb{R})$.

Let now $(\nu_s, s \in \mathbb{R})$ be a $\mathcal{U}^{\text{erg}}$-valued Ξ-Fleming-Viot process defined from $(\bar{\rho}_s, s \in \mathbb{R})$ as recalled in the beginning of this section. As in Theorem 4.2 (or alternatively, by an extension of Remark 12.2), the process $(\nu_t, t \in \mathbb{R})$ is Markovian. By Remark 12.2 and as $\nu_{t+u} = \zeta(\bar{\rho}_{t+u})$ by definition, $\Lambda_\Xi(\cdot, \zeta^{-1}(\cdot))$ is a regular conditional distribution also of ν_{t+u} given ν_t. This implies the assertion. □

List of notation

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

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, (pp. 26, 30, 33)

(Marked) distance matrices

$\mathfrak{U}$: space of semi-ultrametrics on $\mathbb{N}$, (p. 20)
$\hat{\mathfrak{U}}$: space of decomposed semi-ultrametrics on $\mathbb{N}$, (p. 21)
$\mathfrak{D}$, $\hat{\mathfrak{D}}$: spaces of (decomposed) semimetrics on $\mathbb{N}$ (pp. 20, 49)
α: map that retrieves the semi-ultrametric from a decomposed semi-ultrametric (p. 21)
$\beta : \mathfrak{U} \to \hat{\mathfrak{U}}$: decomposition map into the external branches and the remaining subtree (p. 21)
$\Upsilon(\rho)$: vector of the lengths of the external branches in the coalescent tree associated with ρ (p. 21)
$\mathcal{U}$: space of exchangeable distributions on $\mathfrak{U}$ (p. 28)

(Marked) metric measure spaces

$\mathbb{M}$: space of isomorphy classes of metric measure spaces (p. 21)
$\mathbb{U}$: space of isomorphy classes of ultrametric measure spaces (p. 22)
$\hat{\mathbb{M}}$, $\hat{\mathbb{U}}$: spaces of isomorphy classes of marked metric measure spaces (p. 22)
ν^χ: distance matrix distribution of $\chi \in \mathbb{M}$ (pp. 17, 22) or marked distance matrix distribution of $\chi \in \hat{\mathbb{M}}$ (pp. 17, 22)
$\mathcal{U}^{\text{erg}}$: space of distance matrix distributions (p. 28)
$\psi : \mathfrak{D} \to \mathbb{M}$, $\hat{\psi} : \mathfrak{D} \times \mathbb{R}_+^{\mathbb{N}} \to \hat{\mathbb{M}}$: construction of (marked) metric measure spaces (p. 25)
$\mathfrak{D}^*$, $\hat{\mathfrak{D}}^*$: sets of (marked) distance matrices with a good sampling measure (p. 25)
$\mathcal{C}_n$, $\mathcal{C}$, $\hat{\mathcal{C}}_n$, $\hat{\mathcal{C}}$: sets of bounded differentiable functions with bounded uniformly continuous derivative (p. 30)
Π: set of polynomials on $\mathbb{U}$ (p. 30)
$\hat{\Pi}$: set of marked polynomials on $\hat{\mathbb{U}}$ (p. 30)
$\mathscr{C}$: a set of test functions on $\mathcal{U}^{\text{erg}}$ (p. 30)

Partitions and semi-partitions

$\mathcal{P}$: Set of partitions of $\mathbb{N}$
$B_i(\pi)$: i-th block of a partition π (p. 33)
$\#\pi$: number of blocks of a partition π
$K_{i,j}$: partition of $\mathbb{N}$ that contains only $\{i, j\}$ and singleton blocks (p. 36)
$\mathcal{P}_n$: 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. 33)
$\hat{\mathcal{P}}^n$: Set of partitions of $\mathbb{N}$ in which the first n integers are not all in singleton blocks (p. 40)
$\mathcal{S}_n$ set of semi-partitions of $[n]$, associated transformations (pp. 39, 40)
$\Delta = \{x = (x_1, x_2, \ldots) : x_1 \geq x_2 \geq \ldots 0, |x|_1 \leq 1\}$
$\kappa(x, \cdot)$: paintbox distribution associated with $x \in \Delta$ (p. 35)

Genealogy in the lookdown model

η: point measure on $(0, \infty) \times \mathcal{P}$ that encodes the reproduction events, (pp. 33, 36)
$A_s(t, i)$: level of the ancestor at time s of the particle on level i at time t (p. 34)
$\rho_t(i, j)$: genealogical distance (p. 34)
(r_t, v_t): decomposed genealogical distance (p. 39)
$\Xi = \Xi_0 + \Xi\{0\}\delta_0$, equation (6.1)
H_Ξ: characteristic measure of η (p. 36)

Acknowledgments

The author thanks Götz Kersting, Anton Wakolbinger, and two anonymous referees for helpful comments. Partial support from the DFG Priority Programme 1590 "Probabilistic Structures in Evolution" is acknowledged.

Chapter 3

Pathwise construction of tree-valued Fleming-Viot processes

In a random complete and separable metric space that we call the lookdown space, we encode the genealogical distances between all individuals ever alive in a lookdown model with simultaneous multiple reproduction events. We construct families of probability measures on the lookdown space and on an extension of it that allows to include the case with dust. From this construction, we read off the tree-valued Ξ-Fleming-Viot processes and deduce path properties. For instance, these processes usually have a. s. càdlàg paths with jumps at the times of large reproduction events. In the case of coming down from infinity, the construction on the lookdown space also allows to read off a process with values in the space of measure-preserving isometry classes of compact metric measure spaces, endowed with the Gromov-Hausdorff-Prohorov metric. This process has a. s. càdlàg paths with additional jumps at the extinction times of parts of the population.

1 Introduction

Similarly to the measure-valued Fleming-Viot process that is a model for the evolution of the type distribution in a large neutral haploid population, a tree-valued Fleming-Viot process models the evolution of the distribution of the genealogical distances between randomly sampled individuals. The (neutral) tree-valued Fleming-Viot process is introduced in Greven, Pfaffelhuber, and Winter [46]. In Chapter 2, it is generalized to the setting with simultaneous multiple reproduction events. The lookdown model of Donnelly and Kurtz [27, 28] provides a pathwise construction of the measure-valued Fleming-Viot process and more general measure-valued processes. In this chapter, we give a pathwise construction of the tree-valued Fleming-Viot process from the lookdown model.

Let us sketch the lookdown model that we state in more detail in Section 2. The time axis is $\mathbb{R}_+$. In the population model, there are countably infinitely many levels which are labeled by $\mathbb{N}$. Each level is occupied by one particle at each time. As time evolves, the particles undergo reproduction events in which particles can increase their levels. We call a particle at a fixed instant in time an individual. We identify each element (t, i) of $\mathbb{R}_+ \times \mathbb{N}$ with the individual on level i at time t. From the genealogy that is determined by the reproduction events and from given genealogical distances between the individuals at time zero, we define the semi-metric ρ on $\mathbb{R}_+ \times \mathbb{N}$ of the genealogical distances between all individuals. We speak of the case with dust if each particle reproduces at finite rate. In the general case, only the rate at which a particle reproduces and has offspring on a given level is finite. In the case without dust, we introduce the lookdown space (Z, ρ) as the metric completion of $(\mathbb{R}_+ \times \mathbb{N}, \rho)$. We allow for simultaneous multiple reproduction events so that we can obtain any Ξ-coalescent as the genealogy at a fixed time [13, 28, 72, 79, 84, 86].

In Section 3, we state the central results of this chapter. Theorem 3.1 asserts that a. s. in the case without dust, the uniform measures $\mu_t^n = n^{-1} \sum_{i=1}^n \delta_{(t,i)}$ on the individuals on the first n levels at time t converge uniformly in compact time intervals to some probability measures $(\mu_t, t \in \mathbb{R}_+)$ in the Prohorov metric d_{P}^Z on the lookdown space,

$$\lim_{n\to\infty} \sup_{t\in[0,T]} d_{\mathrm{P}}^Z(\mu_t^n, \mu_t) \quad \text{a. s. for all } T \in \mathbb{R}_+. \tag{1.1}$$

We recall that a metric measure space (X, r, μ) is a triple that consists of a complete and separable metric space (X, r) and a probability measure μ on the Borel sigma algebra on (X, r). The Gromov-Prohorov distance between two metric measure spaces (X, r, μ) and (X', r', μ') is defined as

$$d_{\mathrm{GP}}((X, r, \mu), (X', r', \mu')) = \inf_{Y,\varphi,\varphi'} d_{\mathrm{P}}^Y(\varphi(\mu), \varphi'(\mu'))$$

where the infimum is over all isometric embeddings $\varphi : X \to Y$, $\varphi' : X' \to Y$ into complete and separable metric spaces Y. Two metric measure spaces are called isomorphic if their Gromov-Prohorov distance is zero, or equivalently, if there is a measure-preserving isometry between the supports of the measures. The Gromov-Prohorov distance is a complete and separable metric on the space $\mathbb{M}$ of isomorphy classes of metric measure spaces, it induces the Gromov-weak topology in which metric measure spaces converge if and only if the distributions of the matrices of the distances between iid samples (the so-called distance matrix distributions) converge weakly. For the theory of metric measure spaces, we refer to Greven, Pfaffelhuber, and Winter [45] and Gromov [48].

In Section 4, we read off the tree-valued Ξ-Fleming-Viot processes. As in [46, 50], we consider trees as metric measure spaces, as marked metric measure spaces, and as distance matrix distributions. While the lookdown model is used in Chapter 2 to characterize only versions of the tree-valued Ξ-Fleming-Viot processes (see Remark 8.2 in Chapter 2), we obtain the whole paths in the present chapter. Other than in Chapter 2, we do not use ultrametricity of the initial state for the techniques in the present chapter. Therefore, we speak for instance of an $\mathbb{M}$-valued Ξ-Fleming-Viot process when the initial state not necessarily corresponds to an ultrametric tree. Such an $\mathbb{M}$-valued Ξ-Fleming-Viot process is given in the case without dust by $(\chi_t, t \in \mathbb{R}_+)$, where χ_t is the isomorphy class of the metric measure space (Z, ρ, μ_t). The uniform convergence (1.1) also yields that tree-valued processes read off from finite restrictions of the lookdown model converge uniformly in compact time intervals (which is related to [46, Theorem 2], see Remark 4.11 below).

We show first a. s. convergence of probability measures in the Prohorov metric on the lookdown space, then we take isomorphy classes to obtain pathwise a tree-valued process. By contrast, in [46], first isomorphy classes are taken to describe evolving genealogies in finite populations. Then convergence in distribution in the Gromov-Prohorov metric is shown, and the limit process is characterized by a well-posed martingale problem.

From our approach on the lookdown space, it follows readily that tree-valued Ξ-Fleming-Viot processes have a. s. càdlàg paths with jumps at the times of large reproduction events (except for some settings in which there is no right-continuity at initial time). In particular, we retrieve the result from [46] that paths are a. s. continuous in the Gromov-weak topology in the case with only binary reproduction events (which is the case associated with the Kingman coalescent). The Gromov-weak topology emphasizes the typical genealogical distances in a sample from the population.

As initially suggested to the author by G. Kersting and A. Wakolbinger, we also consider a process whose state space is endowed with a stronger topology, the Gromov-Hausdorff-Prohorov topology, which highlights also the overall structure of the population. This process has jumps already in the Kingman case, namely at the times when the shape of the whole genealogical tree changes as all descendants of an ancestor die out (Theorem 3.4 and Proposition 4.2). We call this process a tree-valued evolving Ξ-coalescent, it can be defined in the case of coming down from infinity which is a subcase of the case without dust.

An ($\mathbb{R}_+$-)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 space of isomorphy classes of marked metric measure spaces is introduced in Depperschmidt, Greven, and Pfaffelhuber [24], we recall basic facts in the beginning of Section 4.2. Tree-valued Ξ-Fleming-Viot processes can be defined as processes with values in the space of isomorphy classes of marked metric measure spaces to include the case with dust, as shown in Chapter 2. To give a pathwise construction, we define in Section 2.1 for each individual (t, i) a parent $z(t, i)$, and we introduce an extended lookdown space $(\hat{Z}, \rho)$ which also includes the parents of the individuals at time zero. We denote by $v_t(i)$ the genealogical distance between the individual (t, i) and its parent, and we consider in Section 3.2 the probability measures

$$m_t^n = \sum_{i=1}^{n} \delta_{(z(t,i), v_t(i))}$$

on $\hat{Z} \times \mathbb{R}_+$. By Theorem 3.9, these measures m_t^n converge in the Prohorov metric uniformly for t in compact time intervals to probability measures $(m_t, t \in \mathbb{R}_+)$. In Section 4.2, we obtain the tree-valued Ξ-Fleming-Viot process as the process $(\hat{\chi}_t, t \in \mathbb{R}_+)$, where $\hat{\chi}_t$ is the isomorphy class of the marked metric measure space $(\hat{Z}, \rho, m_t)$. In the case without dust, $z(t, i) = (t, i)$ for all individuals (t, i), which yields consistency with the construction in the case without dust.

In general, we work in one-sided time. In this way, we obtain the path regularity of the processes under consideration for arbitrary initial states, which is applied in Chapter 5. Complementing the results on convergence to equilibrium in Section 10 of Chapter

2 and [46, Theorem 3], we also show in Section 4.1 that the tree-valued evolving Ξ-coalescent, started from any initial state, converges to a unique equilibrium, and we define a stationary tree-valued evolving Ξ-coalescent in two-sided time. We remark that in the Kingman case, the restriction of the lookdown space in two-sided time to the closure of the set of individuals at a fixed time t, endowed with the probability measure μ_t, equals the a. s. compact metric measure space associated with the Kingman coalescent that is studied by Evans [38].

We defer the proofs of the central theorems from Section 3 to the second part of the chapter whose organization is outlined in Section 5.

Now we discuss more relations to the literature. The lookdown graph of Pfaffelhuber and Wakolbinger [77] can be viewed as a semi-metric space whose completion is a lookdown space in two-sided time. A lookdown construction of the measure-valued Ξ-Fleming-Viot process is given by Birkner et al. [13]. Véber and Wakolbinger [90] give a lookdown construction of measure-valued spatial Λ-Fleming-Viot processes with dust using a skeleton structure. To construct the probability measures $(\mu_t, t \in \mathbb{R}_+)$ on the lookdown space in the case without dust, we use the flow of partitions for which we refer to Foucart [43] and Labbé [64]. A related description of evolving genealogies is the flow of bridges of Bertoin and Le Gall [8].

Pfaffelhuber, Wakolbinger, and Weisshaupt [78] and Dahmer, Knobloch, and Wakolbinger [23] study the compensated total tree length of the evolving Kingman coalescent as a stochastic process with jumps, using also the lookdown model. The times of these jumps correspond to the extinction times of parts of the population. Functionals of evolving coalescents such as the external length have been studied in several works, see for example [22, 58].

For the coming down from infinity property in the setting with simultaneous multiple reproduction events, see e. g. [13, 43, 53, 86]. By methods which differ from those in the present work, it is also shown in [26] that a. s., the states of the tree-valued Fleming-Viot process are non-atomic in the Kingman case. We also mention the work of Athreya, Löhr, and Winter [2] where in particular the Gromov-weak topology and the Gromov-Hausdorff-Prohorov topology are compared. Marked metric measure spaces are applied by Depperschmidt, Greven, and Pfaffelhuber [25, 26] to construct the tree-valued Fleming-Viot process with mutation and selection.

2 The lookdown space

We write $\mathbb{R}_+ = [0, \infty)$, $\mathbb{N} = \{1, 2, \ldots\}$, and we denote the set of partitions of $\mathbb{N}$ by $\mathcal{P}$. For $n \in \mathbb{N}$, we write $[n] = \{1, \ldots, n\}$ and we denote the set of partitions of $[n]$ by $\mathcal{P}_n$. Let $\gamma_n : \mathcal{P} \to \mathcal{P}_n$ be the restriction map, $\gamma_n(\pi) = \{B \cap [n] : B \in \pi\} \setminus \{\emptyset\}$. We endow $\mathcal{P}_n$ with the discrete topology, and $\mathcal{P}$ with the topology induced by the restriction maps.

Let us first repeat the lookdown model from Section 5 of Chapter 2 which is determined by the genealogy at time 0 and a point measure that encodes the reproduction events.

In the population model, there are countably infinitely many levels which are labeled by $\mathbb{N}$. The time axis is $\mathbb{R}_+$, and each level is occupied by one particle at each time. To encode the reproduction events that the particles undergo, we use a simple point measure η on $(0,\infty)\times\mathcal{P}$ with

$$\eta((0,T]\times\mathcal{P}^n)<\infty \quad \text{for all } n\in\mathbb{N} \text{ and } T\in(0,\infty), \tag{2.1}$$

where $\mathcal{P}^n$ denotes the subset of those partitions of $\mathbb{N}$ in which not all of the first n integers are in different blocks, that is,

$$\mathcal{P}^n=\{\pi\in\mathcal{P}:\gamma_n(\pi)\neq\{\{1\},\dots,\{n\}\}\}. \tag{2.2}$$

For a partition $\pi\in\mathcal{P}$ and $i\in\mathbb{N}$, we denote by $B_i(\pi)$ the i-th block of π when the blocks are ordered increasingly according to their smallest elements. Each point (t,π) of η is interpreted as a reproduction event as follows. At time $t-$, the particles on levels $i\in\mathbb{N}$ with $i>\#\pi$ are removed. Then, for each $i\in[\#\pi]$, the particle that was on level i at time $t-$ is on level $\min B_i(\pi)$ at time t and has offspring on all other levels in $B_i(\pi)$. In this way, the level of a particle is non-decreasing as time evolves. For each $n\in\mathbb{N}$, only finitely many particles in bounded time intervals are pushed away from one of the first n levels by condition (2.1).

We consider not only the process that describes the genealogical distances between the individuals at each fixed time as in Chapter 2, but we are interested in the genealogical distances between all individuals which we describe by a complete and separable metric space, the lookdown space. We define an individual as a particle at a fixed instant in time. We identify each element (t,i) of $\mathbb{R}_+\times\mathbb{N}$ with the individual on level i at time t. For $s\in[0,t]$, we denote by $A_s(t,i)$ the level of the ancestor of the individual (t,i) such that the maps $s\mapsto A_s(t,i)$ and $t\mapsto A_s(t,i)$ are càdlàg. Let ρ_0 be a semi-metric on $\mathbb{N}$. We define the genealogical distance between the individuals on levels i and j at time 0 by

$$\rho((0,i),(0,j))=\rho_0(i,j).$$

More generally, we define the genealogical distance between individuals $(t,i),(u,j)\in\mathbb{R}_+\times\mathbb{N}$ by

$$\rho((t,i),(u,j))=\begin{cases} t+u-2\sup\{s\le t\wedge u: A_s(t,i)=A_s(u,j)\} & \text{if } A_0(t,i)=A_0(u,j)\\ t+u+\rho_0(A_0(t,i),A_0(u,j)) & \text{else.}\end{cases}$$

The genealogical distance $\rho((t,i),(u,j))$ can be seen as the sum of the distances to the most recent common ancestor of (t,i) and (t,j) if these individuals have a common ancestor after time zero. Else it is the genealogical distance of their ancestors at time zero, augmented by the times at which the individuals live. The distinction between these two cases is needed as we work in one-sided time.

The distance ρ is a semi-metric on $\mathbb{R}_+\times\mathbb{N}$ (offspring individuals from the same parent have genealogical distance zero at the time of the reproduction event). We identify individuals with genealogical distance zero, and we take the metric completion. We call the resulting metric space (Z,ρ) the lookdown space associated with η and ρ_0. In slight abuse of notation, we refer by $(t,i)\in\mathbb{R}_+\times\mathbb{N}$ also to the element of the metric space after the identification of elements with ρ-distance zero, in this sense we also assume $\mathbb{R}_+ \times \mathbb{N} \subset Z$.

For $t \in \mathbb{R}_+$, we define a semi-metric ρ_t on $\mathbb{N}$ by

$$\rho_t(i,j) = \rho((t,i),(t,j)), \quad i,j \in \mathbb{N}. \tag{2.3}$$

Then ρ_t describes the genealogical distances between the particles at fixed times, and the process $(\rho_t, t \in \mathbb{R}_+)$ is the process that is denoted in the same way in Section 5 of Chapter 2.

The remainder of this section is organized as follows. In Subsection 2.1, we replace ρ_0 with a decomposed semi-metric and we enlarge the lookdown space by parents of the individuals at time zero. In Subsection 2.2, we consider the two ways in which particles can die and we define extinction times for parts of the population. The construction is randomized in Subsection 2.3 where η becomes a Poisson random measure.

2.1 Parents and decomposed genealogical distances

We will use the contents of this section to include the case with dust.

Let r_0 be a semi-metric on $\mathbb{N}$ and $v_0 = (v_0(i))_{i\in\mathbb{N}} \in \mathbb{R}_+^{\mathbb{N}}$ such that (r_0, v_0) satisfies

$$\rho_0(i,j) = (r_0(i,j) + v_0(i) + v_0(j))\,\mathbf{1}\{i \neq j\}$$

for all $i,j \in \mathbb{N}$. Then (r_0, v_0) is a decomposition of the semi-metric ρ_0 in the sense of Section 2 in Chapter 2. The trivial decomposition $(r_0, v_0) = (\rho_0, 0)$ always exists.

For each $(t,i) \in \mathbb{R}_+ \times \mathbb{N}$, we define the quantity $v_t(i)$ as in Section 7.1 of Chapter 2: For $j \in \mathbb{N}$, let $\mathcal{P}(j) = \{\pi \in \mathcal{P} : \{j\} \notin \pi\}$ be the set of partitions of $\mathbb{N}$ in which j 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 from the individual (t,i) until the ancestral lineage 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.

Now we enlarge the set of individuals to the disjoint union $(\mathbb{R}_+ \times \mathbb{N}) \sqcup \mathbb{N}$. We call each element i of $\mathbb{N} \subset (\mathbb{R}_+ \times \mathbb{N}) \sqcup \mathbb{N}$ the parent of the individual $(0,i)$. We extend the semi-metric ρ to $(\mathbb{R}_+ \times \mathbb{N}) \sqcup \mathbb{N}$ by

$$\begin{aligned} \rho(i,j) &= r_0(i,j) \quad \text{for } i,j \in \mathbb{N} \\ \text{and} \quad \rho((t,i),j) &= t + v_0(A_0(t,i)) + r_0(A_0(t,i),j) \quad \text{for } (t,i) \in \mathbb{R}_+ \times \mathbb{N}, j \in \mathbb{N}. \end{aligned}$$

Furthermore, we define for each individual $(t,i) \in \mathbb{R}_+ \times \mathbb{N}$ the parent $z(t,i)$ as the individual $(t - v_t(i), A_{t-v_t(i)}(t,i))$ if $v_t(i) < t$, else we set $z(t,i) = A_0(t,i)$. Then $v_t(i)$ equals the genealogical distance between the individual (t,i) and its parent.

We identify the elements of $(\mathbb{R}_+ \times \mathbb{N}) \sqcup \mathbb{N}$ with distance zero and take the metric completion. We call the resulting metric space the extended lookdown space associated with η and (r_0, v_0), and we denote it by $(\hat{Z}, \rho)$. Here we write again $\hat{Z} \supset (\mathbb{R}_+ \times \mathbb{N}) \sqcup \mathbb{N}$ in slight abuse of notation. Note that the lookdown space (Z, ρ) associated with η and ρ_0 is contained in $(\hat{Z}, \rho)$ as a subspace. Figure 2.1 below shows an extended lookdown space.

Remark 2.1. If $v_0 = 0$, then $\rho_0 = r_0$ and the individuals at time zero in the extended lookdown space $(\hat{Z}, \rho)$ are identified with their parents as $\rho((0,i), z(0,i)) = 0$ for all $i \in \mathbb{N}$. In this case, $(\hat{Z}, \rho)$ is equal to the lookdown space (Z, ρ) associated with η and r_0 from the beginning of Section 2.

Remark 2.2 (Relation to the decomposed genealogical distances in Chapter 2)*.* We denote the genealogical distances between the parents of individuals $(t,i), (t,j) \in \mathbb{R}_+ \times \mathbb{N}$ by

$$r_t(i,j) = \rho(z(t,i), z(t,j)).$$

For $t = 0$, this is consistent with the definition of r_0 above as $z(0,i) = i$, $z(0,j) = j$, and $\rho(i,j) = r_0(i,j)$. For all $t \in \mathbb{R}_+$ and $i, j \in \mathbb{N}$,

$$\rho_t(i,j) = (v_t(i) + r_t(i,j) + v_t(j))\, \mathbf{1}\{i \neq j\}\,. \tag{2.4}$$

That is, the process $((r_t, v_t), t \in \mathbb{R}_+)$ of the decomposed genealogical distances between the individuals at fixed times coincides with the process defined from η and (r_0, v_0) in Section 7.1 of Chapter 2. For $t = 0$, equation (2.4) holds by definition of ρ_0, r_0, and v_0. That equation (2.4) holds for all $t \in \mathbb{R}_+$ can be seen from Figure 2.1. For a formal proof, we distinguish four cases. We always assume $i \neq j$ in the following.

Case 1: $v_t(i), v_t(j) < t$, $A_0(t,i) = A_0(t,j)$. In this case, the definition of $v_t(i)$ and $v_t(j)$ implies $A_s(t,i) \neq A_s(t,j)$ for all $s \in (t - v_t(i) \vee v_t(j), t]$. By definition of ρ, it follows

$$\begin{aligned}&\rho(z(t,i), z(t,j)) = t - v_t(i) - 2\sup\{s \leq t : A_0(t,i) = A_0(t,j)\} + t - v_j(t)\\ &= \rho((t,i),(t,j)) - v_t(i) - v_t(j),\end{aligned}$$

which is equation (2.4).

Case 2: $v_t(i), v_t(j) < t$, $A_0(t,i) \neq A_0(t,j)$. In this case, the definition of $z(t,i)$ and $z(t,j)$ yields $A_0(t,i) = A_0(z(t,i))$ and $A_0(t,j) = A_0(z(t,j))$. In particular, it follows $A_0(z(t,i)) \neq A_0(z(t,j))$. With the definition of ρ, it follows

$$\begin{aligned}&\rho(z(t,i), z(t,j)) = t - v_t(i) - \rho(A_0(t,i), A_0(t,j)) + t - v_j(t)\\ &= \rho((t,i),(t,j)) - v_t(i) - v_t(j),\end{aligned}$$

which is equation (2.4).

Case 3: $v_t(i) < t, v_t(j) \geq t$. In this case, it follows that $A_0(t,i) \neq A_0(t,j)$. From the definitions, it follows that

$$\rho(z(t,i), z(t,j)) = t - v_t(i) + v_0(A_0(t,i)) + r_0(A_0(t,i), A_0(t,j)).$$

Using $v_t(j) = t + v_0(A_0(t,j))$ and equation (2.4) for $t = 0$, we deduce that

$$v_t(i) + r_t(i,j) + v_t(j) = t + \rho_0(A_0(t,i), A_0(t,j)) + t$$

which is equation (2.4).

Case 4: $v_t(i), v_t(j) \geq t$. Again by the definitions and by equation (2.4) for $t = 0$, we have

$$\begin{aligned} &v_t(i) + \rho(z(t,i), z(t,j)) + v_t(j) \\ &= t + v_0(A_0(t,i)) + r_0(A_0(t,i), A_0(t,j)) + t + v_0(A_0(t,j)) = 2t + \rho_0(A_0(i,j), A_0(t,j)), \end{aligned}$$

which is equation (2.4).

Remark 2.3 (Parents and starting vertices of external branches). In this remark, we assume $v_0(i) = \frac{1}{2}\inf_{j\in\mathbb{N}\setminus\{i\}} \rho_0(i,j)$ for all $i \in \mathbb{N}$, and that ρ_0 is a semi-ultrametric (that is, $\rho_0(i,j) \vee \rho_0(j,k) \geq \rho_0(i,k)$ for all $i,j,k \in \mathbb{N}$). Then, as in Remark 1.1 of Chapter 2 and the references therein, we associate with ρ_0 the real tree (T_0, d_0) that is obtained by identifying the points with distance zero in the semi-metric space $((-\infty, 0] \times \mathbb{N}, d_0)$, where $d_0((s,i),(t,j)) = \max\{\rho_0(i,j) + s + t, |s-t|\}$. Now we briefly sketch how the space $(\hat{Z}, \rho)$ can be isometrically embedded into a real tree (T, d) that contains the genealogical trees of the individuals at all times and we interpret the parents as starting vertices of external branches.

We define a semi-metric d on $\mathbb{R} \times \mathbb{N}$ that coincides on $\mathbb{R}_+ \times \mathbb{N}$ with the semi-metric ρ from the beginning of Section 2, that coincides with d_0 on $(-\infty, 0] \times \mathbb{N}$, and for $(s,i) \in \mathbb{R}_+ \times \mathbb{N}$, $(t,j) \in (-\infty, 0] \times \mathbb{N}$, we set $d((s,i),(t,j)) = s + d_0(A_0(s,i),(t,j))$. Then we identify points with d-distance zero and define (T, d) as the metric completion. By construction, $\rho_0(i,j) = d((0,i),(0,j))$ and $r_0(i,j) = d((-v_0(i), i), (-v_0(j), j))$ for all $i,j \in \mathbb{N}$. Hence, $(\hat{Z}, \rho)$ is embedded into (T, d) by the isometry that maps $(t,i) \in \mathbb{R}_+ \times \mathbb{N} \subset \hat{Z}$ to (t,i), and $i \in \mathbb{N} \subset \hat{Z}$ to $(-v_0(i), i)$.

For each $t \in \mathbb{R}_+$, the subspace $T_t = (-\infty, t] \times \mathbb{N}$ of (T, d) is the real tree associated with the semi-ultrametric ρ_t. Assume furthermore that $v_t(i) = \frac{1}{2}\inf_{j\in\mathbb{N}\setminus\{i\}} \rho_t(i,j)$ for some $t \in \mathbb{R}_+$. Then as in Remark 2.2 of Chapter 2, the quantity $v_t(i)$ and the parent $z(t,i)$ can be interpreted as the length and the starting vertex, respectively, of the external branch that ends in the leaf (t,i) of T_t.

Remark 2.4. In the context of Sections 3 and 4, the assumption on v_0 in Remark 2.3 can be checked by using Proposition 3.3 in Chapter 2, and the assumption on v_t by Proposition 7.4 in Chapter 2.

2.2 Extinction of parts of the population

Recall the particle picture from the beginning of this section. In reproduction events encoded by points (t, π) of η where the partition π has finitely many blocks, the particles at time t have only finitely many ancestors among the infinite population at time $t-$, hence particles die at time t. A particle can also die due an accumulation of reproduction events in which its level is pushed to infinity.

For $(s,i) \in \mathbb{R}_+ \in \mathbb{N}$ and $t \in [s, \infty)$, let $D_t(s,i)$ be the lowest level that is occupied at time t by a descendant of the individual (s,i), that is,

$$D_t(s,i) = \inf\{j \in \mathbb{N} : A_s(t,j) = i\}$$

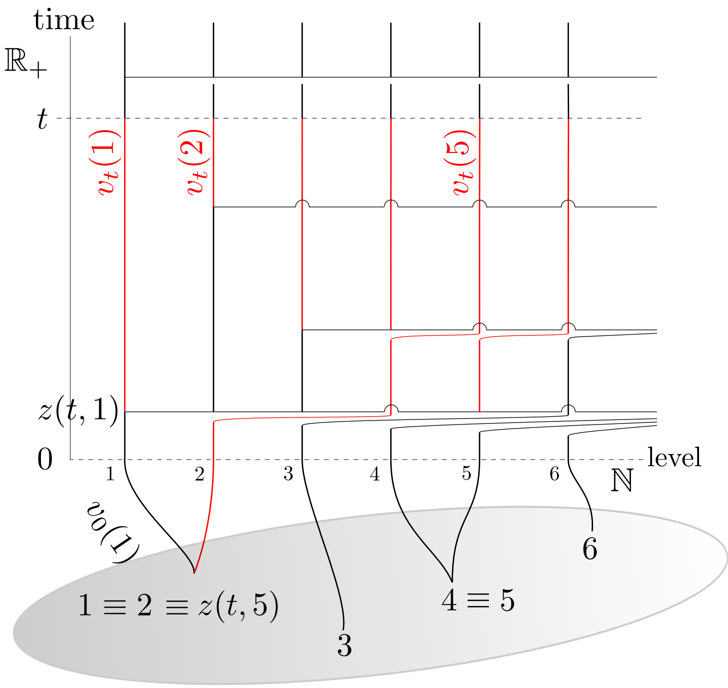

Figure 2.1: Part of an extended lookdown space. The space $\mathbb{R}_+ \times \mathbb{N}$ is represented in the upper part of the figure. Time goes upwards and levels go from the left to the right. In the lower part, the metric space obtained from $\mathbb{N}$, endowed with the semi-metric obtained from r_0 is symbolized. For each $i \in \mathbb{N}$, the junction between the individual $(0, i)$ and its ancestor i has length $v_0(i)$. Individuals that are in the same block in a reproduction event have genealogical distance zero and are identified. In the figure, they are connected by horizontal lines. In this example, there are no simultaneous multiple reproduction events. The genealogical distances between the individuals (t, i) and their respective parents $z(t, i)$ equal $v_t(i)$, they are represented by red lines. The genealogical distance between any two individuals is the sum of the lengths of the vertical parts of the path from one individual to the other, plus the distance in the metric space obtained from r_0 if this space has to be traversed.

with $D_t(s, i) = \infty$ if and only if there exists no $j \in \mathbb{N}$ with $i = A_s(t, j)$. This quantity corresponds to the forward level process in [77] and to the fixation line in [52]. The map $t \mapsto D_t(s, i)$ is non-decreasing. Let $\tau_{s,i}$ be the extinction time of the part of the population that descends from the individual (s, i), that is,

$$\tau_{s,i} = \inf\{t \in [s, \infty) : D_t(s, i) = \infty\}.$$

Then the set of times at which such parts of the population become extinct is given by

$$\Theta^{\text{ext}} := \{\tau_{s,i} : s \in \mathbb{R}_+, i \in \mathbb{N}\}. \tag{2.5}$$

Remark 2.5. In a reproduction event that is encoded by a point (t, π) of η with $\#\pi = \infty$, every individual that sits on a level $i \in \mathbb{N}$ at time $t-$ has a descendant at time t. Hence the two mechanisms mentioned in the beginning of this subsection are the only possibilities for a particle to die.

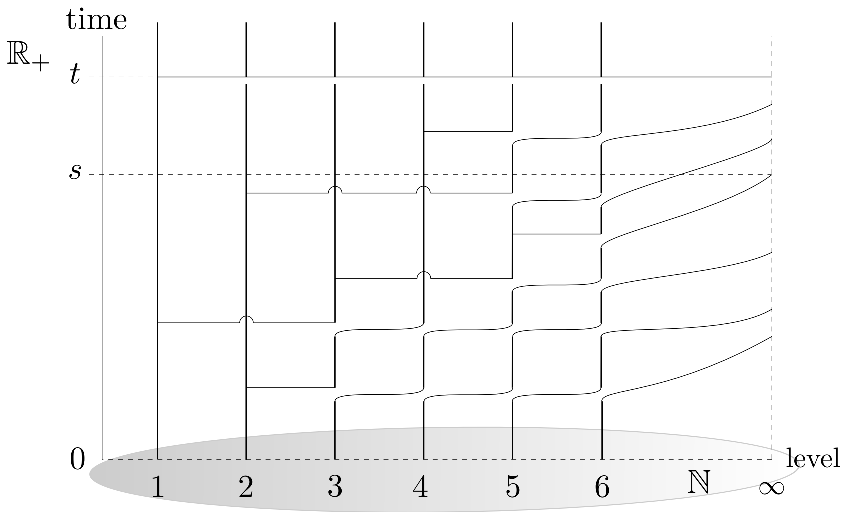

Figure 2.2: Part of a lookdown space. Only the reproduction events with offspring on the first 6 levels are drawn. At time t, a reproduction event occurs in which the whole population is replaced by the offspring of the individual on level 1. The limits $(t-,i) = \lim_{r\uparrow t}(r,i)$, $i \geq 2$ are part of the boundary of the lookdown space. Due to accumulations of jumps, the lines $t' \mapsto D_{t'}(s',i)$ may hit infinity, here symbolized by a dashed line, similarly to illustrations of the lookdown graph in [77]. This occurs for instance at time s, the limit $\lim_{r\uparrow s} D_r(0,3)$ is part of the boundary. Further elements of the boundary are obtained from Cauchy sequences at fixed times.

2.3 The Ξ-lookdown model

We recall here the simple point measure η that is used to drive the lookdown model in Chapter 2, cf. also the references therein.

Let Δ be the simplex

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

where $|x|_p = (\sum_{i\in\mathbb{N}} x_i^p)^{1/p}$. Let κ be the probability kernel from Δ to $\mathcal{P}$ associated with Kingman's correspondence, that is, $\kappa(x,\cdot)$ is the distribution of the paintbox partition associated with $x \in \Delta$, see e. g. Section 2.3.2 in [7].

Let Ξ be a finite measure on Δ. We decompose

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

For distinct integers $i,j \in \mathbb{N}$, we denote by $K_{i,j}$ the partition of $\mathbb{N}$ 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

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

In the following sections, η is always a Poisson random measure on $(0,\infty)$ with intensity $dt\ H_\Xi(d\pi)$. Then η satisfies a. s. condition (2.1), as in Sections 6 and 7.2 of Chapter 2 and the references therein. The lookdown model can therefore be driven by η.

The measure Ξ is called dust-free if and only if

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

In this case, each particle reproduces with infinite rate, hence $v_t(i) = 0$ for all $t \in (0, \infty)$ and $i \in \mathbb{N}$ a. s. We denote by $\mathcal{M}_{\mathrm{nd}}$ the subset of finite measures on Δ that are dust-free, and by $\mathcal{M}_{\mathrm{dust}}$ its complement in the set of finite measures on Δ.

The extended lookdown space in Figure 2.1 could be the extended lookdown space associated with a typical realization of an appropriate Poisson random measure η with dust, and a decomposed semimetric (r_0, v_0). Figure 2.2 illustrates the dust-free case.

We speak of a large reproduction event when a particle has offspring on a positive proportion of the levels. We denote by Θ_0 the set of times at which large reproduction events occur:

$$\Theta_0 = \{t \in (0, \infty) : \text{there exist } \pi \in \mathcal{P} \text{ and } B \in \pi \text{ with } \eta\{(t, \pi)\} > 0 \text{ and } |B| > 0\}. \quad (2.6)$$

Here $|B| = n^{-1} \lim_{n \to \infty} B \cap [n]$ denotes the asymptotic frequency of the block B. The measure Ξ_0 governs the large reproduction events, as opposed to $\Xi\{0\}$, which gives the rates of the binary reproduction events. A. s. by definition of H_Ξ and Kingman's correspondence, each particle that reproduces in a large reproduction event has offspring on a positive proportion of the levels. In case $\Xi \in \mathcal{M}_{\mathrm{nd}}$, the set Θ_0 equals a. s. the set $\{t \in (0, \infty) : \eta(\{t\} \times \mathcal{P}) > 0\}$ of reproduction times, as there are a. s. no binary reproduction events.

3 Sampling measures and jump times

In this section, we consider mathematical objects that are defined from a realization of the Poisson random measure η and a random (decomposed) distance matrix on an event of probability 1. Stochastic processes are read off from these constructions in Section 4. We defer the proofs of the main statements in Section 3 to the second part of the chapter which begins in Section 5.

3.1 The case without dust

We construct a family of probability measures on the lookdown space. We consider regularity of this family in the weak topology. In the case of coming down from infinity, we also consider regularity of a family of subsets of the lookdown space with respect to the Hausdorff distance.

Let the Poisson random measure η be defined from the finite measure Ξ on Δ as in Section 2.3. Let (X, r, μ) be a metric measure space and ρ_0 be an independent $\mathbb{R}^{\mathbb{N}^2}$-valued random variable that has the distance matrix distribution of (X, r, μ). That is, we can assume $\rho_0 = (r(x(i), x(j)))_{i,j \in \mathbb{N}}$ for a μ-iid sequence $x(1), x(2), \ldots$ in (X, r). Then we can view ρ_0 as a random semi-metric on $\mathbb{N}$. Let (Z, ρ) be the lookdown space associated with η and ρ_0 as defined in Section 2.

For each $t \in \mathbb{R}_+$ and $n \in \mathbb{N}$, let the probability measure μ_t^n on (Z, ρ) be the uniform measure on the first n individuals at time t, that is,

$$\mu_t^n = \frac{1}{n} \sum_{i=1}^{n} \delta_{(t,i)}.$$

Let μ_t be the weak limit of μ_t^n, provided it exists. For almost all realizations of η and ρ_0, these weak limits exist simultaneously for all $t \in \mathbb{R}_+$ by Theorem 3.1 below. The convergence is uniform for the Prohorov metric d_{P}^Z on (Z, ρ) for t in compact intervals.

Theorem 3.1. *Assume $\Xi \in \mathcal{M}_{\mathrm{nd}}$. Then there exists an event of probability 1 on which the following assertions hold:*

(i) For all $T \in \mathbb{R}_+$,

$$\lim_{n\to\infty} \sup_{t\in[0,T]} d_{\mathrm{P}}^Z(\mu_t^n, \mu_t) = 0.$$

(ii) The map $t \mapsto \mu_t$ is càdlàg in the weak topology on the space of probability measures on (Z, ρ). The set Θ_0, defined in (2.6), is the set of jump times.

(iii) For all $t \in \Theta_0$, the measure μ_t contains atoms, and the left limit μ_{t-} is non-atomic.

Note that the measure μ_t^n assigns mass $1/n$ to each offspring in a reproduction event. At the times of large reproduction events, the measure μ_t has an atom on each family of individuals at time t that descend from the same individual at time $t-$. Indeed, these individuals have genealogical distance zero and are identified in the lookdown space. In the Kingman case, that is, if $\Xi = \delta_0$, there are a. s. no large reproduction events and $t \mapsto \mu_t$ is a. s. continuous.

For $t \in \mathbb{R}_+$, let X_t be the closure of the set of individuals $\{t\} \times \mathbb{N}$ at time t. Clearly, the probability measures μ_0^n do not depend on η. Their weak limit exists a. s. by the following lemma which is essentially Vershik's proof [91, Theorem 4] of the Gromov reconstruction theorem, see also Proposition 3.12 of Chapter 2. We write supp μ' for the support of a measure μ' (which is closed).

Lemma 3.2. *The weak limit μ_0 of the probability measures μ_0^n on Z exists a. s. The metric measure spaces* (supp μ, r, μ) *and* (X_0, ρ, μ_0) *are a. s. measure-preserving isometric.*

Proof. By the Glivenko-Cantelli theorem, the empirical measures $\mu^n := n^{-1} \sum_{i=1}^n \delta_{x(i)}$ converge to μ a. s. The isometry $\{x(i) : i \in \mathbb{N}\} \to Z$, $x(i) \mapsto (0, i)$ can be extended to an isometry from (supp μ, r) to (X_0, ρ). This yields the assertion as in Proposition 3.12 of Chapter 2. □

In Theorem 3.4 below, we consider measures Ξ that satisfy the "coming down from infinity"-assumption, that is, we assume that there exists an event of probability 1 on which the number $\#\{A_s(t, j) : j \in \mathbb{N}\}$ of ancestors at time s of the individuals at time t is finite for all $0 \leq s < t$. Let $\mathcal{M}_{\mathrm{CDI}}$ denote the subset of those finite measures on Δ that satisfy this assumption. Clearly, $\mathcal{M}_{\mathrm{CDI}} \subset \mathcal{M}_{\mathrm{nd}}$. Indeed, if $\Xi \in \mathcal{M}_{\mathrm{dust}}$, then the rate at which a given ancestral lineage merges with any other ancestral lineage is finite which implies $\{A_s(t,j) : j \in \mathbb{N}\} = \infty$ a. s. for all $0 \leq s < t$, see for instance [86]. Furthermore, if $\Xi \in \mathcal{M}_{\mathrm{CDI}}$, then the set Θ^{ext} of extinction times, defined in (2.5), is a. s. dense in $(0, \infty)$. Indeed, if Θ^{ext} has no points in an interval $[s, t]$, then the individuals on all levels at time s are ancestors of individuals at time t.

Remark 3.3. Under the assumptions of Theorem 3.4 below, there exists an event of probability one on which all subsets $X_t \subset Z$ with $t \in \mathbb{R}_+$ are compact. A. s. compactness of X_0 follows from Lemma 3.2 which also implies assertion (i) of Theorem 3.4 below for $t = 0$. For each $t \in (0, \infty)$ the closed subset X_t is totally bounded (hence compact) if $\#\{A_s(t,j) : j \in \mathbb{N}\} < \infty$ for all $s \in (0, t)$. This follows by definition of the metric ρ on the lookdown space Z.

Recall that the Hausdorff distance between two subsets A, B of a metric space (Y, d) is defined as the infimum over those $\varepsilon > 0$ such that $d(a, B) < \varepsilon$ for all $a \in A$ and $d(A, b) < \varepsilon$ for all $b \in B$. The Hausdorff distance is a metric on the set of closed subspaces of (Y, d), see e. g. [17].

Theorem 3.4. *Assume* $\Xi \in \mathcal{M}_{\mathrm{CDI}}$ *and that* (X, r) *is compact. Then the following assertions hold on an event of probability* 1*:*

(i) For each $t \in \mathbb{R}_+$*, the compact set* X_t *is the support of* μ_t*.*

(ii) The map $t \mapsto X_t$ *is càdlàg for the Hausdorff distance on the set of closed subsets of* (Z, ρ)*. The set* Θ^{ext} *is the set of jump times. For each* $t \in \Theta^{\mathrm{ext}}$*, the set* X_t *and the left limit* X_{t-} *are not isometric.*

Remark 3.5. By Theorems 3.1 and 3.4, there exists an event of probability 1 on which $\mu_t = \mu_{t-}$ and $X_t \subsetneq X_{t-}$ for all $t \in \Theta^{\mathrm{ext}} \setminus \Theta_0$. In particular, the support of μ_{t-} is strictly smaller than X_{t-} for these t a. s. The set $X_{t-} \setminus \operatorname{supp} \mu_{t-} \subset X_{t-}$ is equal to $X_{t-} \setminus X_t$ a. s., this is the part of the population at time $t-$ that dies out at time t.

We conclude this subsection with a side observation (Proposition 3.7) on the intersection of the sets of jump times in Theorems 3.1 and 3.4. Reproduction events in which the whole population is replaced by finitely many particles and their offspring occur at the times in the set

$$\Theta_{\mathrm{f}} := \{t \in (0, \infty) : \text{ there exists } \pi \in \mathcal{P} \text{ with } \eta\{(t, \pi)\} > 0 \text{ and } \#\pi < \infty\}.$$

By the construction in Section 2, all particles with level larger than $\#\pi$ die in a reproduction event that is encoded by a point (t, π) of η.

Remark 3.6. If Ξ is concentrated on $\{(x, 0, 0, \ldots) : x \in [0, 1]\} \subset \Delta$, then a. s., no simultaneous multiple reproduction events occur. This case corresponds to the coalescents with multiple collisions (Λ-coalescents). In this case, Θ_{f} is a. s. the set of times at which the whole population is replaced by a single particle and its offspring. If Ξ is concentrated on $\{(x, 0, 0, \ldots) : x \in [0, 1)\} \subset \Delta$, then $\Theta_{\mathrm{f}} = \emptyset$ a. s. More generally, $\Theta_{\mathrm{f}} = \emptyset$ a. s. if and only if $\Xi\{x \in \Delta : x_1 + \ldots + x_k = 1 \text{ for some } k \in \mathbb{N}\} = 0$.

Proposition 3.7. *A. s.,* $\Theta^{\mathrm{ext}} \cap \Theta_0 = \Theta_{\mathrm{f}}$.

For the proof and for later use in Sections 8 and 10, we now express η in terms of a collection of Poisson processes. Let $\pi'_1, \pi'_2, \ldots$ be an arbitrary enumeration of $\mathcal{P}$ and let $J_{t,k} = \eta((0,t] \times \{\pi'_k\})$ for $t \in \mathbb{R}_+$ and $k \in \mathbb{N}$. Then the processes $(J_{t,k}, t \in \mathbb{R}_+)$ form a collection (indexed by $k \in \mathbb{N}$) of independent Poisson processes. We endow $\mathbb{R}_+^{\mathbb{N}}$ with the product topology and consider the $\mathbb{R}_+^{\mathbb{N}}$-valued stochastic process

$$J = (J_t, t \in \mathbb{R}_+) = ((J_{t,k}, k \in \mathbb{N}), t \in \mathbb{R}_+) \tag{3.1}$$

Note that J is a strong Markov process and by e. g. Theorem (5.1) in Chapter I of Blumenthal [16] quasi-left-continuous. Let $\mathcal{F} = (\mathcal{F}_t, t \in \mathbb{R}_+)$ be the complete filtration induced by J. Then $\mathcal{F}_t$ is the sigma field generated by the random measure $\eta(\cdot \times ((0,t] \times \mathcal{P}))$ and all null events. By condition (2.1), J stays finite a. s. and the set of jump times of J equals a. s. the set of reproduction times $\{t \in (0,\infty) : \eta(\{t\} \times \mathcal{P}) > 0\}$.

Proof. We use notation also from Section 2.2. First we show $\Theta_{\mathrm{f}} \subset \Theta^{\mathrm{ext}} \cap \Theta_0$ a. s. Let $t \in \Theta_{\mathrm{f}}$. Then there exists a point (t, π) of η with $i - 1 := \#\pi < \infty$. By condition (2.1), as the level of each particle is non-decreasing in time, and as offspring always has a higher level than the reproducing particle, there exists a. s. a time $s \in (0,t)$ such that $D_{s'}(s,i) = i$ for all $s' \in [s,t)$. In particular, all descendants of (s,i) at time $t-$ have level at least i. Hence, (s,i) has no descendants at time t, that is, $D_t(s,i) = \infty$, and it follows $t \in \Theta^{\mathrm{ext}}$. Clearly, $\#\pi < \infty$ implies that π contains blocks of infinite size. The definition of H_Ξ and Kingman's correspondence imply $\Theta_{\mathrm{f}} \subset \Theta_0$ a. s. It remains to show that $\Theta^{\mathrm{ext}} \cap \Theta_0 \subset \Theta_{\mathrm{f}}$ a. s.

On the event of probability 1 on which condition (2.1) holds, particles on any level remain on that level for a positive amount of time. This implies

$$\Theta^{\mathrm{ext}} = \{\tau_{s,i} : s \in \mathbb{Q}_+, i \in \mathbb{N}\} \quad \text{a. s.}$$

For $s \in \mathbb{R}_+$, $i, n \in \mathbb{N}$, we define the $\mathcal{F}$-stopping time

$$\tau_{s,i,n} = \inf\{t \geq s : D_t(s,i) \geq n\}.$$

Then $\tau_{s,i,n}$ is non-decreasing in n, and $\tau_{s,i} = \lim_{n\to\infty} \tau_{s,i,n}$. We assume w. l. o. g. $\Xi(\Delta) > 0$. Then $\tau_{s,i,n} \in [s,\infty)$ for all $s \in \mathbb{R}_+$ and $i, n \in \mathbb{N}$ a. s. Let $E_{s,i}$ be the event that $\tau_{s,i,n} < \tau_{s,i}$ for all $n \in \mathbb{N}$.

A. s. by (2.1), on the event $E_{s,i}^c$ that $\tau_{s,i} = \tau_{s,i,n}$ for some $n \in \mathbb{N}$, a particle on a level below n at time $\tau_{s,i,n}-$ dies at time $\tau_{s,i,n}$ due to a reproduction event that lies in Θ_{f}.

To show the assertion of the proposition, it now suffices to show that $\tau_{s,i} \notin \Theta_0$ a. s. on $E_{s,i}$. We define the $\mathcal{F}$-stopping time $\tilde{\tau}_{s,i}$ by $\tilde{\tau}_{s,i} = \tau_{s,i}\mathbf{1}_{E_{s,i}} + \infty\mathbf{1}_{E_{s,i}^c}$. Then the $\mathcal{F}$-stopping times

$$\tilde{\tau}_{s,i,n} := \begin{cases} \tau_{s,i,n} & \text{if } \tau_{s,i,n} < \tau_{s,i} \\ \tau_{s,i,n} \vee n & \text{if } \tau_{s,i,n} = \tau_{s,i} \end{cases}$$

form an announcing sequence for $\tilde{\tau}_{s,i}$, that is, $\tilde{\tau}_{s,i,n} < \tilde{\tau}_{s,i}$ a. s. and $\tilde{\tau}_{s,i} = \lim_{n\to\infty} \tilde{\tau}_{s,i,n}$ a. s. Quasi-left-continuity of J implies $J_{\tilde{\tau}_{s,i}-} = J_{\tilde{\tau}_{s,i}}$ a. s. Hence, a. s. on $E_{s,i}$, no reproduction event occurs at time $\tilde{\tau}_{s,i} = \tau_{s,i}$, and we have $\tau_{s,i} \notin \Theta_0$. □

Remark 3.8. In the Kingman case, the set Θ^{ext} of extinction times is described by Poisson processes by Dahmer, Knobloch, and Wakolbinger [23, Proposition 1], see also the references therein.

3.2 The general case

We construct a family of probability measures on the Cartesian product of the extended lookdown space and the mark space $\mathbb{R}_+$. Let (r_0, v_0) be an independent $\mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}}$-valued random variable that has the marked distance matrix distribution of an $(\mathbb{R}_+)$-marked metric measure space (X, r, m). That is, (X, r) is a complete and separable metric space, m is a probability measure on the Borel sigma algebra on the product space $X \times \mathbb{R}_+$, and we may assume that $(x(i), v(i))_{i \in \mathbb{N}}$ is an m-iid sequence in $X \times \mathbb{R}_+$ and set $r_0(i,j) = r(x(i), x(j))$ for $i, j \in \mathbb{N}$. Then we can view r_0 as a random semi-metric on $\mathbb{N}$. Let $(\hat{Z}, \rho)$ be the extended lookdown space associated with η and (r_0, v_0), as defined in Section 2.1. We endow $\hat{Z} \times \mathbb{R}_+$ with the product metric $d^{\hat{Z} \times \mathbb{R}_+}((z, v), (z', v')) = \rho(z, z') \vee |v - v'|$. Recall from Section 2.1 also the parent $z(t, i)$ and the genealogical distance $v_t(i)$ between the individual (t, i) and its parent. For each $t \in \mathbb{R}_+$ and $n \in \mathbb{N}$, we define a probability measure m_t^n on $\hat{Z} \times \mathbb{R}_+$ by

$$m_t^n = \frac{1}{n} \sum_{i=1}^{n} \delta_{(z(t,i), v_t(i))}$$

The first component $m_t^n(\cdot \times \mathbb{R}_+)$ lays mass on the parents of the first n individuals at time t. The second component $m_t^n(\hat{Z} \times \cdot)$ records the genealogical distances to these parents. Let m_t denote the weak limit of m_t^n provided it exists. This existence is addressed in Theorem 3.9 below in the case with dust, and in Lemma 3.10, Remark 3.12, and Corollary 3.13 below in the case without dust.

Theorem 3.9. *Assume $\Xi \in \mathcal{M}_{\text{dust}}$. Then the following assertions hold on an event of probability* 1*:*

(i) For all $T \in \mathbb{R}_+$,

$$\lim_{n \to \infty} \sup_{t \in [0,T]} d_{\mathrm{P}}^{\hat{Z} \times \mathbb{R}_+}(m_t^n, m_t) = 0.$$

(ii) The map $t \mapsto m_t$ is càdlàg in the weak topology on the space of probability measures on $\hat{Z} \times \mathbb{R}_+$. The set Θ_0, defined in (2.6), is the set of jump times.

(iii) For each $t \in (0, \infty)$, the left limit m_{t-} satisfies $m_{t-}(\hat{Z} \times \{0\}) = 0$. For each $t \in \Theta_0$, it holds $m_t(\hat{Z} \times \{0\}) > 0$.

(iv) If m is purely atomic, then m_t and m_{t-} are purely atomic for all $t \in (0, \infty)$.

In Proposition 9.2, the measures m_t are stated explicitly. At the times of large reproduction events, $v_t(i) = 0$ for all individuals i with levels in a non-singleton block. This yields the positive mass of $m_t(\hat{Z} \times \cdot)$ in zero asserted in Theorem 3.9(iii).

Lemma 3.10. *The weak limit m_0 of the probability measures m_0^n on $\hat{Z} \times \mathbb{R}_+$ exists a. s.*

Proof. This follows analogously to Lemma 3.2. □

Remark 3.11. We recall that case $\Xi \in \mathcal{M}_{\mathrm{nd}}$, the rate at which each particle reproduces is infinite. In this case, there exists an event of probability 1 on which $v_t(i) = 0$ and $(t,i) = z(t,i)$ for all $t \in (0,\infty)$ and $i \in \mathbb{N}$, cf. Section 7.2 of Chapter 2.

As in Subsection 3.1, we define the probability measures $\mu_t^n = \frac{1}{n}\sum_{i=1}^n \delta_{(t,i)}$ on $\hat{Z}$, and we denote their weak limits by μ_t.

Remark 3.12. If $m(\hat{Z} \times \{0\}) = 1$, then $m = \mu \otimes \delta_0$ for a probability measure μ on X, hence $v_0 = 0$ a.s. and r_0 has the distance matrix distribution of the metric measure space (X, r, μ). By Remark 2.1, the extended lookdown space $(\hat{Z}, \rho)$ coincides in this case a.s. with the lookdown space associated with η and r_0. Hence, if $\Xi \in \mathcal{M}_{\mathrm{nd}}$ and $m(\hat{Z} \times \{0\}) = 1$, then the assertions of Theorem 3.1 also hold in the context of the present subsection.

Moreover, in the case without dust, the following corollary to Theorem 3.1 also holds for the extended lookdown space and the more general initial configuration (r_0, v_0) in the present subsection.

Corollary 3.13. *Assume $\Xi \in \mathcal{M}_{\mathrm{nd}}$. Then a.s., the probability measures μ_t exist for all $t \in (0,\infty)$. The map $t \mapsto \mu_t$ is a.s. càdlàg on $(0,\infty)$ in the weak topology on the space of probability measures on $(\hat{Z}, \rho)$, and Θ_0 is a.s. the set of jump times. Moreover, the family of probability measures $(\mu_t, t \in (0,\infty))$ satisfies a.s. assertion (iii) of Theorem 3.1. A.s., also the probability measures m_t exist for all $t \in (0,\infty)$ and satisfy $m_t = \mu_t \otimes \delta_0$.*

Proof. Let (Z', ρ') be the lookdown space associated with η and $(0)_{i,j\in\mathbb{N}}$. By Theorem 3.1, the probability measures μ'_t defined on Z' analogously to μ_t satisfy the assertion.

Let $\varepsilon > 0$, let Z_ε be the closure of $[\varepsilon,\infty)\times\mathbb{N}$ in $(\hat{Z}, \rho)$, and Z'_ε the closure of $[\varepsilon,\infty)\times\mathbb{N}$ in (Z', ρ'). The construction in Section 2 yields $\rho'((t,i),(u,j)) \wedge \varepsilon = \rho((t,i),(u,j)) \wedge \varepsilon$ for all (t,i), $(u,j) \in [\varepsilon,\infty)\times\mathbb{N}$. Hence, the map from $[\varepsilon,\infty)\times\mathbb{N} \subset Z'_\varepsilon$ to Z_ε, given by $(t,i) \mapsto (t,i)$, can be extended to a homeomorphism $h : Z'_\varepsilon \to Z_\varepsilon$. Hence a.s., the weak limits

$$\mu_t = \text{w-}\lim_{n\to\infty} \mu_t^n = \text{w-}\lim_{n\to\infty} h(\mu_t'^n) = h(\mu'_t)$$

exist for all $t \in [\varepsilon,\infty)$ and the assertion on $(\mu_t, t \in (0,\infty))$ follows. The assertion on $(m_t, t \in (0,\infty))$ now follows from Remark 3.11. □

4 Stochastic processes

4.1 The case without dust

From the construction on the lookdown space in Section 3.1, we now read off stochastic processes with values in the space $\mathbb{M}$ of isomorphy classes of metric measure spaces and in the space $\mathbf{M}$ of strong isomorphy classes of compact metric measure spaces. First we recall these state spaces from the literature [41, 45, 70].

As stated in the introduction, we call two metric measure spaces (X', r', μ'), (X'', r'', μ'') isomorphic if there exists an isometry φ from the support supp $(\mu') \subset X'$ to supp $\mu'' \subset X''$ with $\mu'' = \varphi(\mu')$. We denote the isomorphy class by $[\![X', r', \mu']\!]$. We endow the space $\mathbb{M}$ of isomorphy classes of metric measure spaces with the Gromov-Prohorov metric d_{GP} which is complete and separable and induces the Gromov-weak topology, as shown in [45].

Moreover, we call two metric measure spaces (X', r', μ'), (X'', r'', μ'') strongly isomorphic if they are measure-preserving isometric, that is, if there exists a surjective isometry $\varphi : X \to X'$ with $\mu'' = \varphi(\mu')$. We denote the strong isomorphy class by $[X', r', \mu']$. We endow the space $\mathbf{M}$ of strong isomorphy classes of compact metric measure spaces with the Gromov-Hausdorff-Prohorov metric d_{GHP}, given by

$$d_{\mathrm{GHP}}((X', r', \mu'), (X'', r'', \mu'')) = \inf_{Y, \varphi', \varphi''} \{d_{\mathrm{P}}^Y(\varphi'(\mu'), \varphi''(\mu'')) \vee d_{\mathrm{H}}^Y(\varphi'(X'), \varphi''(X''))\}$$

where the infimum is over all isometric embeddings $\varphi' : X' \to Y$, $\varphi'' : X'' \to Y$ into complete and separable metric spaces Y; the Hausdorff distance over Y is denoted by d_{H}^Y. Then $(\mathbf{M}, d_{\mathrm{GHP}})$ is a complete and separable metric space, see [41, 70], and d_{GHP} induces the Gromov-Hausdorff-Prohorov topology on $\mathbf{M}$. The Hausdorff distance in the definition of d_{GHP} compares the metric spaces also where the probability measures charges them with negligible mass.

We work with the lookdown space (Z, ρ), the families of sampling measures $(\mu_t, t \in \mathbb{R}_+)$, and the subspaces X_t of the lookdown space from Section 3.1. Recall that the randomness comes from a Poisson random measure η that is characterized by $\Xi \in \mathcal{M}_{\mathrm{nd}}$, and from an independent random variable ρ_0 with the distance matrix distribution of a metric measure space (X, r, μ).

Proposition 4.1. *Assume $\Xi \in \mathcal{M}_{\mathrm{nd}}$. Then a Markovian stochastic process with values in $\mathbb{M}$ is given a. s. by $([\![Z, \rho, \mu_t]\!], t \in \mathbb{R}_+)$.*

Proposition 4.2. *Assume $\Xi \in \mathcal{M}_{\mathrm{CDI}}$ and that (X, r) is compact. Then a Markovian stochastic process with values in $\mathbf{M}$ is given a. s. by $([X_t, \rho, \mu_t], t \in \mathbb{R}_+)$.*

By Remark 3.3, the assumption $\Xi \in \mathcal{M}_{\mathrm{CDI}}$ ensures that a. s., the spaces X_t are compact for all $t \in \mathbb{R}_+$. We call the process in Proposition 4.1 an $\mathbb{M}$-valued Ξ-Fleming-Viot process with initial state $[\![X, r, \mu]\!]$. By Remark 4.4 below, this process is the $\mathbb{U}$-valued Ξ-Fleming-Viot process from Section 8.1 of Chapter 2 if ρ_0 is a semi-ultrametric. We call the process in Proposition 4.2 an $\mathbf{M}$-valued evolving Ξ-coalescent starting from $[X, r, \mu]$. Note that $[\![Z, \rho, \mu_t]\!]$ in Proposition 4.1 above can be replaced by $[\![X_t, \rho, \mu_t]\!]$ as μ_t is supported by X_t for all $t \in \mathbb{R}_+$ a. s.

Remark 4.3 (Isomorphy classes and strong isomorphy classes). In this remark, we assume $\Xi \in \mathcal{M}_{\mathrm{CDI}}$ and that (X, r) is compact. By construction, the $\mathbf{M}$-valued Ξ-Fleming-Viot process in Proposition 4.2 depends on (X, r, μ) only through the isomorphy class $[\![X, r, \mu]\!]$. In particular, it does not depend on $X \setminus \mathrm{supp}\, \mu$. As supp $\mu_0 = X_0$ a. s. by Lemma 3.2, the strong isomorphy class $[X, r, \mu]$ is not necessarily the initial state.

Moreover, let $\mathbb{M}_{\mathrm{c}}$ be the space of isomorphy classes of compact metric measure spaces, and let $f : \mathbb{M}_{\mathrm{c}} \to \mathbf{M}$, $[\![X', r', \mu']\!] \mapsto [\mathrm{supp}\, \mu', r', \mu']$ be the function that maps an isomorphy class to the strong isomorphy class of a representative where the measure has full support. Using Theorem 3.4(i), we then obtain that $(f(\chi_t), t \in \mathbb{R}_+)$ is an $\mathbf{M}$-valued evolving Ξ-coalescent if $(\chi_t, t \in \mathbb{R}_+)$ is an $\mathbb{M}$-valued Ξ-Fleming-Viot process with initial state $\chi \in \mathbb{M}_c$. Conversely, let $g : \mathbf{M} \to \mathbb{M}$, $[X', r', \mu'] \mapsto [\![X', r', \mu']\!]$. Then $(g(\mathcal{X}_t), t \in \mathbb{R}_+)$ is an $\mathbb{M}$-valued Ξ-Fleming-Viot process if $(\mathcal{X}_t, t \in \mathbb{R}_+)$ is an $\mathbf{M}$-valued evolving Ξ-coalescent. Also note that g is continuous. By Remark 4.6 below, the function f is measurable.

We denote by $\mathfrak{D}$ the space of semi-metrics on $\mathbb{N}$. We do not distinguish between a semi-metric $\rho' \in \mathfrak{D}$ and the distance matrix $(\rho'(i,j))_{i,j\in\mathbb{N}}$, and we consider $\mathfrak{D}$ as a subspace of $\mathbb{R}^{\mathbb{N}^2}$ which we endow with the product topology. Let the $\mathfrak{D}$-valued Markov process $(\rho_t, t \in \mathbb{R}_+)$ be defined from η and ρ_0 as in equation (2.3). For each $t \in \mathbb{R}_+$, the random variable ρ_t is exchangeable by Proposition 6.3 of Chapter 2. That is, its distribution is invariant under the action of the group of bijections $\mathbb{N} \to \mathbb{N}$ which is defined by $p(\rho') = (\rho'(p(i), p(j)))_{i,j\in\mathbb{N}}$ for $\rho' \in \mathfrak{D}$ and a bijection $p : \mathbb{N} \to \mathbb{N}$.

Remark 4.4. Recall the measurable map $\psi : \mathfrak{D} \to \mathbb{M}$ from Section 3.4 in Chapter 2. By construction and Theorem 3.1(i), we have $[\![Z, \rho, \mu_t]\!] = \psi(\rho_t)$ for all $t \in \mathbb{R}_+$ a. s.

Proof of Proposition 4.1. By Remark 4.4, $[\![Z, \rho, \mu_t]\!]$ is a random variable. The Markov property follows, precisely as in Theorem 4.2 of Chapter 2, from an application of Theorem 2 of Rogers and Pitman [82] to the Markov process $(\rho_t, t \in \mathbb{R}_+)$, the measurable map $\psi : \mathfrak{D} \to \mathbb{M}$, and the probability kernel from $\mathbb{M}$ to $\mathfrak{D}$ given by $(\chi, B) \mapsto \nu^\chi(B)$. Here we use the exchangeability of ρ_t. □

The proof of Proposition 4.2 is analogous. Instead of the map ψ, we need another map υ which we now define. Let $\mathfrak{D}_c \subset \mathfrak{D}$ be the space of totally bounded semi-metrics on $\mathbb{N}$,

$$\mathfrak{D}_c = \{\rho' \in \mathfrak{D} : \lim_{n\to\infty} \sup_{j>n} \inf_{i\le n} \rho'(i,j) = 0\}.$$

Let $\upsilon : \mathfrak{D}_c \to \mathbf{M}$ be the function that maps $\rho' \in \mathfrak{D}_c$ to the strong isomorphy class $[X', \rho', \mu']$ of the compact metric measure space (X', ρ', μ') defined as follows: (X', ρ') is the completion of the metric space obtained by identifying the elements with ρ'-distance zero in $(\mathbb{N}, \rho')$. The probability measure μ' on (X', ρ') is the weak limit

$$\mu' = \text{w-}\lim_{n\to\infty} n^{-1} \sum_{i=1}^{n} \delta_i$$

if it exists, else we set $\mu' = \delta_1$. The following lemma is analogous to Proposition 3.8 of Chapter 2.

Lemma 4.5. *The function $\upsilon : \mathfrak{D}_c \to \mathbf{M}$ is measurable.*

Again, we refer by an element of a semi-metric space also to the corresponding element of the completion of the metric space that is obtained by identifying points with distance zero.

Proof. For $n \in \mathbb{N}$, let $\mathfrak{D}_n \subset \mathbb{R}^n$ be the space of semi-metrics on $[n]$, again we do not distinguish between semi-metrics and distance matrices. Let $\upsilon_n : \mathfrak{D}_n \to \mathbf{M}$ be the function that maps $\rho' \in \mathfrak{D}_n$ to the strong isomorphy class of the metric measure space $(X', \rho', n^{-1}\sum_{i=1}^n \delta_n)$, where (X', ρ') is the metric space obtained by identifying the elements of $[n]$ with ρ'-distance zero. Clearly, the map υ_n is continuous. To show this formally, we define analogously a metric measure space $(X'', \rho'', n^{-1}\sum_{i=1}^n \delta_i)$ from another $\rho'' \in \mathfrak{D}_n$. From [70, Proposition 6], it follows

$$d_{\mathrm{GHP}}((X', \rho', n^{-1}\sum_{i=1}^n \delta_i), (X'', \rho'', n^{-1}\sum_{i=1}^n \delta_i)) \le \tfrac{1}{2} \max_{i,j \le n} |\rho'(i,j) - \rho''(i,j)|\,,$$

we use the coupling $\nu = n^{-1}\sum_{i=1}^n \delta_{(i,i)}$ on $X' \times X''$ and the correspondence $\mathfrak{R} = \{(i,i) : i \in [n]\} \subset X' \times X''$.

Now let $\rho' \in \mathfrak{D}_c$, and let (X', ρ', μ') be defined as in the definition of $\upsilon(\rho')$ above. Using the definition of the Gromov-Hausdorff-Prohorov metric, we obtain that

$$d_{\mathrm{GHP}}(\upsilon(\rho'), \upsilon_n((\rho'(i,j))_{i,j \le n})) \le d_{\mathrm{P}}^{X'}(\mu', n^{-1}\sum_{i=1}^n \delta_i) \vee d_{\mathrm{H}}^{X'}(X', [n]) \to 0 \quad (n \to \infty)$$

if the weak limit μ' of the measures $n^{-1}\sum_{i=1}^n \delta_i$ on X' exists. This yields the assertion. □

Proof of Proposition 4.2. By Remark 3.3, there exists an event of probability 1 on which $\rho_t \in \mathfrak{D}_c$ for all $t \in \mathbb{R}_+$. By construction and Theorems 3.1(i) and 3.4(i), we have $[X_t, \rho, \mu_t] = \upsilon(\rho_t)$ for all $t \in \mathbb{R}_+$ a. s. Hence, Lemma 4.5 yields that $[X_t, \rho, \mu_t]$ is a random variable. The Markov property follows as in Theorem 4.2 of Chapter 2 from an application of [82, Theorem 2] to the Markov process $(\rho_t, t \in \mathbb{R}_+)$, the measurable map $\upsilon : \mathfrak{D}_c \to \mathbf{M}$, and the probability kernel from $\mathbf{M}$ to $\mathfrak{D}_c$ given by $(\chi, B) \mapsto \nu^\chi(B)$. Here we use exchangeability of ρ_t. To check that Condition (a) in [82, Theorem 2] is satisfied, we note that $\upsilon(\rho') = \upsilon(\rho_t)$ a. s. for $t \in \mathbb{R}_+$ and a random variable ρ' with conditional distribution $\nu^{\upsilon(\rho_t)}$ given $\upsilon(\rho_t)$. This a. s. equality follows as in the proof of Proposition 3.12 of Chapter 2, we also use Theorem 3.4(i). □

Remark 4.6. As a by-product of Lemma 4.5, let us deduce measurability of the canonical map $f : \mathbb{M}_c \to \mathbf{M}$ from Remark 4.3 (this answers a question posed to the author by H. Sulzbach). We consider $\chi \in \mathbb{M}_c$ and a random variable ρ with the distance matrix distribution ν^χ. As in the proof of Proposition 3.12 of Chapter 2, it follows that $f(\chi) = \upsilon(\rho)$ a. s. Hence, for a Borel subset $A \subset \mathbf{M}$, we obtain the equivalence

$$\mathbf{1}\{f(\chi) \in A\} = 1 \quad \Leftrightarrow \quad \int \nu^\chi(d\rho')\, \mathbf{1}\{\upsilon(\rho') \in A\} = 1 \quad \Leftrightarrow \quad \nu^\chi(\upsilon^{-1}(A)) = 1.$$

The function that maps a metric measure space $\chi \in \mathbb{M}_c$ to its distance matrix distribution ν^χ is continuous by definition of the Gromov-Prohorov topology (see [45]). Lemma 4.5 now implies that $\{\chi \in \mathbb{M}_c : f(\chi) \in A\}$ is a measurable subset of $\mathbb{M}_c$. We remark that measurability of f can also be obtained as a consequence of [2, Corollary 5.6] and e. g. [57, Theorem 15.1].

Path regularity of the $\mathbb{M}$-valued Ξ-Fleming-Viot process and the $\mathbf{M}$-valued evolving Ξ-coalescent follows from Section 3.1:

Corollary 4.7. *Assume $\Xi \in \mathcal{M}_{\mathrm{nd}}$. Then a. s., the process $([\![Z, \rho, \mu_t]\!], t \in \mathbb{R}_+)$ has càdlàg paths in the Gromov-weak topology and Θ_0 is the set of jump times.*

Corollary 4.8. *Assume $\Xi \in \mathcal{M}_{\mathrm{CDI}}$. Then a. s., the process $([X_t, \rho, \mu_t], t \in \mathbb{R}_+)$ has càdlàg paths in the Gromov-Hausdorff-Prohorov topology and $\Theta_0 \cup \Theta^{\mathrm{ext}}$ is the set of jump times.*

Proof of Corollary 4.7. By Theorem 3.1(ii) and the definition of the Gromov-Prohorov metric, it follows that a. s., the map $t \mapsto [\![Z, \rho, \mu_t]\!]$ is càdlàg and the set of jump times is not larger than Θ_0. By Theorem 3.1(iii) and as the atomicity properties only depend on the isomorphy classes, it follows that a. s., the set of jump times is not smaller than Θ_0. □

Proof of Corollary 4.8. This is analogous to Corollary 4.7. We use Theorems 3.1(ii) and 3.4(ii), the definition of the Gromov-Hausdorff-Prohorov metric, and the atomicity properties from Theorem 3.1(iii) which are determined by the strong isomorphy classes. We also use that X_{t-} and X_t are isometric if $[X_{t-}, \rho, \mu_{t-}] = [X_t, \rho, \mu_t]$. □

Now we study further properties of the $\mathbf{M}$-valued evolving Ξ-coalescent. In Proposition 4.9 below, we show Feller continuity, namely the continuity with respect to the weak topology of the law of the process at time t as a function of the initial state (also recall here that the map $g : \mathbf{M} \to \mathbb{M}$ in Remark 4.3 is continuous). Feller continuity for tree-valued Ξ-Fleming-Viot processes is shown in Proposition 9.1 of Chapter 2.

Proposition 4.9. *Let (X^n, r^n, μ^n) be a sequence of compact metric measure spaces such that $[\![X^n, r^n, \mu^n]\!]$ converges to $[\![X, r, \mu]\!]$ in the Gromov-weak topology. Assume $\Xi \in \mathcal{M}_{\mathrm{CDI}}$ and let $(\mathcal{X}_t^n, t \in \mathbb{R}_+)$ be an $\mathbf{M}$-valued evolving Ξ-coalescent starting from $[X^n, r^n, \mu^n]$. Then for each $t \in (0, \infty)$, the random variable $\mathcal{X}_t^n$ converges in distribution to $[X_t, \rho, \mu_t]$ in $\mathbf{M}$, endowed with the Gromov-Hausdorff-Prohorov topology.*

Proof. Let $t, \varepsilon > 0$ and $n \in \mathbb{N}$. Let ρ_0^n be a random variable with distribution $\nu^{[\![X^n, r^n, \mu^n]\!]}$ that is independent of η. Recall the definition of ρ_0 from Section 3.1. Let (Z', ρ') be the lookdown space associated with η and ρ_0^n. Let $X_t' \subset Z'$ be the closure of the set $\{t\} \times \mathbb{N}$ of individuals at time t therein, and define a probability measure μ_t' on Z' analogously to μ_t. Then a. s., the map $X_t \to X_t'$, $(t, i) \mapsto (t, i)$ can be extended to a measure-preserving homeomorphism h between (X_t, ρ, μ_t) and (X_t', ρ', μ_t'). The correspondence $\mathfrak{R} = \{(x, h(x)) : x \in X_t\} \subset X_t \times X_t'$ has distortion $\max\{|\rho_0^n(i,j) - \rho_0(i,j)| : i, j \in A_0(t, \mathbb{N})\}$, where we write $A_0(t, \mathbb{N}) = \{A_0(t, \ell) : \ell \in \mathbb{N}\}$. With the coupling $\nu(dx\, dx') = \mu_t(dx)\delta_{h(x)}(dx')$ of μ_t and μ_t', Proposition 6 in [70] implies

$$\begin{aligned} &\mathbb{P}(d_{\mathrm{GHP}}([X_t', \rho', \mu_t'], [X_t, \rho, \mu_t]) \geq \varepsilon) \\ &\leq \mathbb{P}(\max\{|\rho_0^n(i,j) - \rho_0(i,j)| : i, j \in [k]\} \geq 2\varepsilon) + \mathbb{P}(\# A_0(t, \mathbb{N}) > k) \end{aligned}$$

for all $k \in \mathbb{N}$. W. l. o. g., we may assume $\mathcal{X}_t^n = [X_t', \rho', \mu_t']$ for all $t \in \mathbb{R}_+$ a. s., and that the distance matrices ρ_0^n converge in probability. We let n and then k tend to infinity. □

The tree-valued Ξ-Fleming-Viot process converges to equilibrium, as shown in Proposition 10.1 of Chapter 2. A similar result holds for the $\mathbf{M}$-valued evolving Ξ-coalescent.

Assume $\Xi \in \mathcal{M}_{\mathrm{CDI}}$ and (analogously to Section 2.3), let $\bar{\eta}$ be a Poisson random measure on $\mathbb{R} \times \mathcal{P}$ with intensity $dt\ H_\Xi(d\pi)$. From $\bar{\eta}$, we define a lookdown space $(\bar{Z}, \bar{\rho})$ in two-sided time as the completion of the space of individuals $\mathbb{R} \times \mathbb{N}$ with respect to the semi-metric $\bar{\rho}$, given by

$$\bar{\rho}((t,i),(u,j)) = t + u - 2\sup\{s \in (-\infty, t \wedge u] : \bar{A}_s(t,i) = \bar{A}_s(u,j)\}, \tag{4.1}$$

where $\bar{A}_s(t,i)$ denotes the level of the ancestor of the individual (t,i) when particles and reproduction events are defined precisely as in Section 2.

Analogously to Theorem 3.1, on an event of probability 1, the probability measures $\bar{\mu}_t^n = n^{-1}\sum_{i=1}^n \delta_{(t,i)}$ on $(\bar{Z}, \bar{\rho})$ weakly converge as $n \to \infty$ for all $t \in \mathbb{R}$, we denote the limits by $\bar{\mu}_t$. For $t \in \mathbb{R}$, we denote by $\bar{X}_t$ the closure of $\{t\} \times \mathbb{N}$ in $(\bar{Z}, \bar{\rho})$. A stationary $\mathbf{M}$-valued evolving Ξ-coalescent is given by $([\bar{X}_t, \bar{\rho}, \bar{\mu}_t], t \in \mathbb{R})$. We call a random variable that is distributed as $[\bar{X}_0, \bar{\rho}, \bar{\mu}_0]$ an $\mathbf{M}$-valued Ξ-coalescent measure tree, in analogy to the coalescent measure trees in [46, 50]. As $[\bar{X}_0, \bar{\rho}, \bar{\mu}_0]$ is a. s. an ultrametric measure space, this random variable can be seen as a random tree.

In the next proposition, we show convergence to equilibrium for the $\mathbf{M}$-valued Ξ-Fleming-Viot process $([X_t, \rho, \mu_t], t \in \mathbb{R}_+)$ that is defined from η and $[X, r, \mu]$ in this section.

Proposition 4.10. *The $\mathbf{M}$-valued random variable $[X_t, \rho, \mu_t]$, $t \in \mathbb{R}_+$ converges in distribution in $(\mathbf{M}, d_{\mathrm{GHP}})$ to an $\mathbf{M}$-valued Ξ-coalescent measure tree as $t \to \infty$.*

As in [27] and in Chapter 2, we use a coupling argument in the proof. In the present context, the topology is stronger than in Chapter 2, but as we restrict to $\Xi \in \mathcal{M}_{\mathrm{CDI}}$, there exists a coupling of the tree-valued evolving Ξ-coalescents with arbitrary initial state and of the stationary process such that these processes coincide after an a. s. finite random time.

Proof. Assume that the Poisson random measure η is the restriction of $\bar{\eta}$ to $(0, \infty) \times \mathcal{P}$. Then $[\bar{X}_t, \bar{\rho}, \bar{\mu}_t] = [X_t, \rho, \mu_t]$ on the event $\{\mathrm{diam}\ X_t < 2t\}$. By the properties of the Poisson random measure η and as the event $\{\mathrm{diam}\ X_1 < 2\}$ is independent of ρ_0, the events $\{\mathrm{diam}\ X_t < 2\}$, $t \in \mathbb{N}$ are independent and have the same positive probability. Hence, the random time $\tau = \inf\{t \in \mathbb{R}_+ : \mathrm{diam}\ X_t < 2t\}$ is geometrically bounded. The assertion follows as $\mathrm{diam}\ X_t < 2t$ for all $t > \tau$, and as $[\bar{X}_t, \bar{\rho}, \bar{\mu}_t]$ is an $\mathbf{M}$-valued Ξ-coalescent measure tree. □

Remark 4.11 (Convergence of $\mathbb{M}$-valued Ξ-Cannings processes). As an immediate consequence of the uniform convergence in Theorem 3.1(i), we obtain

$$\lim_{n\to\infty} \sup_{t\in[0,T]} d_{\mathrm{GP}}(\llbracket Z, \rho, \mu_t^n \rrbracket, \llbracket Z, \rho, \mu_t \rrbracket) = 0 \quad \text{a. s.} \tag{4.2}$$

for each $T \in \mathbb{R}_+$. The process $(\llbracket Z, \rho, \mu_t^n \rrbracket, t \in \mathbb{R}_+)$ may be called an $\mathbb{M}$-valued Ξ-Cannings process. In the case without simultaneous multiple reproduction events, it coincides with the tree-valued Λ-Cannings process discussed in [61, Section 4.2], and in the case without multiple reproduction events with the tree-valued Moran process from [46, Definition 2.19]. This can be seen by an application of [82, Theorem 2] similarly to the proof of Lemma 6.8 below, see also Section 2 of [23]. Then the convergence (4.2) implies the assertion of Theorem 2 in [46] for a special choice of the approximating sequence of the initial state.

From the uniform convergence in Theorem 3.9(i), similar statements can be deduced for processes from the next subsection. In Chapter 5, convergence of tree-valued Cannings chains is studied by different methods.

4.2 The general case

We include the case with dust by using marked metric measure spaces. Let us first recall some facts from [24]. An ($\mathbb{R}_+$)-marked metric measure space (X, r, m) is a triple 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 $X \times \mathbb{R}_+$. Two marked metric measure spaces (X, r, m), (X', r', m') are called isomorphic if there exists an isometry φ between the supports supp $m(\cdot \times \mathbb{R}_+) \subset X$ and supp $m'(\cdot \times \mathbb{R}_+) \subset X'$ such that the measurable map $\hat{\varphi}$: supp $m \to$ supp m', given by $\hat{\varphi}(x, v) = (\varphi(x), v)$, satisfies $m' = \hat{\varphi}(m)$. We endow the space $\hat{\mathbb{M}}$ of isomorphy classes of marked metric measure spaces with the marked Gromov-Prohorov metric, which is defined by

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

where the infimum is over all isometric embeddings $\varphi : X \to Y$, $\varphi' : X' \to Y$ into complete and separable metric spaces (Y, d^Y). The maps $\hat{\varphi} : X \times \mathbb{R}_+ \to Y \times \mathbb{R}_+$ and $\hat{\varphi}' : X' \times \mathbb{R}_+ \to Y \times \mathbb{R}_+$ are defined by $\hat{\varphi}(x, v) = (\varphi(x), v)$ and $\hat{\varphi}'(x', v) = (\varphi'(x'), v)$. The space $Y \times \mathbb{R}_+$ is endowed with the product metric $d^{Y \times \mathbb{R}_+}((y, v), (y', v')) = d^Y(y, y') \vee |v - v'|$. Then $(\hat{\mathbb{M}}, d_{\mathrm{mGP}})$ is a complete and separable metric space. The marked distance matrix distribution $\nu^{(X,r,m)}$ of (an isomorphy class of) a marked metric measure space (X, r, m) is defined as the distribution of the random variable $(r(x(i), x(j))_{i,j \in \mathbb{N}}, (v(i))_{i \in \mathbb{N}})$ where $(x(i), v(i))_{i \in \mathbb{N}}$ is an m-iid sequence in $X \times \mathbb{R}_+$. The metric d_{mGP} induces the marked Gromov-weak topology in which a sequence of marked metric measure spaces converges if and only if their marked distance matrix distributions converge weakly.

From the construction on the extended lookdown space in Section 3.2, we now read off a stochastic processes with values in $(\hat{\mathbb{M}}, d_{\mathrm{mGP}})$.

We work with the extended lookdown space $(\hat{Z}, \rho)$ as defined in Section 3.2. This random metric space is constructed from the Poisson random measure η which is characterized by a finite measure Ξ on the simplex Δ, and from an independent random variable (r_0, v_0) with the marked distance matrix distribution of a marked metric measure space (X, r, μ).

Proposition 4.12. *Assume that Ξ is a finite measure on Δ. Then a Markovian stochastic process with values in $\hat{\mathbb{M}}$ is given a. s. by $([\![\hat{Z}, \rho, m_t]\!], t \in \mathbb{R}_+)$.*

We call $([\![\hat{Z}, \rho, m_t]\!], t \in \mathbb{R}_+)$ an $\hat{\mathbb{M}}$-valued Ξ-Fleming-Viot process with initial state $[\![X, r, m]\!]$. By Remark 4.13 below, this process is the tree-valued Ξ-Fleming-Viot process from Section 8.2 of Chapter 2 if the restriction of ρ to $\{0\} \times \mathbb{N}$ is ultrametric.

Recall the space $\mathfrak{D}$ of semi-metrics on $\mathbb{N}$. We define the space $\hat{\mathfrak{D}} = \mathfrak{D} \times \mathbb{R}_+^{\mathbb{N}} \subset \mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}}$, where $\mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}}$ is endowed with the product topology. The elements of $\hat{\mathfrak{D}}$ are called marked distance matrices or decomposed semi-metrics on $\mathbb{N}$. We define a $\hat{\mathfrak{D}}$-valued stochastic process $((r_t, v_t), t \in \mathbb{R}_+)$ from the Poisson random measure η and (r_0, v_0) as in Remark 2.2. By construction, cf. Section 7 of Chapter 2, this process is Markovian. In Lemma 6.4, we will show that for each $t \in \mathbb{R}_+$, the $\hat{\mathfrak{D}}$-valued random variable (r_t, v_t) is exchangeable. That is, its distribution is invariant under the action of the group of bijections $\mathbb{N} \to \mathbb{N}$, defined by $p(r, v) = ((r(p(i), p(j)))_{i,j\in\mathbb{N}}, (v(p(i)))_{i\in\mathbb{N}})$ for $(r, v) \in \hat{\mathfrak{D}}$ and any bijection $p : \mathbb{N} \to \mathbb{N}$.

Remark 4.13. Recall the measurable map $\hat{\psi} : \hat{\mathfrak{D}} \to \hat{\mathbb{M}}$ from Section 3.4 of Chapter 2. Theorem 3.9 yields $[\![\hat{Z}, \rho, m_t]\!] = \hat{\psi}(r_t, v_t)$ for all $t \in \mathbb{R}_+$ a. s.

Proof of Proposition 4.12. The assertion follows as in Theorem 4.2 of Chapter 2, we apply [82, Theorem 2] to the process $((r_t, v_t), t \in \mathbb{R}_+)$, the measurable map $\hat{\psi} : \hat{\mathfrak{D}} \to \hat{\mathbb{M}}$, and the probability kernel from $\hat{\mathbb{M}}$ to $\hat{\mathfrak{D}}$, given by $(\chi, B) \mapsto \nu^{\chi}(B)$. Here we use the exchangeability of (r_t, v_t). □

Let us deduce path regularity:

Corollary 4.14. *Assume that one of the following conditions hold: (i)* $\Xi \in \mathcal{M}_{\text{dust}}$*, or (ii)* $\Xi \in \mathcal{M}_{\text{nd}}$ *and* $m(X \times \{0\}) = 1$*. Then the process* $([\![\hat{Z}, \rho, m_t]\!], t \in \mathbb{R}_+)$ *has a. s. càdlàg paths in the marked Gromov-weak topology and* Θ_0 *is the set of jump times.*

Proof. We argue as in the proof of Corollary 4.7. In case $\Xi \in \mathcal{M}_{\text{dust}}$, we use Theorem 3.9. If $\Xi \in \mathcal{M}_{\text{nd}}$ and $m(\hat{Z} \times \{0\}) = 1$, we use Remark 3.12 and Theorem 3.1. □

Corollary 4.15. *Assume* $\Xi \in \mathcal{M}_{\text{nd}}$ *and* $m(X \times \{0\}) < 1$*. Then the process* $([\![\hat{Z}, \rho, m_t]\!], t \in (0, \infty))$ *has a. s. càdlàg paths in the marked Gromov-weak topology and* Θ_0 *is the set of jump times. The process* $([\![\hat{Z}, \rho, m_t]\!], t \in \mathbb{R}_+)$ *is a. s. not right-continuous at time* 0*.*

Proof. The first assertion follows from Corollary 3.13. The definitions of m_t and of the marked distance matrix distribution, and Remark 3.11 yield that $\nu^{[\![\hat{Z},\rho,m_t]\!]}(\mathfrak{D} \times \{0\}) = m_t(\hat{Z} \times \{0\}) = 1$ for all $t \in (0, \infty)$ a. s. As the marked metric measure spaces (Z, ρ, m_0) and (X, r, m) are a. s. isomorphic (see e. g. Proposition 3.12 of Chapter 2), the assumptions also yield $\nu^{[\![\hat{Z},\rho,m_0]\!]}(\mathfrak{D} \times \{0\}) = m_0(\hat{Z} \times \{0\}) = m(X \times \{0\}) < 1$ a. s. Hence, the probability measures $\nu^{[\![\hat{Z},\rho,m_t]\!]}$ do a. s. not weakly converge to $\nu^{[\![\hat{Z},\rho,m_0]\!]}$. By definition of the marked Gromov-weak topology, it follows that $t \mapsto [\![\hat{Z}, \rho, m_t]\!]$ is a. s. not right-continuous at 0. □

From the construction in Section 3.2, we also read off the $\mathcal{U}^{\text{erg}}$-valued Ξ-Fleming-Viot process from Section 8.3 of Chapter 2. As in Section 2 of Chapter 2, let

$$\alpha : \hat{\mathfrak{D}} \to \mathfrak{D}, \quad (r, v) \to ((r(i, j) + v(i) + v(j))\, \mathbf{1}\{i \neq j\})_{i,j\in\mathbb{N}} \tag{4.3}$$

be the continuous function that maps a decomposed semi-metric to the corresponding semi-metric. The distance matrix distribution of a marked metric measure space χ is defined as $\alpha(\nu^\chi)$, where ν^χ denotes the marked distance matrix distribution of χ. It depends only on the isomorphy class of the marked metric measure space. Let $\rho_0 = \alpha(r_0, v_0)$, then ρ_0 is the restriction of ρ to $\{0\} \times \mathbb{N}$. As we will apply Proposition 3.3 of Chapter 2, we assume in the remainder of this subsection that ρ_0 is a semi-ultrametric (in which case the restriction of ρ to $\{0\} \times \mathbb{N}$ is ultrametric). Then the stochastic process of the distance matrix distributions of the $\hat{\mathbb{U}}$-valued Ξ-Fleming-Viot process, given by $(\nu_t, t \in \mathbb{R}_+) = (\alpha(\nu^{[\![\hat{Z},\rho,m_t]\!]}), t \in \mathbb{R}_+)$ a. s., is the $\mathcal{U}^{\text{erg}}$-valued Ξ-Fleming-Viot process from Section 8.3 of Chapter 2, where $\mathcal{U}^{\text{erg}}$ refers to the space

$$\mathcal{U}^{\text{erg}} = \{\alpha(\nu^\chi) : \chi \in \hat{\mathbb{M}}, \alpha(\nu^\chi)\text{-a. a. } \rho \in \mathfrak{D} \text{ are semi-ultrametrics}\}$$

of ultrametric distance matrix distributions of marked metric measure spaces. We endow $\mathcal{U}^{\text{erg}}$ with the Prohorov metric.

In spite of Corollary 4.15, the $\mathcal{U}^{\text{erg}}$-valued Ξ-Fleming-Viot process is always a. s. right-continuous at time 0 by the following proposition which will be applied in Chapter 5.

Proposition 4.16. *Assume that Ξ is a finite measure on Δ. Then $(\nu_t, t \in \mathbb{R}_+)$ has a. s. càdlàg paths and Θ_0 is a. s. the set of jump times.*

Proof. Corollaries 4.14 and 4.15, the definition of the marked Gromov-weak topology, and continuity of α yield that on an event of probability 1, the path $(0, \infty) \to \mathcal{U}^{\text{erg}}$, $t \mapsto \nu_t$ is a. s. càdlàg in d_{P} with a. s. no jumps outside Θ_0, and $[\![\hat{Z}, \rho, m_t]\!] \neq [\![\hat{Z}, \rho, m_{t-}]\!]$ for all $t \in \Theta_0$. By Propositions 3.3 and 3.12 of Chapter 2, it follows that $\nu_t \neq \nu_{t-}$ for all $t \in \Theta_0$ a. s.

Right continuity at time 0 follows from Corollary 4.14 under the assumptions therein. For the general case, we use that $(\nu_t, t \in \mathbb{R}_+)$ solves the martingale problem $(C, \mathscr{C})$ defined in Section 8.3 of Chapter 2. We briefly recall the definition of the domain $\mathscr{C}$. For $n \in \mathbb{N}$, let γ_n be the restriction from $\mathbb{R}^{\mathbb{N}^2}$ to $\mathbb{R}^{n^2}$, $\gamma_n(\rho') = (\rho'(i,j))_{i,j \in [n]}$. Let $\mathcal{C}_n$ be the set of functions $\phi \circ \gamma_n : \mathbb{R}^{\mathbb{N}^2} \to \mathbb{R}$, where ϕ is a bounded differentiable function $\mathbb{R}^{n^2} \to \mathbb{R}$ with bounded uniformly continuous derivative. Then we set $\mathcal{C} = \bigcup_{n \in \mathbb{N}} \mathcal{C}_n$ and $\mathscr{C} = \{\mathcal{U}^{\text{erg}} \to \mathbb{R}, \nu \mapsto \nu\phi : \phi \in \mathcal{C}\}$. By definition of C in Chapter 2, $C\Psi$ is bounded for each $\Psi \in \mathscr{C}$.

The set $\mathscr{C}$ is convergence determining in $\mathcal{U}^{\text{erg}}$, see Remark 4.5 in Chapter 2. There also exists a countable subset of $\mathscr{C}$ that generates the weak topology on $\mathcal{U}^{\text{erg}}$. Indeed, by smoothing indicator functions of rational intervals, one finds a countable subset $\mathcal{C}' \subset \mathcal{C}$ that generates the product topology on $\mathbb{R}^{\mathbb{N}^2}$. Let $\mathcal{C}''$ be the set of finite products of functions in $\mathcal{C}'$. Then the algebra $\mathcal{C}''$ is convergence determining in $\mathbb{R}^{\mathbb{N}^2}$ by e. g. [69, Theorem 2.7]. That is, $\mathscr{C}' := \{\nu \mapsto \nu\phi : \phi \in \mathcal{C}''\} \subset \mathscr{C}$ generates the weak topology on $\mathcal{U}^{\text{erg}}$.

For $\Psi \in \mathscr{C}$, we consider the process $(M_t, t \in \mathbb{R}_+)$ that is given by

$$M_t = \Psi(\nu_t) - \int_0^t C\Psi(\nu_s)ds$$

and which is a bounded martingale. From

$$\mathbb{E}[\Psi(\nu_t) - \Psi(\nu_0) - \int_0^t C\Psi(\nu_s)ds] = 0,$$

continuity of Ψ, and as $C\Psi$ is bounded, we obtain that ν_t converges in distribution to ν_0 as $t \to 0$. As in the proof of Corollary 4.15, the marked metric measure spaces (X, r, m) and $(\hat{Z}, \rho, m_0)$ are a. s. isomorphic. Hence, $\nu_0 = \alpha(\nu^{(X,r,m)})$ a. s. As (X, r, m) is deterministic, it follows that ν_t converges to ν_0 also in probability. By martingale convergence and as $(M_t, t \in (0, \infty))$ has a. s. càdlàg paths, it follows that the limit $\lim_{t\downarrow 0} M_t$ exists a. s., see e. g. [83, Theorem II.69.4]. By the convergence in probability we already know, the definition of M_t, and as $C\Psi$ are bounded, the limit must be $\Psi(\nu_0)$ a. s., hence $\Psi(\nu_t)$ converges to $\Psi(\nu_0)$ a. s. As $\mathscr{C}$ contains a countable subset that generates the weak topology on $\mathcal{U}^{\text{erg}}$, it follows that ν_t converges to ν_0 a. s. □

5 Outline and some definitions for the proof of the central results

The aim of the remaining sections is to prove Theorems 3.1, 3.4, and 3.9. In Section 6, we prove exchangeability properties for the (decomposed) genealogical distances between the individuals in the lookdown model at various stopping times. In Section 7, we state convergence results for processes of certain asymptotic frequencies that depend on these (decomposed) genealogical distances. We apply these convergence results in particular to families of partitions in Section 8: In the case without dust, we consider the flow of partitions. For the case with dust, we introduce a family of partitions that fits to the decomposition of genealogical distances. Using these families of partitions, we construct the probability measures on the (extended) lookdown space in Section 9. We prove the convergence results from Section 7 in Section 10 using the exchangeability properties from Section 6.

Now we collect some definitions that we will use in the remaining sections.

5.1 Some notation

For $n \in \mathbb{N}$, we continue using the notation $[n] = \{1, \dots, n\}$. We also write $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ and $[0] = \emptyset$. Recall the space $\mathfrak{D} \subset \mathbb{R}^{\mathbb{N}^2}$ of distance matrices, and that we do not distinguish distance matrices from semi-metrics on $\mathbb{N}$. Recall also the space $\hat{\mathfrak{D}} = \mathfrak{D} \times \mathbb{R}_+^{\mathbb{N}} \subset \mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}}$ of marked distance matrices or decomposed semi-metrics on $\mathbb{N}$. We denote by $\mathfrak{D}_n \subset \mathbb{R}^{n^2}$ the space of semi-metrics on $[n]$ which we do not distinguish from distance matrices. We denote by $\hat{\mathfrak{D}}_n = \mathfrak{D}_n \times \mathbb{R}_+^n \subset \mathbb{R}^{n^2} \times \mathbb{R}^n$ the space of decomposed semi-metrics on $[n]$ or marked distance matrices. Recall also the space $\mathcal{P}$ of partitions of $\mathbb{N}$ which is endowed with the topology induced by the restriction maps. We denote by $\mathcal{P}_n$ of set of partitions of $[n]$. We denote by γ_n the restriction maps $\gamma_n : \mathfrak{D} \to \mathfrak{D}_n$, $\rho \mapsto (\rho(i,j))_{i,j\in[n]}$, $\gamma_n : \hat{\mathfrak{D}} \to \hat{\mathfrak{D}}_n$, $(r, v) \mapsto ((r(i,j))_{i,j\in[n]}, (v(i))_{i\in[n]})$, and $\gamma_n : \mathcal{P} \to \mathcal{P}_n$. For a partition $\pi \in \mathcal{P}$ and $i \in \mathbb{N}$, we denote the block of π that contains i by $B(\pi, i)$.

We denote the set of the minimal elements of the blocks of π by $M(\pi) = \{\min B : B \in \pi\}$. For a subset $B \subset \mathbb{N}$, we denote the relative frequency by $|B|_n = n^{-1}\#(B \cap [n])$ for $n \in \mathbb{N}$, and the asymptotic frequency by $|B| = \lim_{n\to\infty} |B|_n$.

5.2 Two-step construction of the point measure of reproduction events

We will use the following definitions from Section 6.2 onwards. As in Section 2.3, let Ξ be a finite measure on the simplex Δ. We decompose $\Xi = \Xi_0 + \Xi\{0\}\delta_0$. Let η be a Poisson random measure on $(0,\infty) \times \mathcal{P}$ with intensity $dt\, H_\Xi(d\pi)$ as in Section 2.3. We assume w. l. o. g. that η is constructed as the sum $\eta = \eta_{\mathrm{K}} + \eta_0$ of two independent Poisson random measures, defined as follows. We define η_{K} as a Poisson random measure on $(0,\infty) \times \mathcal{P}$ with intensity

$$dt\; a \sum_{1 \le i < j} \delta_{K_{i,j}}(d\pi),$$

where $K_{i,j}$ denotes the partition in $\mathcal{P}$ that contains the block $\{i,j\}$ and apart from that only singletons. The point measure η_{K} encodes the Kingman part, that is, the binary reproduction events. If $\Xi_0(\Delta) = 0$, we set $\eta_0 = 0$. Let ξ_0 be a Poisson random measure on $(0,\infty) \times \Delta$ with intensity $dt\; |x|_2^{-2}\, \Xi_0(dx)$. In case $\Xi_0(\Delta) > 0$, let $((t^k, y^k), k \in \mathbb{N})$ be a collection of ξ_0-measurable random variables with values in $(0,\infty) \times \Delta$ such that

$$\xi_0 = \sum_{k \in \mathbb{N}} \delta_{(t^k, y^k)} \quad \text{a. s.}$$

Let $(\pi^k, k \in \mathbb{N})$ be a collection of $\mathcal{P}$-valued random variables that are conditionally independent given $((t^k, y^k), i \in \mathbb{N})$ such that π^k has conditional distribution $\kappa(y^k, \cdot)$ given $((t^k, y^k), k \in \mathbb{N})$. Thereby, κ is the probability kernel from Δ to $\mathcal{P}$ associated with Kingman's correspondence as in Section 2.3. We define the Poisson random measure η_0 on $(0,\infty) \times \mathcal{P}$ by

$$\eta_0 = \sum_{k \in \mathbb{N}} \delta_{(t^k, \pi^k)}.$$

For each point (t, y) of ξ_0 and the associated point (t, π) of η_0, the vector y gives the asymptotic frequencies (in decreasing order) of the blocks of the partition π, that is, the relative family sizes in the large reproduction event at time t.

Recall that $\Xi \in \mathcal{M}_{\mathrm{dust}}$ implies $\Xi\{0\} = 0$. Also note that the set Θ_0 of times of large reproduction events, defined in (2.6), satisfies

$$\Theta_0 = \{t \in (0,\infty) : \xi_0(\{t\} \times \Delta) > 0\} \quad \text{a. s.}$$

5.3 The general setting

Let (r_0, v_0) be a $\hat{\mathfrak{D}}$-valued random variable that is independent of η. Using the map $\alpha : \hat{\mathfrak{D}} \to \mathfrak{D}$ from (4.3), we define a $\mathfrak{D}$-valued random variable by $\rho_0 = \alpha(r_0, v_0)$. Then ρ_0

can also be considered as an arbitrary $\mathfrak{D}$-valued random variable that is independent of η. In this way, we unify the settings of Sections 3.1 and 3.2.

Let $(\hat{Z}, \rho)$ be the extended lookdown space associated with η and (r_0, v_0) as in Section 2.1. Recall that $(\hat{Z}, \rho)$ contains the lookdown space (Z, ρ) associated with η and ρ_0 as a subspace. We endow $\hat{Z} \times \mathbb{R}_+$ with the product metric $d^{\hat{Z}\times\mathbb{R}_+}$ defined in Section 3.2. We define the $\mathfrak{D}$-valued process $(\rho_t, t \in \mathbb{R}_+)$ of the genealogical distances between the individuals at fixed times as in equation (2.3). Then $(\rho_t, t \in \mathbb{R}_+)$ is a Markov process with a. s. càdlàg paths, see Sections 5 and 6 of Chapter 2. We denote the left limits by ρ_{t-}.

We define the $\hat{\mathfrak{D}}$-valued Markov process $((r_t, v_t), t \in \mathbb{R}_+)$ of the decomposed genealogical distances between the individuals at fixed times as in Section 2.1 (including Remark 2.2). As in Section 7 of Chapter 2, the following condition is a. s. satisfied in case $\Xi \in \mathcal{M}_{\text{dust}}$:

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

Here we denote by $\hat{\mathcal{P}}^n = \{\pi \in \mathcal{P} : \{\{1\}, \dots, \{n\}\} \not\subset \pi\}$ the set of those partitions of $\mathbb{N}$ in which the first n integers are not all in singleton blocks. As in Chapter 2, on the event that condition (5.1) is satisfied, the particles on the first $n \in \mathbb{N}$ levels reproduce at only finitely many times in bounded time intervals. For $i \in [n]$, the maps $t \mapsto (z(t, i), v_t(i))$ have jumps only at such reproduction times. Between such jumps, the parent $z(t, i)$ remains constant and the quantity $v_t(i)$ grows linearly with slope 1. In particular, the process $((r_t, v_t), t \in \mathbb{R}_+)$ is a. s. càdlàg if $\Xi \in \mathcal{M}_{\text{dust}}$. Also recall that $\Xi \in \mathcal{M}_{\text{nd}}$ implies $v_t = 0$ and $\rho_t = r_t$ for all $t \in (0, \infty)$ a. s. by Remark 3.11. Hence, $((r_t, v_t), t \in (0, \infty))$ is a. s. càdlàg. We denote the left limits by (r_{t-}, v_{t-}). Note that the left limits ρ_{t-} and (r_{t-}, v_{t-}) can be defined like ρ_t and (r_t, v_t), respectively, by ignoring a possible reproduction event at time t.

6 Preservation of exchangeability

In this section, we extend the exchangeability results on genealogical distances in the lookdown model from Section 6.1 of Chapter 2: We consider invariance under permutations that leave the first $b \in \mathbb{N}$ levels unchanged, and we show exchangeability properties of the decomposed genealogical distances at various stopping times. In Subsection 6.1, we show that exchangeability properties are preserved in single reproduction events. In Subsection 6.2, we concatenate these reproduction events to show exchangeability properties in the lookdown model. Exchangeability in the lookdown model is also studied in [13, 27, 28, 64].

Let us first repeat from Sections 5 and 7.1 of Chapter 2 the effect of a reproduction event on the (decomposed) genealogical distances. For $n \in \mathbb{N}$, $\pi \in \mathcal{P}_n$ and $i \in [n]$, let $\pi(i)$ be the integer k such that i belongs to the k-th block of π when blocks are ordered increasingly according to their respective smallest element. With each element π of $\mathcal{P}_n$, we associate a transformation $\mathfrak{D}_n \to \mathfrak{D}_n$, which we also denote by π, by

$$\pi(\rho) = (\rho(\pi(i), \pi(j)))_{i,j\in\mathbb{N}}.$$

By comparison with the construction in Section 2, we see that a reproduction event that is encoded by a point (t, π) of η results in a jump of the genealogical distances that is given by

$$\gamma_n(\rho_t) = \gamma_n(\pi)(\gamma_n(\rho_{t-})),$$

which is equation (5.3) in Chapter 2. Recall the set $\mathcal{P}^n$ from equation (2.2). Clearly, $\gamma_n(\pi)(\rho) = \rho$ for each $\pi \in \mathcal{P} \setminus \mathcal{P}^n$ and $\rho \in \mathfrak{D}_n$.

To account for the decomposed genealogical distances, we use the set $\mathcal{S}_n$ of semi-partitions of $[n]$. A semi-partition σ of $[n]$ is a system of nonempty disjoint subsets of $[n]$, which we call blocks. The union $\cup\sigma$ of the blocks needs not comprise all elements of $[n]$. For each semi-partition $\sigma \in \mathcal{S}_n$, there exists a unique partition $\pi \in \mathcal{P}_n$ that has the same non-singleton blocks as σ, that is, $\{B \in \pi : \#B \geq 2\} = \{B \in \sigma : \#B \geq 2\}$; we define $\sigma(i) = \pi(i)$ for $i \in [n]$. With each element σ of $\mathcal{S}_n$, we associate a transformation $\hat{\mathfrak{D}}_n \to \hat{\mathfrak{D}}_n$, which we also denote by σ, by $\sigma(r, v) = (r', v')$, where

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

and

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

for $i, j \in [n]$. Furthermore, we define the map

$$\varsigma_n : \mathcal{P} \to \mathcal{S}_n, \quad \pi \mapsto \{B \cap [n] : B \in \pi, \#B \geq 2\} \setminus \{\emptyset\} \tag{6.1}$$

which removes all singleton blocks from π and restricts the obtained semi-partition to $[n]$. In particular, $\hat{\mathcal{P}}^n = \{\pi \in \mathcal{P} : \varsigma_n(\pi) \neq \emptyset\}$ for all $n \in \mathbb{N}$. By construction, we obtain for each point (t, π) of η that

$$\gamma_n(r_t, v_t) = \varsigma_n(\pi)(\gamma_n(r_{t-}, v_{t-})),$$

this is equation (7.1) in Chapter 2. For each point $\pi \in \mathcal{P} \setminus \hat{\mathcal{P}}^n$ and $(r, v) \in \hat{\mathfrak{D}}_n$, we have $\varsigma_n(\pi)(r, v) = (r, v)$.

For $n \in \mathbb{N}$ and $b \in [n] \cup \{0\}$, we denote by $\mathcal{S}_{n,b}$ the set of semi-partitions $\sigma \in \mathcal{S}_n$ that satisfy $\sigma(i) = i$ for all $i \in [b]$. These are the $\sigma \in \mathcal{S}_n$ such that no element of $[b]$ is in a non-singleton block, that is, for each $i \in [b]$, either $i \notin \cup\sigma$ or $\{i\} \in \sigma$. Therefore,

$$\mathcal{S}_{n,b} = \{\varsigma_n(\pi) : \pi \in \mathcal{P} \setminus \mathcal{P}^b\}. \tag{6.2}$$

Similarly, we define $\mathcal{P}_{n,b}$ as the set of partitions of $[n]$ such that none of the first b integers are in non-singleton blocks. This is the set of partitions $\pi \in \mathcal{P}_n$ with $\pi(i) = i$ for all $i \in [b]$. We have

$$\mathcal{P}_{n,b} = \{\gamma_n(\pi) : \pi \in \mathcal{P} \setminus \mathcal{P}^b\}. \tag{6.3}$$

Let S_n be the group of permutations of $[n]$. We define the action of S_n on $\mathfrak{D}_n$ by $p(\rho) = (\rho(p(i), p(j)))_{i,j \in [n]}$ for $p \in S_n$ and $\rho \in \mathfrak{D}_n$. Analogously, we set $p(r, v) =$

$((r(p(i),p(j)))_{i,j\in[n]},(v(p(i)))_{i\in[n]})$ for $(r,v) \in \hat{\mathfrak{D}}_n$, and $p(\sigma) = \{p(B) : B \in \sigma\}$ for $\sigma \in \mathcal{S}_n \supset \mathcal{P}_n$. For $n \in \mathbb{N}$ and $b \in [n] \cup \{0\}$, we define the group

$$S_{n,b} = \{p \in S_n : \ p(i) = i \text{ for all } i \in [b]\}$$

of permutations of $[n]$ that leave the first b integers unchanged. We say that a random (marked) distance matrix in $\mathfrak{D}_n$ or $\hat{\mathfrak{D}}_n$, or a random (semi-)partition of $[n]$ is (n,b)-exchangeable if its distribution is invariant under the action of $S_{n,b}$. For $n \in \mathbb{N}$, the usual exchangeability is recovered as $(n,0)$-exchangeability. For $b \in \mathbb{N}$, no restriction is meant by (b,b)-exchangeability.

6.1 Single reproduction events

In Lemma 6.1 below, we show that (n,b)-exchangeability of distance matrices is preserved under the transformations associated with independent (n,b)-exchangeable partitions which will later encode reproduction events in the lookdown model. This extends Lemma 6.5 of Chapter 2 to (n,b)-exchangeability. In the subsequent Lemma 6.2, we consider marked distance matrices and semi-partitions.

Lemma 6.1. *Let $n \in \mathbb{N}$ and $b \in [n]\cup\{0\}$. Let ρ be an (n,b)-exchangeable random variable with values in $\mathfrak{D}_n$ and let π be an independent (n,b)-exchangeable random variable with values in $\mathcal{P}_{n,b}$. Then $\pi(\rho)$ is (n,b)-exchangeable.*

Proof. Let $p \in S_{n,b}$. As in the proof of Lemma 6.5 in Chapter 2, see equations (6.2) and (6.3) therein, there exists a map $f : \mathcal{P}_n \to S_n$ that satisfies

$$\pi'(i) = f(\pi')(p(\pi')(p(i))) \tag{6.4}$$

for all $\pi' \in \mathcal{P}_n$ and $i \in [n]$, and

$$\pi'(\rho') = p(p(\pi')(f(\pi')(\rho'))) \tag{6.5}$$

for all $\pi' \in \mathcal{P}_n$ and $\rho' \in \mathfrak{D}_n$.

For each $\pi' \in \mathcal{P}_{n,b}$, the definition of p implies $p(\pi') \in \mathcal{P}_{n,b}$ and $\pi'(i) = i = p(\pi')(p(i))$ for all $i \in [b]$. Hence, $f(\pi') \in S_{n,b}$ for each $\pi' \in \mathcal{P}_{n,b}$.

This allows to conclude analogously to the proof of Lemma 6.5 of Chapter 2. □

Lemma 6.2. *Let $n \in \mathbb{N}$ and $b \in [n] \cup \{0\}$. Let (r,v) be a (n,b)-exchangeable random variable with values in $\hat{\mathfrak{D}}_n$ and let σ be an independent (n,b)-exchangeable random variable with values in $\mathcal{S}_{n,b}$. Then $\sigma(r,v)$ is (n,b)-exchangeable.*

Proof. Recall from (4.3) the map $\alpha : \hat{\mathfrak{D}}_n \to \mathfrak{D}_n$. Let $(r',v') \in \hat{\mathfrak{D}}_n$, $\rho' = \alpha(r',v')$, $\sigma' \in \mathcal{S}_n$, and let π' be the partition in $\mathcal{P}_n$ with the same non-singleton blocks as σ'. Then, $\alpha(\sigma'(r',v')) = \pi'(\rho')$ by definition of the transformations on $\mathfrak{D}_n$ and $\hat{\mathfrak{D}}_n$ associated with each element of $\mathcal{P}_n$ and $\mathcal{S}_n$, respectively. Writing $\sigma'(r',v') = (r'',v'')$, we obtain from the definition of the map α that

$$r'' = (((\pi'(\rho'))_{i,j} - v''(i) - v''(j))\,\mathbf{1}\{i \neq j\})_{i,j\in[n]}. \tag{6.6}$$

Let $p \in S_{n,b}$ and let the map $f : \mathcal{P}_n \to S_n$ be defined as in the proof of Lemma 6.1. For $i \in [n]$, it holds $i \in \cup\sigma'$ if and only if $p(i) \in \cup p(\sigma')$. From equation (6.4), we obtain

$$v''(i) = v'(\pi'(i))\,\mathbf{1}\{i \notin \cup\sigma'\} = v'(f(\pi')(p(\pi')(p(i))))\,\mathbf{1}\{p(i) \notin \cup p(\sigma')\}\,.$$

Using also equations (6.5) and (6.6), we deduce

$$\sigma'(r', v') = p\left(p(\sigma')\left(f(\pi')(r', v')\right)\right).$$

We conclude analogously to the proof of Lemma 6.5 of Chapter 2. □

Recall the partitions of the form $K_{i,j}$ which contain only $\{i, j\}$ as a non-singleton block. These partitions encode binary reproduction events. In the next lemma, we consider the exchangeability after a transformation associated with a partition of the form $K_{i,j}$ is applied to an (n, b)-exchangeable distance matrix.

Lemma 6.3. *Let $n \in \mathbb{N}$, $b \in [n-1] \cup \{0\}$, $i, j \in [b+1]$ with $i < j$, and $\pi = K_{i,j}$. Let ρ be an (n, b)-exchangeable random variable with values in $\mathfrak{D}_n$. Then $\pi(\rho)$ is $(n, b+1)$-exchangeable.*

Proof. For all $k \in [n]$,

$$\pi(k) = \begin{cases} k & \text{if } k < j \\ i & \text{if } k = j \\ k-1 & \text{if } k > j. \end{cases}$$

Let $p \in S_{n,b+1}$. Then,

$$\pi(p^{-1}(k)) = \begin{cases} \pi(k) & \text{if } k \le b+1 \\ p^{-1}(k) - 1 & \text{if } k > b+1. \end{cases}$$

Let $\tilde{p}$ be the permutation in $S_{n,b}$ such that $\tilde{p}(k-1) = p^{-1}(k) - 1$ for all $k \in [n]$ with $k > b+1$. The permutation $\tilde{p}$ indeed exists and is unique as $\tilde{p} \in S_{n,b}$ implies $\tilde{p}(k-1) = k-1$ for all $k \in [b+1]$ with $k \ge 2$, as $p \in S_{n,b+1}$ implies $b < p^{-1}(k) - 1 \le n$ for all $k \in [n]$ with $k > b+1$, and and only one possibility remains for $\tilde{p}(n)$ as p^{-1} is injective.

It follows

$$\pi(p^{-1}(k)) = \tilde{p}(\pi(k))$$

for all $k \in [n]$. To see this, we use that $\pi(k) \le b$ for all $k \in [b+1]$, and that $\pi(k) = k-1$ for all $k \in [n]$ with $k > b+1$. From $\pi(k) = \tilde{p}(\pi(p(k)))$ for all $k \in [n]$, and by definition of the transformations on $\mathfrak{D}_n$ associated with the elements of $\mathcal{P}_n$, it follows $\pi(\rho) = p(\pi(\tilde{p}(\rho)))$. The assertion follows as $\tilde{p}(\rho)$ and ρ are equal in distribution. □

6.2 In the lookdown model

We formulate most results in this subsection for the process of marked distance matrices $((r_t, v_t), t \in \mathbb{R}_+)$. To apply these results to the distance matrices $(\rho_t, t \in \mathbb{R}_+)$, note that the construction in Section 2.1 implies that $\rho_t = \alpha(r_t, v_t)$ for all $t \in \mathbb{R}_+$, with α defined in (4.3). Also recall from Remark 3.11 that $\Xi \in \mathcal{M}_{\mathrm{nd}}$ implies that $v_t = 0$ and $\rho_t = r_t$ for all $t \in (0, \infty)$ a.s.

For $n \in \mathbb{N}$ and $b \in [n] \cup \{0\}$, we say a random variable is conditionally (n, b)-exchangeable (given a sigma-algebra or a random variable) if its conditional distribution is a.s. invariant under the action of the group $S_{n,b}$.

The following lemma generalizes Proposition 6.3 of Chapter 2 and shows that conditioned on the event that until time t no reproduction events affect the genealogical distances between the first b individuals, the marked distance matrix $\gamma_n(r_t, v_t)$ is (n, b)-exchangeable if this holds for $\gamma_n(r_0, v_0)$. This assertion also holds conditionally given the point measure ξ_0.

Lemma 6.4. *Let $b \in \mathbb{N}_0$, $n \in \mathbb{N}$, and $t \in \mathbb{R}_+$. Assume that $\gamma_n(r_0, v_0)$ is (n, b)-exchangeable. Then conditionally given ξ_0, the marked distance matrix*

$$\mathbf{1}\big\{\eta((0, t] \times \mathcal{P}^b) = 0\big\}\, \gamma_n(r_t, v_t)$$

is (n, b)-exchangeable.

The proof is analogous to Proposition 6.3 of Chapter 2. It relies on representations (for instance, equation (6.7)) of the (decomposed) genealogical distances in terms of reproduction events and the growth of the (decomposed) genealogical distances. We use that these constituents preserve exchangeability and that reproduction events that affect the (decomposed) genealogical distances do a.s. not accumulate. The latter property is ensured in case $\Xi \in \mathcal{M}_{\mathrm{nd}}$ by condition (2.1), and in case $\Xi \in \mathcal{M}_{\mathrm{dust}}$ by condition (5.1). To account for the growth of the (decomposed) genealogical distances, we define for $t \in \mathbb{R}_+$ and $n \in \mathbb{N}$ the maps

$$\lambda_t : \mathfrak{D}_n \to \mathfrak{D}_n, \quad \rho \mapsto \rho + \underline{\underline{2}}_n t$$

and

$$\hat{\lambda}_t : \hat{\mathfrak{D}}_n \to \hat{\mathfrak{D}}_n, \quad (r, v) \mapsto (r, v + \underline{1}_n t),$$

where we write $\underline{1}_n = (1)_{i \in [n]}$ and $\underline{\underline{2}}_n = 2(\mathbf{1}\{i \neq j\})_{i,j \in [n]}$. Recall also the restriction $\gamma_n : \mathcal{P} \to \mathcal{P}_n$ and the map ς_n defined in (6.1).

Proof. In this proof, we always condition on the event $\{\eta((0, t] \times \mathcal{P}^b) = 0\}$. This does not affect the distribution of the Poisson random measure $\eta(\cdot \cap ((0, \infty) \times (\mathcal{P} \setminus \mathcal{P}^b)))$.

On an event of probability 1, let (t_1, π_1), $(t_2, \pi_2), \ldots$ with $0 < t_1 < t_2 < \ldots$ be the points of η in $(0, \infty) \times (\mathcal{P}^n \setminus \mathcal{P}^b)$. Let $L = \eta((0, t] \times (\mathcal{P}^n \setminus \mathcal{P}^b))$. Conditionally given ξ_0 and $(t_1, t_2, \ldots)$, the random partitions $\pi_1, \pi_2, \ldots$ are independent and $\gamma_n(\pi_k)$ is (n, b)-exchangeable for each $k \in \mathbb{N}$. Moreover, equation (6.3) yields $\gamma_n(\pi_k) \in \mathcal{P}_{n,b}$ for all $k \in [L]$ a.s. By construction (cf. also Section 5 of Chapter 2) and as we condition on $\{\eta((0, t] \times \mathcal{P}^b) = 0\}$, we have

$$\gamma_n(\rho_t) = \lambda_{t - t_L} \circ \gamma_n(\pi_L) \circ \lambda_{t_L - t_{L-1}} \circ \ldots \circ \gamma_n(\pi_1) \circ \lambda_{t_1}(\gamma_n(\rho_0)) \quad \text{a.s.} \tag{6.7}$$

on $\{L \geq 1\}$, and $\gamma_n(\rho_t) = \lambda_t(\gamma_n(\rho_0))$ a.s. on $\{L = 0\}$. Lemma 6.1 implies that the distance matrix $\gamma_n(\rho_t)$ is (n, b)-exchangeable conditionally given ξ_0. In case $\Xi \in \mathcal{M}_{\mathrm{nd}}$, this also holds for the marked distance matrix (r_t, v_t) as $v_t = 0$ and $r_t = \rho_t$ a. s. by Remark 3.11.

An analogous argument applies in case $\Xi \in \mathcal{M}_{\mathrm{dust}}$. As (5.1) holds a. s. in this case, we can now define $(t_1, \pi_1), (t_2, \pi_2), \ldots$ with $0 < t_1 < t_2 < \ldots$ to be the points of η in $(0, \infty) \times (\hat{\mathcal{P}}^n \setminus \mathcal{P}^b)$ a. s. Now let $L = \eta((0,t] \times (\hat{\mathcal{P}}^n \setminus \mathcal{P}^b)$. Conditionally given ξ_0 and $(t_1, t_2, \ldots)$, the random partitions $\pi_1, \pi_2, \ldots$ are independent and the random semi-partition $\varsigma_n(\pi_k)$ is (n, b)-exchangeable for all $k \in \mathbb{N}$. Moreover, equation (6.2) yields $\varsigma_n(\pi_k) \in \mathcal{S}_{n,b}$ for all $k \in [L]$ a. s. By construction (cf. also Section 7.1 of Chapter 2), we obtain

$$\gamma_n(r_t, v_t) = \hat{\lambda}_{t-t_L} \circ \varsigma_n(\pi_L) \circ \hat{\lambda}_{t_L - t_{L-1}} \circ \ldots \circ \varsigma_n(\pi_1) \circ \hat{\lambda}_{t_1}(\gamma_n(r_0, v_0)) \quad \text{a. s.}$$

on $\{L \geq 1\}$, and $\gamma_n(r_t, v_t) = \hat{\lambda}_t(\gamma_n(r_0, v_0))$ a. s. on $\{L = 0\}$. The assertion follows from Lemma 6.2. □

In the next corollary, we set $(r_{0-}, v_{0-}) = (r_0, v_0)$.

Corollary 6.5. *Assume that (r_0, v_0) is exchangeable. Let τ be a ξ_0-measurable and a. s. finite random time. Then (r_τ, v_τ) and $(r_{\tau-}, v_{\tau-})$ are exchangeable.*

Proof. For $k \in \mathbb{N}$, we define a ξ_0-measurable random time τ^k that assumes countably many values by $\tau^k = (j+1)/k$ on the event $\{\tau \in [j/k, (j+1)/k)\}$ for $j \in \mathbb{N}_0$. For $n \in \mathbb{N}$, $p \in S_n$, and bounded continuous ϕ, Lemma 6.4 with $b = 0$ yields

$$\begin{aligned}\mathbb{E}[\phi(\gamma_n(r_{\tau^k}, v_{\tau^k}))] &= \sum_{j \in \mathbb{N}_0} \mathbb{E}[\phi(\gamma_n(r_{j/k}, v_{j/k})); \tau^k = j/k] \\ &= \sum_{j \in \mathbb{N}_0} \mathbb{E}[\phi(p(\gamma_n(r_{j/k}, v_{j/k}))); \tau^k = j/k] = \mathbb{E}[\phi(p(\gamma_n(r_{\tau^k}, v_{\tau^k})))].\end{aligned}$$

We let k tend to infinity. The assertion follows as $t \mapsto \gamma_n(r_t, v_t)$ is càdlàg a. s. To prove the assertion for $(r_{\tau-}, v_{\tau-})$, we replace τ^k with $\tilde{\tau}^k = \lfloor \tau k \rfloor / k$. □

In the next lemma, we see that at the time of the first reproduction event that affects the genealogical distances between the first b individuals, conditioned on this reproduction event being binary, the matrix of the genealogical distances between the first n individuals is $(n, b+1)$-exchangeable if it is (n, b)-exchangeable at time zero.

Lemma 6.6. *Let $n \in \mathbb{N}$ and $b \in [n-1]$ with $n, b \geq 2$. Assume $\Xi\{0\} > 0$ and that $\gamma_n(\rho_0)$ is (n, b)-exchangeable. Let*

$$\tau = \inf\{t > 0 : \eta((0,t] \times \mathcal{P}^b) > 0\}.$$

Then, $\tau < \infty$ a. s. and the distance matrix

$$\mathbf{1}\{\xi_0(\{\tau\} \times \Delta) = 0\}\, \gamma_n(\rho_\tau)$$

is $(n, b+1)$-exchangeable.

Proof. The assumptions $b \geq 2$ and $\Xi\{0\} > 0$ imply $\tau < \infty$ a.s. and $\mathbb{P}(\xi_0(\{\tau\} \times \Delta) = 0) > 0$. In this proof, we always condition on the event $\{\xi_0(\{\tau\} \times \Delta) = 0\}$.

On an event of probability 1, let $(t_1, \pi_1), (t_2, \pi_2) \ldots$ with $0 < t_1 < t_2 < \ldots$ be the points of η in $(0, \infty) \times \mathcal{P}^n$, and let $L = \eta((0, \tau] \times \mathcal{P}^n)$. Then, $\tau = t_L$ a.s. The partitions $\gamma_n(\pi_1), \gamma_n(\pi_2), \ldots$ are conditionally independent given ξ_0, L, and $(t_1, \ldots, t_L)$. Conditionally given ξ_0, L, and $(t_1, \ldots, t_L)$, the partitions $\gamma_n(\pi_k)$ for $k \in [L-1]$ are also (n, b)-exchangeable and by equation (6.3) in $\mathcal{P}_{n,b}$. On the event $\{\xi_0(\{\tau\} \times \Delta) = 0\}$, conditionally given ξ_0, L and $(t_1, \ldots, t_L)$, the partition $\gamma_n(\pi_L)$ a.s. contains one block with two elements in $[b]$ and apart from that only singleton blocks. By construction (cf. also Section 5 of Chapter 2), we obtain

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

The assertion follows from Lemmas 6.1 and 6.3. □

We also consider exchangeability properties at certain stopping times. For $n \in \mathbb{N}$ and $b \in [n] \cup \{0\}$, we define equivalence classes on $\hat{\mathfrak{D}}_n$ by

$$[\![(r, v)]\!]_{n,b} = \{p(r, v) : p \in S_{n,b}\}$$

for $(r, v) \in \hat{\mathfrak{D}}_n$. For $t \in \mathbb{R}_+$, we denote by $\mathcal{F}_t^{n,b}$ the sigma-algebra generated by

$$([\![\gamma_n(r_s, v_s)]\!]_{n,b}, s \in [0, t]).$$

A filtration is defined by $\mathcal{F}^{n,b} = (\mathcal{F}_t^{n,b}, t \in \mathbb{R}_+)$. We set $D_{n,b} = \{[\![(r, v)]\!]_{n,b} : (r, v) \in \hat{\mathfrak{D}}_n\}$. We will use the following lemma in Section 10 where we consider some relative frequencies in the lookdown model that do not change under permutations that leave the first b individuals fixed.

Lemma 6.7. *Let $n \in \mathbb{N}$ and $b \in [n] \cup \{0\}$. Assume that $\gamma_n(r_0, v_0)$ is (n, b)-exchangeable. Let τ be a finite $\mathcal{F}^{n,b}$-stopping time. Then the marked distance matrix*

$$\mathbf{1}\big\{\eta((0, \tau] \times \mathcal{P}^b) = 0\big\}\, \gamma_n(r_\tau, v_\tau)$$

is (n, b)-exchangeable.

Proof. We show that for each $t \in \mathbb{R}_+$, the marked distance matrix

$$\mathbf{1}\big\{\eta((0, t] \times \mathcal{P}^b) = 0\big\}\, \gamma_n(r_t, v_t)$$

is (n, b)-exchangeable conditionally given $\mathcal{F}_t^{n,b}$. The assertion then follows for stopping times that assume countably many values, and by an approximation argument as in the proof of Corollary 6.5 also for all finite stopping times.

We enlarge the spaces $\hat{\mathfrak{D}}_n$ and $D_{n,b}$ by a coffin state ∂. Let K be the probability kernel from $D_{n,b}$ to $\hat{\mathfrak{D}}_n$ such that $K(\partial, \{\partial\}) = 1$, and such that for all $(r, v) \in \hat{\mathfrak{D}}_n \setminus \{\partial\}$, the probability measure $K([\![(r, v)]\!]_{n,b}, \cdot)$ is the uniform distribution on $\{p(r, v) : p \in S_{n,b}\}$.

Let $\tau' = \inf\{t > 0 : \eta((0,t] \times \mathcal{P}^b) > 0\}$ and set $R_t = \gamma_n(r_t, v_t)$ for $t < \tau'$, and $R_t = \partial$ for $t \geq \tau'$. By assumption, K is a regular conditional distribution of $\gamma_n(r_0, v_0)$ given $[\![\gamma_n(r_0, v_0)]\!]_{n,b}$, analogously to the proof of Lemma 6.8 below. For all $t \in \mathbb{R}_+$, Lemma 6.4 implies that K is a regular conditional distribution of R_t given $[\![R_t]\!]_{n,b}$, where we set $[\![\partial]\!]_{n,b} = \partial$. We apply Theorem 2 of Rogers and Pitman [82] to the Markov process $(R_t, t \in \mathbb{R}_+)$, the measurable function $\hat{\mathfrak{D}}_n \to D_{n,b}$, $R \mapsto [\![R]\!]_{n,b}$, and the probability kernel K to obtain that for each $t \in \mathbb{R}_+$, the random variable R_t has the same conditional distribution given $\mathcal{F}_t^{n,b}$ as given $[\![R_t]\!]_{n,b}$. This now follows from equation (1) in [82] and implies the assertion. □

Furthermore, for $n \in \mathbb{N}$, we denote by S_∞^n the group of bijections $p : \mathbb{N} \to \mathbb{N}$ with $p(i) = i$ for all $i > n$. We define equivalence classes on $\hat{\mathfrak{D}}$ by

$$[\![(r,v)]\!]_n = \{p(r,v) : p \in S_\infty^n\}$$

for $(r,v) \in \hat{\mathfrak{D}}$. For $t \in \mathbb{R}_+$, we denote by $\mathcal{F}_t^n$ the sigma-algebra generated by $([\![r_s, v_s]\!]_n, s \in [0,t])$. A filtration is defined by $\mathcal{F}^n = (\mathcal{F}_t^n, t \in \mathbb{R}_+)$. We set $D_n = \{[\![(r,v)]\!]_n : (r,v) \in \hat{\mathfrak{D}}\}$. We will use the following lemma in Section 8 to study asymptotic frequencies that are invariant under permutation of the first n individuals.

Lemma 6.8. *Let $n \in \mathbb{N}$ and assume that (r_0, v_0) is exchangeable. Let τ be a finite $\mathcal{F}^n$-stopping time. Then $\gamma_n(r_\tau, v_\tau)$ is exchangeable.*

Proof. We show that for each $t \in \mathbb{R}_+$, the marked distance matrix $\gamma_n(r_t, v_t)$ is exchangeable conditionally given $\mathcal{F}_t^n$. Let K be the probability kernel from D_n to $\hat{\mathfrak{D}}$ such that for each $(r,v) \in \hat{\mathfrak{D}}$, the probability measure $K([\![(r,v)]\!]_n, \cdot)$ is the uniform distribution on $\{p(r,v) : p \in S_\infty^n\}$. By Lemma 6.4, the marked distance matrix (r_t, v_t) is exchangeable, hence

$$\begin{aligned}&\mathbb{P}((r_t, v_t) \in B, [\![(r_t, v_t)]\!]_n \in B')\\&= \mathbb{P}(p(r_t, v_t) \in B, [\![p(r_t, v_t)]\!]_n \in B')) = \mathbb{P}(p(r_t, v_t) \in B, [\![(r_t, v_t)]\!]_n \in B'))\end{aligned}$$

for all measurable B, B' and all $p \in S_\infty^n$. This implies that K is a regular conditional distribution of (r_0, v_0) given $[\![(r_0, v_0)]\!]_n$, and of (r_t, v_t) given $[\![(r_t, v_t)]\!]_n$. Now we apply Theorem 2 of [82] to the Markov process $((r_t, v_t), t \in \mathbb{R}_+)$, the measurable function $\hat{\mathfrak{D}} \to D_n$, $(r,v) \mapsto [\![(r,v)]\!]_n$, and the probability kernel K to obtain that for each $t \in \mathbb{R}_+$, the marked distance matrix (r_t, v_t) has the same conditional distribution given $[\![(r_t, v_t)]\!]_n$ as given $\mathcal{F}_t^n$. This implies the assertion as in the proof of Lemma 6.7. □

7 Uniform convergence in the lookdown model

Donnelly and Kurtz [28] prove that the measure-valued processes whose states are the uniform measures of the types on the first n levels in the lookdown model converge a. s. as n tends to infinity. Lemmas 7.1 and 7.2 below give the existence of and uniform convergence to asymptotic frequencies of subsets of individuals that are characterized by the (marked) genealogical distances.

For $\ell \in \mathbb{N}$, let

$$b_\ell : \mathcal{P} \to \mathbb{N}_0, \quad b_\ell(\pi) = \ell - \#(\gamma_\ell(\pi))$$

and

$$\hat{b}_\ell : \mathcal{P} \to \mathbb{N}_0, \quad \hat{b}_\ell(\pi) = \#\{i \in [\ell] : \{i\} \notin \pi\}.$$

Moreover, let

$$N^\ell(I) = \int_{I \times \mathcal{P}} b_\ell(\pi)\, \eta(ds\, d\pi)$$

and

$$\hat{N}^\ell(I) = \int_{I \times \mathcal{P}} \hat{b}_\ell(\pi)\, \eta(ds\, d\pi)$$

for each interval $I \subset (0, \infty)$. The random variable $N^\ell(I)$ is the number of newborn particles on the first ℓ levels in the time interval I. The random variable $\hat{N}^\ell(I)$ counts in the reproducing particles in each reproduction event, in particular, it also takes account of reproduction events in which only the reproducing particles occupy levels in $[\ell]$.

Lemma 7.1. *Assume that ρ_0 is exchangeable. Let $b \in \mathbb{N}$ and let f be a measurable function from $\mathbb{R}_+ \times \mathbb{R}_+^b$ to $\{0,1\}$. For $t \in \mathbb{R}_+$ and $i, n \in \mathbb{N}$, define*

$$Y_i(t) = f(t, (\rho_t(j, b+1+i))_{j \in [b]})$$

and

$$X_n(t) = \frac{1}{n} \sum_{i=1}^n Y_i(t).$$

Impose the assumption on f that

$$|X_n(t) - X_n(s)|\, \mathbf{1}\{\eta((s,t] \times \mathcal{P}^b) = 0\} \le n^{-1} N^{b+1+n}(s,t]$$

for all $n \in \mathbb{N}$ and $0 \le s < t$. Then there exists a process $(X(t), t \in \mathbb{R}_+)$ with $\lim_{n\to\infty} \sup_{t \in [0,T]} |X_n(t) - X(t)| = 0$ *a. s. for all $T \in \mathbb{R}_+$.*

Lemma 7.2. *Assume $\Xi \in \mathcal{M}_{\text{dust}}$ and that (r_0, v_0) is exchangeable. Let $b \in \mathbb{N}$ and let f be a measurable function from $\mathbb{R}_+ \times \mathbb{R}_+^{2b} \times \mathbb{R}_+$ to $\{0,1\}$. For $t \in \mathbb{R}_+$ and $i, n \in \mathbb{N}$, define*

$$Y_i(t) = f(t, (r_t(j, b+i), v_t(j))_{j \in [b]}, v_t(b+i))$$

and

$$X_n(t) = \frac{1}{n} \sum_{i=1}^n Y_i(t).$$

Impose the assumption on f that

$$|X_n(t) - X_n(s)|\, \mathbf{1}\Big\{\eta((s,t] \times \hat{\mathcal{P}}^b) = 0\Big\} \le n^{-1} \hat{N}^{b+n}(s,t]$$

for all $n \in \mathbb{N}$ and $0 \le s < t$. Then there exists a process $(X(t), t \in \mathbb{R}_+)$ with $\lim_{n\to\infty} \sup_{t \in [0,T]} |X_n(t) - X(t)| = 0$ *a. s. for all $T \in \mathbb{R}_+$.*

We defer the proofs of these lemmas to Section 10.

8 Two families of partitions

From the process $(\rho_t, t \in \mathbb{R}_+)$, we now read off the flow of partitions $(\Pi_{s,t}, 0 \leq s \leq t)$. This process corresponds to the dual flow of partitions in Foucart [43] and to the flow of partitions in Labbé [64]. We define the random partition $\Pi_{s,t}$ of $\mathbb{N}$ by

$$i \text{ and } j \text{ are in the same block of } \Pi_{s,t} \quad \Leftrightarrow \quad \rho_t(i,j) < 2(t-s)$$

for all $i, j \in \mathbb{N}$ with $i \neq j$. That is, i and j are in the same block of $\Pi_{s,t}$ if and only if $A_s(t,j) = A_s(t,i)$ which means that the individuals (t,i) and (t,j) have a common ancestor at time s. For each $s \in \mathbb{R}_+$ and $n \in \mathbb{N}$, the process $t \mapsto \gamma_n(\Pi_{s,s+t})$ jumps only at the times of reproduction events that are encoded by a partition in $\mathcal{P}^n$. These times do not accumulate on the event of probability 1 on which condition (2.1) is satisfied. For all $0 \leq s \leq t$, the random partition $\Pi_{s,t}$ is exchangeable. This follows from Lemma 6.4 (where we may assume w. l. o. g., as $\Pi_{s,t}$ is η-measurable, that ρ_0 is exchangeable). We will apply the flow of partitions in the dust-free case.

For application in the case with dust, we define for each $a \in \mathbb{N}$ and $\varepsilon > 0$ a family $(\Pi_t^{a,\varepsilon}, t \in \mathbb{R}_+)$ of partitions of $\mathbb{N}$. Similarly to the partition induced by $\sim^\varepsilon$ in Section 11.2 of Chapter 2, our intention is that individuals at time t whose levels are the same block of $\Pi_t^{a,\varepsilon}$ should have parents that are close to each other in the extended lookdown space (the definition of a parent is given Section 2.1). We will define $\Pi_t^{a,\varepsilon}$ accordingly except for at most one block.

First, we define for each $t \in \mathbb{R}_+$ and $I \subset [a]$ the subset

$$C_t^{a,\varepsilon,I} = \{i \in \mathbb{N} : v_t(i) \geq t\} \cap \bigcap_{k \in I:\, v_t(k) \geq t} \{i \in \mathbb{N} : r_t(i,k) \vee |v_t(i) - v_t(k)| < \varepsilon\}$$
$$\cap \bigcap_{\ell \in [a] \setminus I:\, v_t(\ell) \geq t} \{r_t(i,\ell) \vee |v_t(i) - v_t(\ell)| \geq \varepsilon\}$$

Remark 8.1. Clearly, $(C_t^{a,\varepsilon,I}, I \subset [a])$ is a family of subsets whose union is $\{i \in \mathbb{N} : v_t(i) \geq t\}$. Any two such subsets are either disjoint or equal. For $I \subset [a]$ and $i, j \in C_t^{a,\varepsilon,I}$, the construction in Section 2.1 implies that the parents of the individuals (t,i) and (t,j) are also the parents of the individuals $(0, A_0(t,i))$ and $(0, A_0(t,j))$, respectively. If moreover $I \neq \emptyset$, then the genealogical distance $r_t(i,j)$ between these parents is less than 2ε, and $|v_t(i) - v_t(j)| < 2\varepsilon$.

Now we let $i, j \in \mathbb{N}$ be in the same block of $\Pi_t^{a,\varepsilon}$ if and only if one of the following two conditions is satisfied:

(i) $v_t(i) = v_t(j) < t$ and $r_t(i,j) = 0$

(ii) There exists $I \subset [a]$ such that $i, j \in C_t^{a,\varepsilon,I}$.

Condition (i) means that the individuals (t,i) and (t,j) have the same parent in the (extended) lookdown space, and that this parent lives after time zero. That is, (t,i) and (t,j) have a common ancestor at time zero, and the individuals on each of the ancestral lineages of the individuals (t,i) and (t,j) are in singleton blocks in each reproduction event until these ancestral lineages merge. The individual in which these ancestral lineages merge is the parent of both the individuals (t,i) and (t,j), when we identify individuals with genealogical distance zero. In this sense, the individuals (t,i) and (t,j) may be called siblings.

Using the definitions of $v_t(i)$ and $r_t(i,j)$, it can be seen that for each $n \in \mathbb{N}$, on the event of probability 1 on which condition (5.1) is satisfied, the process $t \mapsto \gamma_n(\Pi_t^{a,\varepsilon})$ jumps only at the times of reproduction events that are encoded by a partition in $\hat{\mathcal{P}}^n$, and that these times do not accumulate.

In the next two lemmas, we show that the asymptotic frequencies in $\Pi_{s,t}$ and $\Pi_t^{a,\varepsilon}$, respectively, exist simultaneously for uncountably many t on an event of probability 1, and that the relative frequencies converge uniformly for t in compact intervals. On an event of probability 1, the left limits

$$\Pi_{s,t-} := \lim_{s'\uparrow t} \Pi_{s,s'} = \{\{j \in \mathbb{N} : \rho_{t-}(i,j) < 2(t-s)\} : i \in \mathbb{N}\}$$

and

$$\Pi_{t-}^{a,\varepsilon} := \lim_{s\uparrow t} \Pi_s^{a,\varepsilon} = \{\{j \in \mathbb{N} : v_{t-}(i) = v_{t-}(j) < t, r_{t-}(i,j) = 0\} : i \in \mathbb{N}\}$$

exist for all $t \in (0,\infty)$ and $s \in [0,t)$. The partitions $\Pi_{s,t-}$ and $\Pi_{t-}^{a,\varepsilon}$ are defined like $\Pi_{s,t}$ and $\Pi_t^{a,\varepsilon}$, respectively, except that a possible reproduction event at time t is ignored. In the next two lemmas, we also show regularity properties in t, and that taking (left) limits in t commutes with taking asymptotic frequencies.

Lemma 8.2. *Let $s, T \in \mathbb{R}_+$ and $b \in \mathbb{N}$. Then,*

$$\lim_{n\to\infty} \sup_{t\in[s,s+T]} \big| |B(\Pi_{s,t}, b)|_n - |B(\Pi_{s,t}, b)| \big| = 0 \quad a.\,s. \tag{8.1}$$

The paths $[s,\infty) \to [0,1]$, $t \mapsto |B(\Pi_{s,t}, b)|$ are càdlàg a. s.

Furthermore, $\lim_{\varepsilon\downarrow 0} |B(\Pi_{s,t-\varepsilon}, b)| = |B(\Pi_{s,t-}, b)|$ for all $t \in (s,\infty)$ a. s.

Lemma 8.3. *Let $T \in \mathbb{R}_+$, $a, k \in \mathbb{N}$, and $\varepsilon > 0$. Assume $\Xi \in \mathcal{M}_{\text{dust}}$ and that (r_0, v_0) is exchangeable. Then,*

$$\lim_{n\to\infty} \sup_{t\in[0,T]} \big| |B(\Pi_t^{a,\varepsilon}, k)|_n - |B(\Pi_t^{a,\varepsilon}, k)| \big| = 0 \quad a.\,s. \tag{8.2}$$

The paths $t \mapsto |B(\Pi_t^{a,\varepsilon}, k)|$ are càdlàg a. s.

Furthermore, $\lim_{s\uparrow t} |B(\Pi_s^{a,\varepsilon}, k)| = |B(\Pi_{t-}^{a,\varepsilon}, k)|$ for all $t \in (0,\infty)$ a. s.

A result similar to Lemma 8.2 is Proposition 2.13 in Labbé [64] which is applied there to study relations between the lookdown model and flows of bridges.

Proof of Lemma 8.2. By η-measurability of the random variables in the assertion, we can assume w. l. o. g. that ρ_0 is exchangeable. By homogeneity, it suffices to consider the case $s = 0$. We choose f in Lemma 7.1 such that

$$f(t, (\rho_t(j, b+1+i))_{j\in[b]}) = \mathbf{1}\{\rho_t(b, b+1+i) < 2(t-s)\}$$

for all $t \in \mathbb{R}_+$ and $i \in \mathbb{N}$. On the right-hand side, we have the indicator variable of the event that b and $b+1+i$ are in the same block of $\Pi_{s,t}$. Hence,

$$X_n(t) \le \frac{b+1+n}{n} \left|B(\Pi_{s,t}, b)\right|_{b+1+n} \le X_n(t) + \frac{b+1}{n}$$

and $X(t) = |B(\Pi_{0,t}, b)|$ for all $t \in \mathbb{R}_+$ with $X_n(t)$ and $X(t)$ as defined in Lemma 7.1. Hence, the convergence (8.1) holds a. s. by Lemma 7.1.

On the event of probability 1 on which condition (2.1) holds, the paths $t \mapsto |B(\Pi_{0,t}, b)|_n$ are càdlàg and $\lim_{s' \uparrow t} |B(\Pi_{0,s'}, b)|_n = |B(\Pi_{0,t-}, b)|_n$ for all $t \in (0,\infty)$ and $n \in \mathbb{N}$. This implies that the paths $t \mapsto |B(\Pi_{0,t}, b)|$ are càdlàg a. s., and that $\lim_{s'\uparrow t} |B(\Pi_{0,s'}, b)| = c_t$ for some $c_t \in [0,1]$ for each $t \in (0,\infty)$ a. s. To show the assertion on the left limits, let $\varepsilon > 0$, and choose on an event of probability 1 a sufficiently large integer n_0 such that $\left||B(\Pi_{0,t}, b)|_n - |B(\Pi_{0,t}, b)|\right| < \varepsilon$ for all $t \in [0,T]$ and $n \ge n_0$. Then, $\limsup_{s'\uparrow t} |B(\Pi_{0,s'}, b)|_n \le c_t + \varepsilon$ and $\liminf_{s'\uparrow t} |B(\Pi_{0,s'}, b)|_n \ge c_t - \varepsilon$ for all $n \ge n_0$ and $t \in [0,T]$. It follows $c_t = |B(\Pi_{0,t-}, b)|$ for all $t \in (0,\infty)$ a. s. □

Proof of Lemma 8.3. We choose $b = a \vee k$ and f in Lemma 7.2 such that

$$\begin{aligned} &f(t, (r_t(j, b+i), v_t(j))_{j\in[b]}, v_t(b+i)) \\ &= \mathbf{1}(\{r_t(k, b+i) = 0, v_t(b+i) < t\} \cup \bigcup_{I \subset [a]} \{k, b+i \in C_t^{a,\varepsilon,I}\}) \end{aligned}$$

for all $t \in \mathbb{R}_+$ and $i \in \mathbb{N}$. Here, the event that k and $b+i$ are in the same block of $\Pi_t^{a,\varepsilon}$ stands in the the indicator variable. Hence,

$$X_n(t) \le \frac{b+n}{n} \left|B(\Pi_t^{a,\varepsilon}, k)\right|_{b+n} \le X_n(t) + \frac{b}{n}$$

and $X(t) = |B(\Pi_t^{a,\varepsilon}, k)|$ for all $n \in \mathbb{N}$ and $t \in \mathbb{R}_+$ a. s., with $X_n(t)$ and $X(t)$ as defined in Lemma 7.2. On the event of probability 1 on which condition (5.1) holds, the processes $t \mapsto |B(\Pi_t^{a,\varepsilon}, k)|_n$ are càdlàg for all $n \in \mathbb{N}$ a. s. Lemma 7.2 implies the assertion. □

To construct probability measures in the next section, we will need families of partitions with proper frequencies. A partition π is said to have proper frequencies if $\sum_{B\in\pi} |B| = 1$, that is, the asymptotic frequencies of its blocks exist and sum up to 1. In case $\Xi \in \mathcal{M}_{\mathrm{nd}}$, the partition $\Pi_{s,t}$ has proper frequencies a. s. for each $0 \le s < t$. This follows from [86] as the $\mathcal{P}$-valued process $(\Pi_{t,(t-s)-}, s \in [0,t))$ is a Ξ-coalescent up to time t. This can be seen, for instance, by comparing the construction from the point measure η in Section 2 and the Poisson construction of Schweinsberg [86]. The next two lemmas show that uncountably many partitions have proper frequencies on an event of probability 1. They give lower bounds on the number of blocks whose asymptotic frequencies add up to $1-\varepsilon$ that are uniform for t in compact intervals.

Lemma 8.4. *Assume $\Xi \in \mathcal{M}_{\mathrm{nd}}$. Let $s \in \mathbb{R}_+$, $T \in (0,\infty)$, and $\varepsilon \in (0,T)$. Then, on an event of probability 1, there exists an integer k such that*

$$\sum_{i \in M(\Pi_{s,t}) \cap [k]} |B(\Pi_{s,t}, i)| > 1 - \varepsilon$$

for all $t \in [s+\varepsilon, s+T]$. *In particular, the partition* $\Pi_{s,t}$ *has proper frequencies for all* $t \in (s,\infty)$ *a. s.*

Lemma 8.5. *Assume* $\Xi \in \mathcal{M}_{\mathrm{dust}}$ *and that* (r_0, v_0) *is exchangeable. Let* $a \in \mathbb{N}$, $\varepsilon, \tilde{\varepsilon} > 0$, *and* $T \in \mathbb{R}_+$. *Then, on an event of probability* 1, *there exists an integer* k *such that*

$$\sum_{i \in M(\Pi_t^{a,\varepsilon}) \cap [k]} |B(\Pi_t^{a,\varepsilon}, i)| > 1 - \tilde{\varepsilon}$$

for all $t \in [0,T]$. *In particular, the partition* $\Pi_t^{a,\varepsilon}$ *has proper frequencies for all* $t \in \mathbb{R}_+$ *a. s.*

A property like the assertion of Lemma 8.4 is also considered in Section 6.1 of Labbé [64].

Proof of Lemma 8.4. Again we assume w. l. o. g. that ρ_0 is exchangeable, and it suffices to consider the case $s = 0$.

Step 1. There exists an event of probability 1 on which for all $t \in (\varepsilon, \infty)$ with $\#\Pi_{0,t} = \infty$, the partitions $\Pi_{0,t}$ and $\Pi_{0,t-}$ do not contain any singleton blocks. Indeed, the partition $\Pi_{0,\varepsilon}$ contains a. s. no singletons. For $t \in (\varepsilon, \infty)$, an implication of $\#\Pi_{0,t} = \infty$ ($\#\Pi_{0,t-} = \infty$) is that $\#\Pi_{\varepsilon,t} = \infty$ ($\#\Pi_{\varepsilon,t-} = \infty$, respectively). This also implies that all individuals at time ε have a descendant at time t (at time $t-$) as the trajectories of the particles in the population model do not cross, see Remark 5.1 of Chapter 2. Hence, the assertion of Step 1 holds on the event of probability 1 that $\Pi_{0,\varepsilon}$ contains no singletons.

Step 2. In this step, we assume $\Xi_0(\Delta) > 0$ and we show that Π_{0,t^k-} has proper frequencies for all $k \in \mathbb{N}$ a. s. By Corollary 6.5, the random partition Π_{0,t^k-} is exchangeable. On the event that the partition $\#\Pi_{0,t^k-}$ has finitely many blocks, it has proper frequencies a. s. If $\mathbb{P}(\#\Pi_{0,t^k-} = \infty) > 0$, then also Π_{0,t^k-}, conditioned on the event $\{\#\Pi_{0,t^k-} = \infty\}$, is exchangeable. The assertion of step 2 now follows from step 1 and Kingman's correspondence.

Step 3. Recall the filtration $\mathcal{F}^\ell$ from Section 6.2. For each $\ell \in \mathbb{N}$, the process

$$\left(\sum_{i \in M(\Pi_{0,t}) \cap [\ell]} |B(\Pi_{s,t}, i)|, t \in \mathbb{R}_+ \right)$$

is adapted with respect to the usual augmentation of $\mathcal{F}^\ell$ and has a. s. càdlàg paths by Lemma 8.2. Hence,

$$\vartheta_{\varepsilon,\ell} := \inf \left\{ t \geq \varepsilon : \sum_{i \in M(\Pi_{0,t}) \cap [\ell]} |B(\Pi_{0,t}, i)| < 1 - \varepsilon \right\}$$

is a stopping time with respect to the usual augmentation of $\mathcal{F}^n$ for all $n \in [\ell]$. As $\vartheta_{\varepsilon,j} \geq \vartheta_{\varepsilon,\ell}$ for integers $j \geq \ell$, it follows that

$$\vartheta_\varepsilon := \sup_{\ell \in \mathbb{N}} \vartheta_{\varepsilon,\ell} = \lim_{\ell \to \infty} \vartheta_{\varepsilon,\ell}$$

is a stopping time with respect to the usual augmentation of $\mathcal{F}^n$ for all $n \in \mathbb{N}$.

By Lemma 6.8, the distance matrix $\rho_{\vartheta_\varepsilon \wedge T}$ is exchangeable, hence the partition $\Pi_{0,\vartheta_\varepsilon \wedge T}$ is exchangeable. On the event that the partition $\Pi_{0,\vartheta_\varepsilon \wedge T}$ has finitely many blocks, it has proper frequencies a. s. If $P(\#\Pi_{0,\vartheta_\varepsilon \wedge T} = \infty) > 0$, then $\Pi_{0,\vartheta_\varepsilon \wedge T}$, conditioned on the event $\{\#\Pi_{0,\vartheta_\varepsilon \wedge T} = \infty\}$, remains exchangeable. It follows from step 1 and Kingman's correspondence that $\Pi_{0,\vartheta_\varepsilon \wedge T}$ has a. s. proper frequencies.

Step 4. A. s., $\big||B_i(\Pi_{0,t})|_n - |B_i(\Pi_{0,t-})|_n\big| \leq 1/n$ for all $n, i \in \mathbb{N}$ and all $t \in (0,\infty) \setminus \{t^k : k \in \mathbb{N}\}$ as only binary reproduction events can occur at these times. Lemma 8.2 now implies $|B_i(\Pi_{0,t})| = |B_i(\Pi_{0,t-})|$ for all $i \in \mathbb{N}$ and $t \in (0,\infty) \setminus \{t^k : k \in \mathbb{N}\}$ a. s. Hence, $|\Pi_{0,t}|_1 = |\Pi_{0,t-}|_1$ for all $t \in (0,\infty) \setminus \{t^k : k \in \mathbb{N}\}$ a. s., where $|\pi|_1$ denotes the sum of the asymptotic frequencies of the blocks of a partition $\pi \in \mathcal{P}$.

It follows that the partitions $\Pi_{0,(\vartheta_\varepsilon \wedge T)-}$ and $\Pi_{0,\vartheta_\varepsilon \wedge T}$ have proper frequencies a. s. Hence, there exists a. s. $\ell \in \mathbb{N}$ such that

$$\sum_{i \in M(\Pi_{0,\vartheta_\varepsilon \wedge T}) \cap [\ell]} |B(\Pi_{0,\vartheta_\varepsilon \wedge T}, i)| > 1 - \varepsilon$$

and

$$\sum_{i \in M(\Pi_{0,(\vartheta_\varepsilon \wedge T)-}) \cap [\ell]} |B(\Pi_{0,(\vartheta_\varepsilon \wedge T)-}, i)| > 1 - \varepsilon.$$

By Lemma 8.2, there exists a. s. $\delta > 0$ such that

$$\sum_{i \in M(\Pi_{0,t}) \cap [\ell]} |B(\Pi_{0,t}, i)| > 1 - \varepsilon$$

for all $t \in (\vartheta_\varepsilon \wedge T - \delta, \vartheta_\varepsilon \wedge T + \delta)$. This implies $\vartheta_{\varepsilon,j} \notin (\vartheta_\varepsilon - \delta, \vartheta_\varepsilon + \delta)$ for all $j \geq \ell$ a. s. on the event $\{\vartheta_\varepsilon < T\}$, hence $\{\vartheta_\varepsilon < T\}$ is a null event.

The assertion follows as $T \in (0,\infty)$ and $\varepsilon \in (0,T)$ can be chosen arbitrarily. □

In the proof of Lemma 8.5, we will use the following lemma which strengthens Lemma 7.6 of Chapter 2.

Lemma 8.6. *Assume $\Xi \in \mathcal{M}_{\text{dust}}$ and let $t \in (0,\infty)$, $i \in \mathbb{N}$. Then a. s. on the event $\{v_t(i) < t\}$, there exists an integer $j \in \mathbb{N} \setminus \{i\}$ with $v_t(i) = v_t(j)$ and $r_t(i,j) = 0$.*

Proof. We use the points (t^k, π^k) of η_0 from Section 5.2. W. l. o. g., we assume $\Xi_0(\Delta) > 0$ and that the times $t^k \wedge t$ are stopping times with respect to the filtration $(\mathcal{F}_s, s \in \mathbb{R}_+)$ that is defined in the end of Section 3.1. The latter property can be achieved e. g. by setting $\Delta_\ell = \{x \in \Delta : 1/(\ell+1) < x_1 \leq 1/\ell\}$ and $t^{\ell,0} = 0$ for each $\ell \in \mathbb{N}$, and $t^{\ell,n} = \inf\{t > t^{\ell,n-1} : \xi_0((t^{\ell,n-1}, t] \times \Delta_\ell) > 0\} \wedge t$ for each $n \in \mathbb{N}$, and associating each $k \in \mathbb{N}$ with a pair (ℓ, n).

Lemma 6.4 yields for each $k \in \mathbb{N}$ that the sequence $(\mathbf{1}\{t - v_t(j) \leq t^k\}, j \in \mathbb{N})$ is exchangeable. Here we also use that this sequence is η-measurable so that we can assume w. l. o. g. that (r_0, v_0) is exchangeable. The de Finetti theorem implies that a. s. either no or infinitely many elements of this sequence equal 1. Note that the random subset

$A_k := \{A_{t^k}(t,j) : j \in \mathbb{N}, t - v_t(j) \leq t^k\} \subset \mathbb{N}$ is measurable with respect to t^k and the process $(J_{t^k+s} - J_{t^k}, s \in \mathbb{R}_+)$, where J is defined in (3.1). Now the strong Markov property of J, our assumption on t^k, and the definition of π^k yield that A_k and π^k are independent. As $\Xi\{0\} = 0$, by Kingman's correspondence, and as A_k and π^k are independent, all non-singleton blocks of π^k have an infinite intersection with A_k a. s. on the event $\{\#A_k = \infty\}$. Furthermore, the definition of v_{t^k} yields that $A_{t^k}(t,j) \neq A_{t^k}(t,j')$ for all distinct $j, j' \in \mathbb{N}$ with $t - v_t(j) \leq t^k$, $t - v_t(j') \leq t^k$. Hence, $\#A_k = \#\{j \in \mathbb{N} : t - v_t(j) \leq t^k\}$.

By definition of $v_t(i)$ and condition (5.1), there exists a. s. on $\{v_t(i) < t\}$ an integer $k \in \mathbb{N}$ such that $t - v_t(i) = t^k$ and $A_{t^k}(t,i)$ is in a non-singleton block of π^k. The above implies that there exists a. s. on $\{v_t(i) < t\}$ an integer $j \in \mathbb{N} \setminus \{i\}$ such that $t - v_t(j) \leq t^k$ and $A_{t^k}(t,j)$ is in the same block of π^k as $A_{t^k}(t,i)$. The definition of $v_t(j)$ now yields $v_t(j) = v_t(i)$ and $r_t(i,j)$=0. □

We will apply Lemma 8.6 through the following corollary.

Corollary 8.7. *Assume $\Xi \in \mathcal{M}_{\text{dust}}$ and that (r_0, v_0) is exchangeable. Let $a \in \mathbb{N}$ and $\varepsilon > 0$. Then on an event of probability 1, none of the partitions $\Pi_t^{a,\varepsilon}$, $t \in \mathbb{R}_+$ contains singleton blocks.*

Proof. Let $q \in \mathbb{R}_+$. A. s. by Lemma 8.6, only integers $i \in \mathbb{N}$ with $v_q(i) \geq q$ can form singleton blocks of $\Pi_q^{a,\varepsilon}$. As those integers belong to finitely many blocks by definition of $\Pi_q^{a,\varepsilon}$, it follows that $\Pi_q^{a,\varepsilon}$ contains a. s. at most finitely many singleton blocks. As the partition $\Pi_q^{a,\varepsilon}$ is exchangeable by Lemma 6.4, Kingman's correspondence implies that $\Pi_q^{a,\varepsilon}$ contains a. s. no singleton blocks.

On the event of probability 1 on which condition (5.1) holds, there exists for each $t \in \mathbb{R}_+$ and $i \in \mathbb{N}$ a time $q(t,i) \in (t,\infty) \cap \mathbb{Q}$ with $\eta((t, q(t,i)] \times \hat{\mathcal{P}}^i) = 0$, as (5.1) implies that the points of $\eta(\cdot \times \hat{\mathcal{P}}^i)$ do not accumulate. By construction, if i forms a singleton block in $\Pi_t^{a,\varepsilon}$, then i forms a singleton block also in $\Pi_{q(t,i)}^{a,\varepsilon}$, as the particle remains on level i and does not reproduce. This implies the assertion. □

Proof of Lemma 8.5. We proceed similarly to the proof of Lemma 8.4. We assume w. l. o. g. $\Xi_0(\Delta) > 0$.

Step 1. For all $k \in \mathbb{N}$, the partition $\Pi_{t^k-}^{a,\varepsilon}$ is exchangeable by Corollary 6.5. By Corollary 8.7 and Kingman's correspondence, it follows that $\Pi_{t^k-}^{a,\varepsilon}$ has proper frequencies a. s.

Step 2. We set for $\ell \in \mathbb{N}$

$$\vartheta_{\tilde{\varepsilon},\ell} = \inf\left\{ t \geq 0 : \sum_{i \in M(\Pi_t^{a,\varepsilon}) \cap [\ell]} |B(\Pi_t^{a,\varepsilon}, i)| < 1 - \tilde{\varepsilon} \right\},$$

We deduce from Lemma 8.3 that $\vartheta_{\tilde{\varepsilon},\ell}$ is a stopping time with respect to the usual augmentation of $\mathcal{F}^n$ for all $n \in [\ell]$. Then we define $\vartheta_{\tilde{\varepsilon}} = \sup_{\ell \in \mathbb{N}} \vartheta_{\tilde{\varepsilon},\ell}$ which is for all $n \in \mathbb{N}$ a stopping time with respect to the usual augmentation of $\mathcal{F}^n$. Let $T \in \mathbb{R}_+$. The partition $\Pi_{\vartheta_{\tilde{\varepsilon}} \wedge T}^{a,\varepsilon}$ is exchangeable by Lemma 6.8. We deduce as in step 1 that it has proper frequencies a. s.

Step 3. We conclude as in the proof of Lemma 8.4, using Lemma 8.3. □

Recall the set of measures $\mathcal{M}_{\text{CDI}}$ from Section 3. The following lemma will be used in the proof of Theorem 3.4(i).

Lemma 8.8. *Assume* $\Xi \in \mathcal{M}_{\text{CDI}}$. *Let* $s \in \mathbb{R}_+$. *Then a. s.,* $|B(\Pi_{s,t}, i)| > 0$ *for all* $t \in (s, \infty)$ *and* $i \in \mathbb{N}$.

Proof. Again we assume w. l. o. g. that ρ_0 is exchangeable, and it suffices to consider the case $s = 0$. Let $\varepsilon > 0$. For each $k \in \mathbb{N}$, the process

$$\left(\min_{i\in[k]} |B(\Pi_{0,t}, i)|, t \in \mathbb{R}_+\right)$$

is adapted with respect to the usual augmentation of the filtration $\mathcal{F}^k$ (defined in Section 6.2) and has a. s. càdlàg paths by Lemma 8.2. Consequently,

$$\vartheta_k := \inf\left\{t \geq \varepsilon : \min_{i\in[k]} |B(\Pi_{0,t}, i)| = 0\right\}$$

is a stopping time with respect to the usual augmentation of $\mathcal{F}^n$ for all $n \in [k]$. It follows that

$$\vartheta := \inf_{k\in\mathbb{N}} \vartheta_k = \lim_{k\to\infty} \vartheta_k$$

is a stopping time with respect to the usual augmentation of $\mathcal{F}^n$ for all $n \in \mathbb{N}$. By Lemma 6.8, the distance matrix $\gamma_n(\rho_{\vartheta\wedge T})$ is exchangeable for each $T \in [\varepsilon, \infty)$ and $n \in \mathbb{N}$. Hence, the partition $\Pi_{0,\vartheta\wedge T}$ is exchangeable. The assumption $\Xi \in \mathcal{M}_{\text{CDI}}$ implies $\#\Pi_{0,\vartheta\wedge T} < \infty$ a. s. Kingman's correspondence now implies that each block of $\Pi_{0,\vartheta\wedge T}$ has a positive asymptotic frequency a. s. Hence, by Lemma 8.2, there exists an event of probability 1 on which all blocks of the partitions $\Pi_{0,t}$ with t in a right neighborhood of $\vartheta \wedge T$ have positive asymptotic frequencies. By definition of ϑ, it follows $\mathbb{P}(\vartheta < T) = 0$. The assertion follows as T and ε can be chosen arbitrarily. □

9 The construction on the lookdown space

Now we apply the results from the last section to prove the assertions from Section 3. We use the coupling characterization of the Prohorov distance, namely (see e. g. [36, Theorem 3.1.2]) that in a separable metric space (X, r), the Prohorov distance between probability measures μ and μ' on the Borel sigma algebra is given by

$$d_{\text{P}}^X(\mu, \mu') = \inf_{\nu} \inf\{\varepsilon > 0 : \nu\{(x, y) \in X^2 : r(x, y) > \varepsilon\} < \varepsilon\},$$

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

9.1 The case with dust

In this subsection, we always consider the case $\Xi \in \mathcal{M}_{\text{dust}}$. Let $(x(i), v(i))_{i\in\mathbb{N}}$ be an m-iid sequence in a marked metric measure space (X, r, m), independent of η. We set $(r_0, v_0) =$

$((r(x(i),x(j)))_{i,j\in\mathbb{N}},v)$. Then (r_0,v_0) is distributed according to the marked distance matrix distribution of (X,r,m). With the extended lookdown space $(\hat{Z},\rho)$ associated with η and (r_0,v_0), we are in the setting of Subsection 3.2.

Proof of Theorem 3.9 (beginning). We begin with the proof of item (i). For $\varepsilon>0$ and $a,n\in\mathbb{N}$, we define the probability measures

$$m_t^{a,\varepsilon,n}=\sum_{i\in M(\Pi_t^{a,\varepsilon})}|B(\Pi_t^{a,\varepsilon},i)|_n\,\delta_{(z(t,i),v_t(i))}.$$

on $\hat{Z}\times\mathbb{R}_+$. Clearly, $|B(\Pi_t^{a,\varepsilon},i)|_n=0$ for all $i\in M(\Pi_t^{a,\varepsilon})$ with $i>n$.

Let $T\in\mathbb{R}_+$. Using that the Prohorov distance is bounded from above by the total variation distance, we obtain

$$\begin{aligned}
&\lim_{n,\ell\to\infty}\sup_{t\in[0,T]}d_{\mathrm{P}}(m_t^{a,\varepsilon,n},m_t^{a,\varepsilon,\ell})\\
&\le\lim_{n,\ell\to\infty}\sup_{t\in[0,T]}\sum_{i\in M(\Pi_t^{a,\varepsilon})}\big|\,|B(\Pi_t^{a,\varepsilon},i)|_n-|B(\Pi_t^{a,\varepsilon},i)|_\ell\,\big|\\
&\le\lim_{k\to\infty}\lim_{n,\ell\to\infty}\sup_{t\in[0,T]}\sum_{i\in M(\Pi_t^{a,\varepsilon})\cap[k]}\big|\,|B(\Pi_t^{a,\varepsilon},i)|_n-|B(\Pi_t^{a,\varepsilon},i)|_\ell\,\big|\\
&\quad+2\lim_{k\to\infty}\lim_{n\to\infty}\sup_{t\in[0,T]}\sum_{\substack{i\in M(\Pi_t^{a,\varepsilon}):\\ i>k}}|B(\Pi_t^{a,\varepsilon},i)|_n=0\quad\text{a.s.}
\end{aligned}$$

The first summand on the right-hand side equals zero a. s. by Lemma 8.3. A. s., the second summand equals zero as for each $\tilde{\varepsilon}>0$, there exist integers k and n_0 such that

$$\inf_{t\in[0,T]}\sum_{i\in M(\Pi_t^{a,\varepsilon})\cap[k]}|B(\Pi_t^{a,\varepsilon},i)|_n\ge 1-\tilde{\varepsilon}$$

for all $n\ge n_0$ by Lemmas 8.5 and 8.3.

Now we compare the probability measures m_t^n and $m_t^{a,\varepsilon,n}$. A coupling ν of these probability measures is given by

$$\nu\{((z(t,i)),v_t(i)),(z(t,j),v_t(j)))\}=1/n$$

for $i\in[n]$ and $j=\min B(\Pi_t^{a,\varepsilon},i)$. By definition of $\Pi_t^{a,\varepsilon}$ and Remark 8.1,

$$d^{\hat{Z}\times\mathbb{R}_+}((z(t,i),v_t(i)),(z(t,j),v_t(j)))=r_t(i,j)\vee|v_t(i)-v_t(j)|<2\varepsilon$$

for all $i,j\in\mathbb{N}$ that are in the same block of $\Pi_t^{a,\varepsilon}$ and not in $C_t^{a,\varepsilon,\emptyset}$. The coupling characterization of the Prohorov metric implies

$$\begin{aligned}
&d_{\mathrm{P}}^{\hat{Z}\times\mathbb{R}_+}(m_t^n,m_t^{a,\varepsilon,n})\\
&\le\nu\{(y,y')\in(\hat{Z}\times\mathbb{R}_+)^2:d^{\hat{Z}\times\mathbb{R}_+}(y,y')\ge 2\varepsilon\}+2\varepsilon\\
&\le|C_t^{a,\varepsilon,\emptyset}|_n+2\varepsilon.
\end{aligned}$$

By construction, $|C_t^{a,\varepsilon,\emptyset}|_n \leq |C_0^{a,\varepsilon,\emptyset}|_n$. This follows from the definition of $C_t^{a,\varepsilon,\emptyset}$ in Section 8, from the definition of (r_t, v_t), as a particle at time s on a level i loses the property that $v_s(i) \geq s$ if it reproduces, and as it can only increase its level in a reproduction event. By exchangeability (or Lemma 8.3), we have $\lim_{n\to\infty} |C_0^{a,\varepsilon,\emptyset}|_n = |C_0^{a,\varepsilon,\emptyset}|$ a. s.

The triangle inequality yields

$$\lim_{n,\ell\to\infty} \sup_{t\in[0,T]} d_{\mathrm{P}}^{\hat{Z}\times\mathbb{R}_+}(m_t^n, m_t^\ell) \leq 2|C_0^{a,\varepsilon,\emptyset}| + 4\varepsilon \quad \text{a. s.} \tag{9.1}$$

Letting first $a \to \infty$ and then $\varepsilon \downarrow 0$, we obtain from Lemma 9.1 below that the left-hand side of (9.1) equals zero a. s. By definition of m_t and as $\hat{Z} \times \mathbb{R}_+$ is complete, this implies assertion (i).

As recalled in Section 5.3, the map $t \mapsto (z(t,i), v_t(i))$ is a. s. càdlàg with respect to $d^{\hat{Z}\times\mathbb{R}_+}$, hence the map $t \mapsto m_t^n$ is a. s. càdlàg in the weak topology on $\hat{Z} \times \mathbb{R}_+$. Jump times can only lie in the set Θ_0, which equals a. s. the set of reproduction times. The uniformity of the convergence in assertion (i) implies that also $t \mapsto m_t$ is a. s. càdlàg in the weak topology on $\hat{Z} \times \mathbb{R}_+$ with no jump times outside Θ_0.

W. l. o. g., we assume $\Xi(\Delta) > 0$. As $\Xi \in \mathcal{M}_{\mathrm{dust}}$, this implies $\Xi_0(\Delta) > 0$ and $\Theta_0 = \{t^k : k \in \mathbb{N}\}$ a. s. Now we deduce that $m_{t^k}(\hat{Z} \times \{0\}) > 0$ a. s. all for $k \in \mathbb{N}$, which is part of assertion (iii). By Corollary 6.5, the sequence $(v_{t^k}(i), i \in \mathbb{N})$ is exchangeable. From the definition (2.6) of Θ_0, the definition of $v_t(i)$ in Section 2.1, and as condition (5.1) is a. s. satisfied, it follows that $|\{i \in \mathbb{N} : v_{t^k}(i) = 0\}| > 0$ a. s. Hence, the de Finetti theorem and the definition of m_t yield that a. s., the empirical measures $n^{-1}\sum_{i=1}^n \delta_{v_{t^k}(i)}$ converge weakly to the directing measure $m_{t^k}(\hat{Z} \times \cdot)$ on $\mathbb{R}_+$ which satisfies $m_{t^k}(\hat{Z} \times \{0\}) > 0$.

On the event of probability 1 on which condition (5.1) is satisfied, $v_{t^k-}(i) > 0$ for all $i \in \mathbb{N}$ a. s. Analogously to the above, this yields $m_{t^k-}(\hat{Z} \times \{0\}) = 0$ a. s. This is another part of assertion (iii). As a consequence, the set of jump times is a. s. not smaller than Θ_0, which yields assertion (ii). After completing the proof of assertion (i) by proving Lemma 9.1 below, we give in Proposition 9.2 a representation of the probability measures m_t from which assertion (iv) will follow (see Remark 9.4). A. s., $v_t(i) > 0$ (by condition (5.1)) and $m_t = m_{t-}$ (by item (ii)) for all $t \in (0,\infty) \setminus \Theta_0$ and $i \in \mathbb{N}$. Hence, the remainder of assertion (iii) also follows from Proposition 9.2 below. □

Lemma 9.1. *Let $\varepsilon > 0$, and let $C_0^{a,\varepsilon,\emptyset}$ be defined as in this subsection from a random variable (r_0, v_0) that has the marked distance matrix distribution of a marked metric measure space (X, r, m). Then, $\lim_{a\to\infty} |C_0^{a,\varepsilon,\emptyset}| = 0$ a. s.*

Proof. From the definitions of $C_0^{a,\varepsilon,\emptyset}$ and of the extended lookdown space, we obtain

$$C_0^{a,\varepsilon,\emptyset} = \{j \in \mathbb{N} : (x(j), v(j)) \in (X \times \mathbb{R}_+) \setminus \bigcup_{i=1}^a \mathcal{U}_\varepsilon^{X\times\mathbb{R}_+}(x(i), v(i))\}$$

for all $a \in \mathbb{N}$, where ε-balls are defined by $\mathcal{U}_\varepsilon^{X\times\mathbb{R}_+}(x', v') = \{(x'', v'') \in X \times \mathbb{R}_+ : r(x', x'') \vee |v' - v''| < \varepsilon\}$ for $(x', v') \in X \times \mathbb{R}_+$.

As $(x(a+j), v(a+j))_{j\in\mathbb{N}}$ is an m-iid sequence in $X \times \mathbb{R}_+$ that is independent of $(x(i), v(i))_{i\in[a]}$, the law of large numbers yields

$$|C_0^{a,\varepsilon,\emptyset}| = m((X \times \mathbb{R}_+) \setminus \bigcup_{i=1}^{a} \mathcal{U}_\varepsilon^{X\times\mathbb{R}_+}(x(i), v(i))) \quad \text{a. s.}$$

By separability, $X \times \mathbb{R}_+$ can be covered by countably many balls of diameter $\varepsilon/2$, and each ball with positive mass contains elements of the sequence $(x(i), v(i))_{i\in\mathbb{N}}$ a. s. Using also continuity of m from above, this implies

$$\lim_{a\to\infty} m((X \times \mathbb{R}_+) \setminus \bigcup_{i=1}^{a} \mathcal{U}_\varepsilon^{X\times\mathbb{R}_+}(x(i), v(i))) = 0 \quad \text{a. s.}$$

This yields the assertion. □

Now we give an explicit representation of the probability measures m_t. For $t \in \mathbb{R}_+$, we denote by Π_t the partition of $\mathbb{N}$ in which integers i, j are in the same block if and only if $v_t(i) = v_t(j) < t$ and $r_t(i, j) = 0$, which is condition (i) on p. 96. We may call Π_t the partition of siblings.

The individuals at time t whose parents are also parents of individuals at time 0 are on the levels in the set

$$C_t = \{i \in \mathbb{N} : v_t(i) \geq t\}.$$

All elements of C_t form singleton blocks in Π_t. For $a \in \mathbb{N}, \varepsilon > 0$, the non-singleton blocks of Π_t are also blocks of $\Pi_t^{a,\varepsilon}$. Hence, Corollary 8.7 implies that a. s., C_t equals the union of the singleton blocks of Π_t for each $t \in \mathbb{R}_+$. We also define the map

$$\theta_t : \hat{Z} \times \mathbb{R}_+ \to \hat{Z} \times \mathbb{R}_+, \quad (z, s) \mapsto (z, s+t)$$

with is continuous for $d^{\hat{Z}\times\mathbb{R}_+}$.

Proposition 9.2. *Assume $\Xi \in \mathcal{M}_{\text{dust}}$. Then on an event of probability* 1*,*

$$m_t = \sum_{i\in M(\Pi_t)} |B(\Pi_t, i)|\, \delta_{(z(t,i), v_t(i))} + |C_t|\, \theta_t(m_0) \tag{9.2}$$

for all $t \in \mathbb{R}_+$.

Remark 9.3. Proposition 9.2 allows to describe the probability measure $m_t(\cdot \times \mathbb{R}_+)$ on the extended lookdown space $\hat{Z}$ as follows. With probability given by the asymptotic frequency of the individuals at time t whose ancestral lineages do not coalesce with other ancestral lineages within the time interval $(0, t]$, we sample according to $m_0(\cdot \times \mathbb{R}_+)$. For each block in a reproduction event at a time τ' in $(0, t]$, we draw the individual on, say, the lowest level in this block (which is identified with the individuals on all other levels in this block, as they have genealogical distance zero) with probability given by the asymptotic frequency of the individuals at time t that descend from this block and whose ancestral lineages do not coalesce with any other ancestral lineages in the time interval $(\tau', t]$.

The $\mathbb{R}_+$-component of m_t is determined by the $\hat{Z}$-component if this property holds for m_0. That is, if the marked metric measure space (X, r, m) admits a mark function (see [26, Definition 3.13], [61]), then also $(\hat{Z}, \rho, m_t)$ admits a mark function (cf. also Proposition 3.4 of Chapter 2).

Remark 9.4. The marked metric measure spaces (X, r, m) and $(\hat{Z}, \rho, m_0)$ are isomorphic (as defined in Section 4.2). This follows from the definition of $(\hat{Z}, \rho)$ and the Gromov reconstruction theorem (cf. e.g. Proposition 3.12 in Chapter 2). Hence, m_0 is purely atomic if and only if m is purely atomic, and the assertion on the atomicity of m_t in Theorem 3.9(iv) follows from Proposition 9.2.

Proposition 9.2 and Lemmas 8.3 and 8.5 imply that on an event of probability 1, each weak limit m_{t-} with $t \in (0, \infty)$ is also the sum of countably many atoms and a multiple of $\theta_t(m_0)$. Here we also use that the map $\mathbb{R}_+ \times \hat{Z} \times \mathbb{R}_+ \to \hat{Z} \times \mathbb{R}_+$, $(s, z', v') \mapsto \theta_s(z', v')$ is continuous also in s. This yields the assertion on the atomicity of the left limits m_{t-} in Theorem 3.9(iv).

Proof of Proposition 9.2. For all $a \in \mathbb{N}$, $\varepsilon > 0$, and $t \in \mathbb{R}_+$, the definition of $C_t^{a,\varepsilon,I}$ implies

$$C_t = \bigcup_{I \subset [a]} C_t^{a,\varepsilon,I}.$$

Lemma 8.3 implies the existence of the asymptotic frequencies

$$|C_t| = \sum_{\substack{C \colon C = C_t^{a,\varepsilon,I} \\ \text{for some } I \subset [a]}} |C|$$

for all $t \in \mathbb{R}_+$ a. s. Let us denote the right-hand side of equation (9.2) by m_t'. On an event of probability one, m_t' is a well-defined probability measure for all $t \in \mathbb{R}_+$, and $t \mapsto m_t'$ is càdlàg with respect to $d_{\mathrm{P}}^{\hat{Z} \times \mathbb{R}_+}$. This follows from Lemmas 8.3 and 8.5, and as a. s., C_t equals the union of the singleton blocks of Π_t for each t. As also $t \mapsto m_t$ is a. s. càdlàg, it suffices to show that (9.2) holds a. s. for a fixed $t \in \mathbb{R}_+$.

For $i, j \in \mathbb{N}$ that are in the same block of Π_t, we have $v_t(i) = v_t(j)$ and $z(t, i) = z(t, j)$ in $\hat{Z}$ by definition. Hence,

$$m_t^n = \sum_{i \in M(\Pi_t)} |B(\Pi_t, i)|_n \delta_{(z(t,i), v_t(i))}.$$

Let $A_t^n = \{A_0(t, i) : i \in C_t \cap [n]\}$ be the set of the ancestral levels at time 0 of the individuals on the levels in $C_t \cap [n]$ at time t. For finite sets $A \subset \mathbb{N}$, let

$$m_0^A = \frac{1}{\#A \vee 1} \sum_{i \in A} \delta_{(z(0,i), v_0(i))}.$$

Using the definitions of $z(t,i)$ and $v_t(i)$, we can write

$$m_t^n = \sum_{i \in M(\Pi_t) \setminus C_t} |B(\Pi_t, i)|_n \delta_{(z(t,i), v_t(i))} + |C_t|_n \theta_t(m_0^{A_t^n}). \tag{9.3}$$

Note that A_t^n is the set of those levels at time zero that are occupied by particles that do not reproduce until time t and that are not above level n at time t. This implies that $C_t \cap [n]$ and A_t^n are bijective. In particular, $\bigcup_{n \in \mathbb{N}} A_t^n$ is infinite on the event $\{|C_t| > 0\}$.

Now we work with arguments that we encountered already in Lemmas 3.2 and 3.10. Recall that $(r_0, v_0) = ((r(x(i), x(j)))_{i,j \in \mathbb{N}}, (v(i)_{i \in \mathbb{N}})$ is an m-iid sequence $(x(i), v(i))_{i \in \mathbb{N}}$ in $X \times \mathbb{R}_+$ that is independent of η. For $n \in \mathbb{N}$, we define the empirical measures

$$m^{A_t^n} = \frac{1}{\# A_t^n \vee 1} \sum_{i \in A_t^n} \delta_{(x(i), v(i))}$$

on $X \times \mathbb{R}_+$. As the sets A_t^n and C_t are η-measurable, the Glivenko-Cantelli theorem implies that the weak convergence

$$m = \text{w-}\lim_{n \to \infty} m^{A_t^n}$$

holds a. s. on $\{|C_t| > 0\}$. By construction of the lookdown space $\hat{Z}$, the map $\{(x(i), v(i)) : i \in \mathbb{N}\} \to \hat{Z} \times \mathbb{R}_+$, $(x(i), v(i)) \mapsto (i, v(i))$ can be extended a. s. to an isometry $\hat{\varphi}$ from the support of m in $X \times \mathbb{R}_+$ to $\hat{Z} \times \mathbb{R}_+$ such that $m_0^{A_t^n} = \hat{\varphi}(m^{A_t^n})$, $m_0^n = \varphi(m^{[n]})$, and $\varphi(m) = m_0$. Thus the weak convergence

$$m_0 = \text{w-}\lim_{n \to \infty} m^{A_t^n}$$

holds a. s. on $\{|C_t| > 0\}$.

By exchangeability (or Lemma 8.3), also the relative frequencies in expression (9.3) converge a. s. to the corresponding asymptotic frequencies. Hence, m_t^n converges weakly to m_t' on an event of probability 1. This yields the assertion. □

9.2 The case without dust

In this subsection, we always consider the case $\Xi \in \mathcal{M}_{\mathrm{nd}}$. Let $(x(i), i \in \mathbb{N})$ be an iid sequence in a metric measure space (X, r, μ) that is independent of η. We assume $\rho_0 = (r(x(i), x(j)))_{i,j \in \mathbb{N}}$. With the lookdown space (Z, ρ) associated with η and ρ_0, we are in the setting of Section 3.1.

Proof of Theorem 3.1 (beginning). We work on an event of probability 1 on which in particular the assertions of Lemmas 8.2 and 8.4 hold simultaneously for all $s \in \mathbb{Q}_+$, and we mostly omit 'a. s.' We define for each $t \in (0, \infty)$, $s \in (0,t) \cap \mathbb{Q}$, and $n \in \mathbb{N}$ a probability measure $\mu_t^{(n,s)}$ on Z by

$$\mu_t^{(n,s)} = \sum_{i \in M(\Pi_{s,t})} |B(\Pi_{s,t}, i)|_n \, \delta_{(t,i)}.$$

There exists a coupling ν of the probability measures $\mu_t^{(n,s)}$ and μ_t^n given by

$$\nu\{((t, \min B(\Pi_{s,t}, i)), (t, i))\} = 1/n$$

for all $i \in [n]$. As $\rho((t, \min B(\Pi_{s,t}, i)), (t, i)) \leq 2(t-s)$, the coupling characterization of the Prohorov metric implies

$$d_{\mathrm{P}}^Z(\mu_t^{(n,s)}, \mu_t^n) \leq 2(t-s). \tag{9.4}$$

Let $\varepsilon \in (0, \infty) \cap \mathbb{Q}$, $T \in (\varepsilon, \infty)$, and $\tau \in (0, \varepsilon) \cap \mathbb{Q}$. First we consider $t \in [\tau, T]$. Let $s_0 = 0$ and $s_j = \tau + (j-1)\varepsilon$ for $j \in \mathbb{N}$. By Lemma 8.4, there exists for each $j \in \mathbb{N}_0$ an integer ℓ_j such that

$$\sum_{i \in M(\Pi_{s_j,t}) \cap [\ell_j]} |B(\Pi_{s_j,t}, i)| > 1 - \varepsilon$$

for all $t \in [s_{j+1}, s_{j+2}]$. We set $\ell = \max\{\ell_j : j \in \mathbb{N}_0, s_{j+1} \leq T\}$. By Lemma 8.2, there exists an integer n' such that

$$\sum_{i \in M(\Pi_{s_j,t}) \cap [\ell]} |B(\Pi_{s_j,t}, i)|_n > 1 - \varepsilon \tag{9.5}$$

for all $n \geq n'$, $j \in \mathbb{N}_0$ with $s_{j+1} \leq T$, and $t \in [s_{j+1}, s_{j+2}]$. For all $t \in [\tau, T]$, all $k, n \geq n'$, and $j \in \mathbb{N}_0$ such that $t \in [s_{j+1}, s_{j+2}]$, the bound (9.5) yields

$$d_{\mathrm{P}}^Z(\mu_t^{(n,s_j)}, \mu_t^{(k,s_j)}) \leq \sum_{\substack{i \in M(\Pi_{s_j,t}): \\ i \leq \ell}} \big| |B(\Pi_{s_j,t}, i)|_n - |B(\Pi_{s_j,t}, i)|_k \big| + \varepsilon$$

as the Prohorov distance is bounded from above by the total variation distance. By Lemma 8.2, this expression converges to ε uniformly in $t \in [\tau, T]$ as $n, k \to \infty$. Furthermore, for all $n, k \in \mathbb{N}$, $t \in [\tau, T]$, and j with $t \in [s_{j+1}, s_{j+2}]$, the bound (9.4) yields

$$\begin{aligned} d_{\mathrm{P}}^Z(\mu_t^n, \mu_t^k) &\leq d_{\mathrm{P}}^Z(\mu_t^n, \mu_t^{(n,s_j)}) + d_{\mathrm{P}}^Z(\mu_t^{(n,s_j)}, \mu_t^{(k,s_j)}) + d_{\mathrm{P}}^Z(\mu_t^{(k,s_j)}, \mu_t^k) \\ &\leq 8\varepsilon + d_{\mathrm{P}}^Z(\mu_t^{(n,s_j)}, \mu_t^{(k,s_j)}). \end{aligned}$$

Hence, $\lim_{n\to\infty} \sup_{t \in [\tau, T]} d_{\mathrm{P}}^Z(\mu_t^n, \mu_t) \leq 8\varepsilon$ for all $\tau \in (0, \varepsilon) \cap \mathbb{Q}$ a. s. □

In the next part of the proof, we choose a random τ and show uniform convergence of μ_t^n in $[0, \tau]$. To this aim, we define for each $b \in \mathbb{N}$, $\varepsilon > 0$, and $t \in \mathbb{R}_+$ a collection $(A_t^{b,\varepsilon,I}, I \subset [b])$ of disjoint subsets of $\mathbb{N}$ (whose union is $\mathbb{N}$) by

$$A_t^{b,\varepsilon,I} = \bigcap_{j \in I} \{i \in \mathbb{N} : \rho_t(j, i) < \varepsilon + 2t\} \cap \bigcap_{k \in [b] \setminus I} \{i \in \mathbb{N} : \rho_t(k, i) \geq \varepsilon + 2t\}.$$

This partitions the set of individuals at time t into blocks such that, if t is small, then two individuals are close if they are in a common block $A_t^{b,\varepsilon,I}$ with $I \neq \emptyset$.

Lemma 9.5. *Let $\tilde{T}, \varepsilon > 0$, $b \in \mathbb{N}$, and $I \subset [b]$. Then*

$$\lim_{n\to\infty} \sup_{t\in[0,\tilde{T}]} \big| |A_t^{b,\varepsilon,I}|_n - |A_t^{b,\varepsilon,I}| \big| = 0 \quad a.\,s.$$

Proof. In Lemma 7.1, choose f such that

$$f(t, (\rho_t(j, b{+}1{+}i))_{j\in[b]}) = \prod_{j\in I} \mathbf{1}\{\rho_t(j, b+1+i) < \varepsilon + 2t\} \prod_{k\in[b]\setminus I} \mathbf{1}\{\rho_t(k, b+1+i) \geq \varepsilon + 2t\}$$

for all $t \in \mathbb{R}_+$ and $i \in \mathbb{N}$. The product on the right-hand side is the indicator variable of $\{b + 1 + i \in A_t^{b,\varepsilon,I}\}$. The assertion now follows from Lemma 7.1 as

$$X_n(t) \leq \frac{b+1+n}{n} |A_t^{b,\varepsilon,I}|_{b+1+n} \leq X_n(t) + \frac{b+1}{n}$$

and $X(t) = |A_t^{b,\varepsilon,I}|$ for all $t \in \mathbb{R}_+$ a. s., where $X_n(t)$ and $X(t)$ are defined in Lemma 7.1. □

In a metric space (Y, d), we denote by $\mathcal{U}_\varepsilon^Y(z) = \{y' \in Y : d(y', y) < \varepsilon\}$ for $y \in Y$ some ε-balls.

Lemma 9.6. *Let $b \in \mathbb{N}$ and $\varepsilon > 0$. Then*

$$|A_0^{b,\varepsilon,\emptyset}| = 1 - \mu(\bigcup_{i=1}^{b} \mathcal{U}_\varepsilon^X(x(i))) \quad a.\,s.$$

Proof. By the definitions of $A_0^{b,\varepsilon,\emptyset}$ and Z,

$$A_0^{b,\varepsilon,\emptyset} = \{j \in \mathbb{N} : x(j) \notin \bigcup_{i=1}^{b} \mathcal{U}_\varepsilon^X(x(i))\}.$$

Clearly, $(x(b + i), i \in \mathbb{N})$ is a μ-iid sequence in X that is independent of $(x(i), i \in [b])$. The assertion follows from the law of large numbers. □

Proof of Theorem 3.1 (end). By separability of (X, r) and continuity of μ from above (analogously to the proof of Lemma 9.1), we can (and do) choose a (random) integer b such that

$$\mu\left(\bigcup_{i=1}^{b} \mathcal{U}_\varepsilon^X(x(i))\right) > 1 - \varepsilon.$$

By condition (2.1) and Lemmas 9.6 and 9.5, we can (and do) choose (random) $\tau \in (0, \varepsilon) \cap \mathbb{Q}$ and $n_0 \in \mathbb{N}$ such that $\eta((0, \tau] \times \mathcal{P}^b) = 0$ and $|A_t^{b,\varepsilon,\emptyset}|_n < \varepsilon$ for all $t \in [0, \tau]$ and $n \geq n_0$.

By construction, $\rho((t, i), (t, j)) < 2\varepsilon + 4t$ for all $t \in \mathbb{R}_+$, $I \neq \emptyset$, and $i, j \in A_t^{b,\varepsilon,I}$. Using a coupling ν of μ_t^n and μ_t^k such that

$$\nu((\{t\} \times A_t^{b,\varepsilon,I}) \times (\{t\} \times A_t^{b,\varepsilon,I})) = \mu_t^n(\{t\} \times A_t^{b,\varepsilon,I}) \wedge \mu_t^k(\{t\} \times A_t^{b,\varepsilon,I})$$

for all $I \subset [b]$, we obtain for $n, k \geq n_0$ and $t \in [0, \tau]$ from the coupling characterization of the Prohorov metric that

$$\begin{aligned}
&d_{\mathrm{P}}^{Z}(\mu_t^n, \mu_t^k) \\
&\leq \nu((z, z') \in Z^2 : \rho(z, z') > 2\varepsilon + 4t) + 2\varepsilon + 4t \\
&\leq \nu(Z \times (\{t\} \times A_t^{b,\varepsilon,\emptyset})) + \sum_{\substack{I \subset [b]: \\ I \neq \emptyset}} \nu((Z \setminus (\{t\} \times A_t^{b,\varepsilon,I})) \times (\{t\} \times A_t^{b,\varepsilon,I})) + 2\varepsilon + 4t \\
&\leq \mu_t^k(\{t\} \times A_t^{b,\varepsilon,\emptyset}) + \sum_{\substack{I \subset [b]: \\ I \neq \emptyset}} \left| \mu_t^n(\{t\} \times A_t^{b,\varepsilon,I}) - \mu_t^k(\{t\} \times A_t^{b,\varepsilon,I}) \right| + 2\varepsilon + 4t \\
&\leq 7\varepsilon + \sum_{\substack{I \subset [b]: \\ I \neq \emptyset}} \big| |A_t^{b,\varepsilon,I}|_n - |A_t^{b,\varepsilon,I}|_k \big|.
\end{aligned}$$

Lemma 9.5 implies $\lim_{n,k\to\infty} \sup_{t\in[0,\tau]} d_{\mathrm{P}}^{Z}(\mu_t^n, \mu_t^k) = 7\varepsilon$. Altogether,

$$\lim_{n,k\to\infty} \sup_{t\in[0,T]} d_{\mathrm{P}}^{Z}(\mu_t^n, \mu_t^k) = 8\varepsilon.$$

Assertion (i) follows as ε can be chosen arbitrarily small, and as (Z, ρ) is complete.

Now we come to assertion (ii). By condition (2.1), the map $\mathbb{R}_+ \to Z$, $t \mapsto (t, i)$ is càdlàg for each $i \in \mathbb{N}$. Hence, the map $t \mapsto \mu_t^n$ is càdlàg for each $n \in \mathbb{N}$, and by item (i) also the map $t \mapsto \mu_t$ is càdlàg. For $n \in \mathbb{N}$ and $t \in (0, \infty)$, we define the probability measure

$$\mu_{t-}^n = \frac{1}{n} \sum_{i=1}^{n} \delta_{(t-,i)}$$

on Z, with the left limit $(t-, i) = \lim_{s\uparrow t}(s, i)$. Then $\mu_{t-}^n = \text{w-}\lim_{s\uparrow t} \mu_s^n$. Let $\mu_{t-} = \text{w-}\lim_{s\uparrow t} \mu_s$. From the uniform convergence in item (i), it follows that μ_{t-}^n converges weakly to μ_{t-} as $n \to \infty$. Note that $d_{\mathrm{P}}^{Z}(\mu_{t-}^n, \mu_t^n) \leq 1/n$ for all $t \in (0, \infty) \setminus \Theta_0$ and all $n \in \mathbb{N}$ a. s., as only a binary reproduction event can occur at such at time t. It follows that $\mu_{t-} = \mu_t$ for all $t \in (0, \infty) \setminus \Theta_0$ a. s. That the set of jump times is not smaller than Θ_0 follows from assertion (iii) which we now prove.

W. l. o. g., we assume $\Xi_0(\Delta) > 0$, then we have $\Theta_0 = \{t^k : k \in \mathbb{N}\}$ a. s. (For $\Xi_0(\Delta) = 0$, nothing remains to prove as $\Theta_0 = \emptyset$ a. s. in this case.) Condition (2.1) implies that $\rho_{t^k-}(i, j) > 0$ for all $k, i, j \in \mathbb{N}$ a. s. For each $k \in \mathbb{N}$, the random variable ρ_{t^k-} is exchangeable by Corollary 6.5. On an event of probability 1, let χ be the isomorphy class of the metric measure space (Z, ρ, μ_{t^k-}). Remark 4.4 yields $\chi = \psi(\rho_{t^k-})$. Let ρ' be a random variable whose conditional distribution given χ is the distance matrix distribution of χ. Then by Remark 3.11 of Chapter 2, the random variables ρ' and ρ_{t^k-} are (unconditionally) equal in distribution. Hence, $\rho'(i, j) > 0$ a. s. for all $i, j \in \mathbb{N}$ which implies that μ_{t^k-} is a. s. non-atomic.

By the definition of the population model in Section 2 and the definition (2.6) of Θ_0, there exists for each $k \in \mathbb{N}$ an $i \in \mathbb{N}$ such that $|\{j \in \mathbb{N} : \rho_{t^k}(i, j) = 0\}| = |B(\pi^k, i)| > 0$.

It can now be shown as above that μ_{t^k} contains an atom. More simply, the Portmanteau theorem and item (i) imply

$$\mu_{t^k}((t,i)) \geq \limsup_{n\to\infty} \mu^n_{t^k}((t,i)) = |\{j \in \mathbb{N} : \rho_{t^k}(i,j) = 0\}| > 0.$$

□

Proof of Theorem 3.4. Recall that supp μ_t denotes the support of μ_t. Up to null events,

$$\begin{aligned}
&\{\text{supp } \mu_t \neq X_t \text{ for some } t \in (0,\infty)\} \\
&\subset \bigcup_{i\in\mathbb{N}} \{\mu_t(\mathcal{U}^Z_{2(t-s)}(t,i)) = 0 \text{ for some } t \in (0,\infty), s \in (0,t) \cap \mathbb{Q}\} \\
&= \bigcup_{i\in\mathbb{N}} \{|B(\Pi_{s,t}, i)| = 0 \text{ for some } t \in (0,\infty), s \in (0,t) \cap \mathbb{Q}\}
\end{aligned}$$

and we have a null event in the last line by Lemma 8.8. Together with Lemma 3.2, this shows assertion (i).

For $z \in Z$ and $\varepsilon > 0$, we denote ε-balls in Z by $\mathcal{B}^Z_\varepsilon = \{z' \in Z : \rho(z', z) \leq \varepsilon\}$. For $t \in (0,\infty)$ and $s \in [0,t)$, we denote the closure of $(s,t) \times \mathbb{N}$ in Z by $X_{s,t}$. For $t \in \mathbb{R}_+$ and $n \in \mathbb{N}$, we denote the subset $\{t\} \times [n]$ of Z by X^n_t, and more generally, for $M \subset \mathbb{N}$, we denote the subset $\{t\} \times M$ of Z by X^M_t. We mostly omit 'a. s.' in the following.

In the proof of assertion (ii), we begin with right continuity. Let $t \in \mathbb{R}_+$ and $\varepsilon > 0$. By construction, $X_u \subset \mathcal{B}^Z_{u-t}(X_t)$ for all $u \geq t$. As X_t is compact and as $\{t\} \times \mathbb{N}$ is dense in X_t, there exists $n \in \mathbb{N}$ such that

$$X_t \subset \mathcal{B}^Z_\varepsilon(X^n_t).$$

By condition (2.1), we may choose $\delta \in (0,\varepsilon)$ sufficiently small such that $\eta((t, t+\delta)\times\mathcal{P}^n) = 0$. Then,

$$\mathcal{B}^Z_\varepsilon(X^n_t) \subset \mathcal{B}^Z_{2\varepsilon}(X^n_u) \subset \mathcal{B}^Z_{2\varepsilon}(X_u)$$

for all $u \in [t, t+\delta)$. Thus, we obtain the bound

$$d^Z_{\mathrm{H}}(X_t, X_u) \leq 2\varepsilon$$

for the Hausdorff distance d^Z_{H} over Z. This proves right continuity of the map $t \mapsto X_t$ in d^Z_{H}.

We now turn to the left limits. We write $A_s(t,\mathbb{N}) = \{A_s(t,i) : i \in \mathbb{N}\}$. Let $t \in (0,\infty)$ and $\varepsilon \in (0,t)$. We define the closed subset

$$X_{t-} = \bigcap_{s\in(0,t)} X_{s,t}$$

of Z. By construction, $X_{t-} \supset X_t$. We claim that $d^Z_{\mathrm{H}}(X_s, X_{t-}) \to 0$ as $s \uparrow t$. From the definition of X_{t-}, we have $X_{t-} \subset \mathcal{B}^Z_{t-s}(X_s)$ for all $s \in (0,t)$. Let $M = \bigcap_{s\in[t-\varepsilon,t)} A_{t-\varepsilon}(s,\mathbb{N})$. As $\Xi \in \mathcal{M}_{\mathrm{CDI}}$, it holds $\#M < \infty$. That is, at time $t-$, the number of families of individuals that descend from the same ancestor at time $t-\varepsilon$ is finite. As the map $(t-\varepsilon/2,\infty)\to 2^{\mathbb{N}}$, $s\mapsto A_{t-\varepsilon}(s,\mathbb{N})$ is non-increasing and by condition (5.1) piecewise constant, there exists $\delta\in(0,\varepsilon)$ such that $A_{t-\varepsilon}(s,\mathbb{N})=M$ for all $s\in(t-\delta,t)$. For all $i\in M$ and $s,s'\in(t-\varepsilon,t)$, the definitions of M and $D_s(t-\varepsilon,i)$ yield $D_s(t-\varepsilon,i)<\infty$ and $\rho((s,D_s(t-\varepsilon,i)),(s',D_{s'}(t-\varepsilon,i)))=|s-s'|$, hence the limit $x:=\lim_{s\uparrow t}(s,D_s(t-\varepsilon,i))$ exists in the complete subspace $X_{s',t}$. Also note that $\rho(x,(t-\varepsilon,i))=\varepsilon$ and $x\in X_{t-}$. Thus, $X^M_{t-\varepsilon}\subset\mathcal{B}^Z_\varepsilon(X_{t-})$,

$$X_s\subset\mathcal{B}^Z_\varepsilon(X^M_{t-\varepsilon})\subset\mathcal{B}^Z_{2\varepsilon}(X_{t-}),$$

and $d^Z_{\mathrm{H}}(X_s,X_{t-})\le 2\varepsilon$ for all $s\in(t-\delta,t)$.

Now we show $X_t=X_{t-}$ for $t\in(0,\infty)\setminus\Theta^{\mathrm{ext}}$. Let $x\in X_{t-}$. Then there exists a sequence $((s_k,i_k):k\in\mathbb{N})$ in $(0,t)\times\mathbb{N}\subset Z$ with $0<s_1<s_2<\ldots$ that converges to x. For each $k\in\mathbb{N}$, there exists $\ell\in\mathbb{N}$ such that $\rho((s_n,i_n),(s_\ell,i_\ell))<2(s_\ell-s_k)$ for all $n\ge\ell$. This implies $A_{s_k}(s_n,i_n)=A_{s_k}(s_\ell,i_\ell)$ for all $n\ge\ell$. We set $j_k=A_{s_k}(s_\ell,i_\ell)$. Then $j_k=A_{s_k}(s_n,j_n)$ for all $n\ge\ell\in\mathbb{N}$, hence $D_s(s_k,j_k)<\infty$ for all $s\in[s_k,t)$. By our assumption on t, the definition (2.5) of Θ^{ext}, and condition (5.1), this implies $j'_k:=D_t(s_k,j_k)<\infty$ for all $k\in\mathbb{N}$. The sequence $((s_k,j_k):k\in\mathbb{N})$ converges to x as

$$\rho((s_k,j_k),x)=\lim_{n\to\infty}\rho((s_k,j_k),(s_n,i_n))\le\lim_{n\to\infty}\rho((s_k,i_k),(s_n,i_n))=\rho((s_k,i_k),x)$$

for all $k\in\mathbb{N}$. Also the sequence $((t,j'_k):k\in\mathbb{N})$ converges to x as

$$\rho((t,j'_k),(s_k,j_k))=t-s_k$$

for all $k\in\mathbb{N}$. This implies $x\in X_t$.

Now let $t\in\Theta^{\mathrm{ext}}$. Then by (2.5), there exists $\varepsilon\in(0,t)$ and $i\in\mathbb{N}$ with $D_t(t-\varepsilon,i)=\infty$ and $D_s(t-\varepsilon,i)<\infty$ for all $s\in[t-\varepsilon,t)$. That is, the descendants of some ancestor at time $t-\varepsilon$ die out at time t. The space $(X_t,\rho\wedge t)$ is ultrametric, cf. Remark 5.2 of Chapter 2 and equation (2.3). Hence, a semi-metric $\rho^{(\varepsilon)}$ on X_t is given by $\rho^{(\varepsilon)}=(\rho\wedge t-\varepsilon)\vee 0$. We denote by $X^{(\varepsilon)}_t$ the space obtained from X_t by identifying elements with $\rho^{(\varepsilon)}$-distance 0. As X_t is compact, the space $X^{(\varepsilon)}_t$ is finite. Also the space $(X_{t-},\rho\wedge t)$ is ultrametric. Indeed, for each $x,y,z\in X_{t-}$ and each $s\in(0,t)$, there exist $x',y',z'\in X_s$ with $\rho(x,x')\le t-s$, $\rho(y,y')\le t-s$, and $\rho(z,z')\le t-s$. This implies

$$\begin{aligned}(\rho\wedge t)(x,z)&\le(\rho\wedge s)(x',z')+2(t-s)\\&\le\max\{(\rho\wedge s)(x',y'),(\rho\wedge s)(y',z')\}+2(t-s)\\&\le\max\{(\rho\wedge t)(x,y),(\rho\wedge t)(y,z)\}+4(t-s).\end{aligned}$$

Hence, a semi-metric $\rho^{(\varepsilon)}$ on X_{t-} is given by $\rho^{(\varepsilon)}=(\rho\wedge t-\varepsilon)\vee 0$. We denote by $X^{(\varepsilon)}_{t-}$ the space obtained from X_{t-} by identifying elements with $\rho^{(\varepsilon)}$-distance 0. With i as above, the limit $x:=\lim_{s\uparrow t}(s,D_s(t-\varepsilon,i))$ exists in X_{t-}. As $A_{t-\varepsilon}(t,j)\ne i$ for all $j\in\mathbb{N}$, it follows that $\rho((s,A_s(t,j)),(s,D_s(t-\varepsilon,i)))\ge 2(s-t+\varepsilon)$ for all $s\in(t-\varepsilon,t)$. Taking the limit $s\uparrow t$, we obtain that $\rho((t,j),x)\ge 2\varepsilon$. As $\{t\}\times\mathbb{N}$ is dense in X_t, it also follows that $\rho(x',x)\ge 2\varepsilon$ for all $x'\in X_t$. As $X_t\subset X_{t-}$, this implies that the cardinality of $X^{(\varepsilon)}_{t-}$ is greater than the cardinality of $X^{(\varepsilon)}_t$. Hence, $X^{(\varepsilon)}_{t-}$ and $X^{(\varepsilon)}_t$ are not isometric. It follows that X_{t-} and X_t are not isometric. As a consequence, $d^Z_{\mathrm{H}}(X_{t-},X_t)>0$, that is, the map $t\mapsto X_t$ is discontinuous in Θ^{ext}. □

10 Proof of Lemmas 7.1 and 7.2

We work in the context of Section 7, using also the definitions from Section 5. The proofs in the present section are adaptations of the proofs of Lemmas 3.4 and 3.5 of Donnelly and Kurtz [28] and of Lemma 3.2 of Birkner et al. [13]. We also mention Lemma 6.2 of Labbé [64].

We define stochastic processes $(U(t), t \in \mathbb{R}_+)$ and $(\hat{U}(t), t \in \mathbb{R}_+)$ by

$$U(t) = \int_{(0,t]\times\Delta} |x|_2^2 \; \xi_0(ds\; dx)$$

and

$$\hat{U}(t) = \int_{(0,t]\times\Delta} |x|_1 \; \xi_0(ds\; dx).$$

For $t \in \mathbb{R}_+$, the random variable $U(t)$ equals the sum of the squared asymptotic frequencies of the blocks that encode the reproduction events up to time t. The random variable $\hat{U}(t)$ equals the sum of the asymptotic frequencies of these blocks. By the properties of the Poisson random measure ξ_0,

$$\mathbb{E}[U(t)] = t \int_\Delta |x|_2^2 \, |x|_2^{-2} \, \Xi_0(dx) < \infty,$$

hence the random variable $U(t)$ is a. s. finite. In case $\Xi \in \mathcal{M}_{\mathrm{dust}}$, we also have that

$$\mathbb{E}[\hat{U}(t)] = t \int_\Delta |x|_1 \, |x|_2^{-2} \, \Xi_0(dx) < \infty,$$

and $\hat{U}(t)$ is a. s. finite.

Proof of Lemma 7.1. We assume w. l. o. g. $\Xi(\Delta) > 0$. Let $T, \varepsilon, c > 0$. For $\ell \in \mathbb{N}$, we set $\alpha_0^\ell = 0$ and inductively for $k \in \mathbb{N}_0$

$$\alpha_{k+1}^\ell = \inf\{t > \alpha_k^\ell : U(t) > U(\alpha_k^\ell) + \ell^{-4}\} \wedge (\alpha_k^\ell + \ell^{-4}).$$

We also set $k_\ell = 2\lceil (c+T)\ell^4 \rceil$, then we have that $\mathbb{P}(\alpha_{k_\ell}^\ell \le T, U(T) \le c) = 0$.

For an arbitrary integer $n_\ell \ge \ell$, we define for $k \in \mathbb{N}_0$

$$\beta_k^\ell = \inf\{t > \alpha_k^\ell : \eta((\alpha_k^\ell, t] \times \mathcal{P}^b) > 0\},$$

$$\tilde{\alpha}_k^\ell = \inf\{t > \alpha_k^\ell : \left|X_{n_\ell}(t) - X_{n_\ell}(\alpha_k^\ell)\right| \ge 4\varepsilon\} \wedge (\alpha_k^\ell + 1),$$

and

$$\tilde{\beta}_k^\ell = \inf\{t > \beta_k^\ell : \left|X_{n_\ell}(t) - X_{n_\ell}(\beta_k^\ell)\right| \ge 4\varepsilon\} \wedge (\beta_k^\ell + 1).$$

We define the event

$$E^\ell = \bigcap_{k\in\mathbb{N}_0} \{\eta((\alpha_k^\ell \wedge (T+1), \alpha_{k+1}^\ell \wedge (T+1)) \times \mathcal{P}^b) \le 1\}$$
$$\cap \{|\pi|_2^2 > \ell^{-4} \text{ for all } \pi \in \mathcal{P}^b \text{ with } \eta_0((0, T+1] \times \{\pi\}) > 0\},$$

where $|\pi|_2^2 = \sum_{B\in\pi} |B|^2$. We use the notation $(c, c') = \emptyset$ for $c' \le c$. As $\eta((0, T+1]\times\mathcal{P}^b) < \infty$ a. s., the event E^ℓ occurs for all sufficiently large ℓ a. s.

Recall the strong Markov property of the process J from (3.1). For each $k \in \mathbb{N}_0$, the sequence $(Y_i(\alpha_k^\ell), i \in \mathbb{N})$ is exchangeable by Corollary 6.5 as α_k^ℓ is ξ_0-measurable. The distance matrix

$$\mathbf{1}\big\{\eta((\alpha_k^\ell, \tilde\alpha_k^\ell] \times \mathcal{P}^b) = 0\big\}\, \gamma_{b+1+n_\ell}(\rho_{\tilde\alpha_k^\ell})$$

is $(b+1+n_\ell, b)$-exchangeable by Lemma 6.7 and the strong Markov property of J at α_k^ℓ. Hence, the vector

$$\big(\mathbf{1}\big\{\eta((\alpha_k^\ell, \tilde\alpha_k^\ell] \times \mathcal{P}^b) = 0\big\}\, Y_i(\tilde\alpha_k^\ell), i \in [n_\ell]\big)$$

is exchangeable.

If $b \ge 2$ and $\Xi\{0\} > 0$, then the distance matrix

$$\mathbf{1}\big\{\xi_0(\{\beta_k^\ell\} \times \Delta) = 0\big\}\, \gamma_{b+1+n_\ell}(\rho_{\beta_k^\ell})$$

is $(b+1+n_\ell, b+1)$-exchangeable by Lemma 6.6 and the strong Markov property of J at α_k^ℓ. Hence, the vector

$$(\mathbf{1}\big\{\xi_0(\{\beta_k^\ell\} \times \Delta) = 0\big\}\, Y_i(\beta_k^\ell), i \in n_\ell)$$

is exchangeable. The distance matrix

$$\mathbf{1}\Big\{\xi_0(\{\beta_k^\ell\} \times \Delta) = 0, \eta((\beta_k^\ell, \tilde\beta_k^\ell] \times \mathcal{P}^b) = 0\Big\}\, \gamma_{b+1+n_\ell}(\rho_{\tilde\beta_k^\ell})$$

is $(b+1+n_\ell, b+1)$-exchangeable by Lemma 6.7 and the strong Markov property of J at β_k^ℓ. Hence, the vector

$$\Big(\mathbf{1}\Big\{\xi_0(\{\beta_k^\ell\} \times \Delta) = 0, \eta((\beta_k^\ell, \tilde\beta_k^\ell] \times \mathcal{P}^b) = 0\Big\}\, Y_i(\tilde\beta_k^\ell), i \in [n_\ell]\Big)$$

is exchangeable. If $b < 2$ or $\Xi\{0\} = 0$, then it suffices to work with the stopping times α_k^ℓ and $\tilde\alpha_k^\ell$.

By Lemma A.2 in [28], there exists a number η_ε that depends only on ε (not on n_ℓ) such that

$$\begin{aligned}
&\mathbb{P}(|X_{n_\ell}(\alpha_k^\ell) - X_\ell(\alpha_k^\ell)| \ge \varepsilon) \le 2\,\mathrm{e}^{-\eta_\varepsilon \ell},\\
&\mathbb{P}(|X_{n_\ell}(\tilde\alpha_k^\ell) - X_\ell(\tilde\alpha_k^\ell)| \ge \varepsilon, \eta((\alpha_k^\ell, \tilde\alpha_k^\ell] \times \mathcal{P}^b) = 0) \le 2\,\mathrm{e}^{-\eta_\varepsilon \ell},\\
&\mathbb{P}(|X_{n_\ell}(\beta_k^\ell) - X_\ell(\beta_k^\ell)| \ge \varepsilon, \xi_0(\{\beta_k^\ell\} \times \Delta) = 0) \le 2\,\mathrm{e}^{-\eta_\varepsilon \ell},\\
\text{and}\quad &\mathbb{P}(|X_{n_\ell}(\tilde\beta_k^\ell) - X_\ell(\tilde\beta_k^\ell)| \ge \varepsilon, \xi_0(\{\beta_k^\ell\} \times \Delta) = 0, \eta((\beta_k^\ell, \tilde\beta_k^\ell] \times \mathcal{P}^b) = 0) \le 2\,\mathrm{e}^{-\eta_\varepsilon \ell}.
\end{aligned}$$

Let

$$\begin{aligned}
H_k = &\big|X_{n_\ell}(\alpha_k^\ell) - X_\ell(\alpha_k^\ell)\big| \vee \big|X_{n_\ell}(\tilde\alpha_k^\ell \wedge \beta_k^\ell \wedge \alpha_{k+1}^\ell) - X_\ell(\tilde\alpha_k^\ell \wedge \beta_k^\ell \wedge \alpha_{k+1}^\ell)\big|\\
&\vee \big|X_{n_\ell}(\beta_k^\ell \wedge \alpha_{k+1}^\ell) - X_\ell(\beta_k^\ell \wedge \alpha_{k+1}^\ell)\big| \vee \big|X_{n_\ell}(\tilde\beta_k^\ell \wedge \alpha_{k+1}^\ell) - X_\ell(\tilde\beta_k^\ell \wedge \alpha_{k+1}^\ell)\big|.
\end{aligned}$$

As

$$\{\beta_k^\ell < \alpha_{k+1}^\ell\} \cap \{\alpha_k^\ell \le T\} \cap E^\ell \subset \{\xi_0(\{\beta_k^\ell\} \times \Delta) = 0\},$$
$$\{\tilde\beta_k^\ell < \alpha_{k+1}^\ell\} \cap \{\alpha_k^\ell \le T\} \cap E^\ell \subset \{\eta((\beta_k^\ell, \tilde\beta_k^\ell] \times \mathcal{P}^b) = 0\},$$

and

$$\{\tilde\alpha_k^\ell < \beta_k^\ell\} \subset \{\eta((\alpha_k^\ell, \tilde\alpha_k^\ell] \times \mathcal{P}^b) = 0\}$$

up to null events for all $k \in \mathbb{N}_0$, the above implies

$$\mathbb{P}(\max_{k<k_\ell:\, \alpha_k^\ell \le T} H_k \ge \varepsilon, E^\ell) \le 16\lceil (c+T)\ell^4 \rceil \,\mathrm{e}^{-\eta_\varepsilon \ell}.$$

For $k \in \mathbb{N}_0$, we have that $\tilde\alpha_k^\ell \ge \beta_k^\ell \wedge \alpha_{k+1}^\ell$ a. s. on the event

$$\{H_k < \varepsilon\} \cap \left\{ \sup_{t \in [\alpha_k^\ell, \beta_k^\ell \wedge \alpha_{k+1}^\ell)} \left| X_\ell(t) - X_\ell(\alpha_k^\ell) \right| < \varepsilon \right\}.$$

Indeed, the intersection of this event with $\{\tilde\alpha_k^\ell < \beta_k^\ell \wedge \alpha_{k+1}^\ell\}$ is a null event as it holds on this event that

$$|X_{n_\ell}(\tilde\alpha_k^\ell) - X_{n_\ell}(\alpha_k^\ell)| < 3\varepsilon,$$

whereas we have

$$|X_{n_\ell}(\tilde\alpha_k^\ell) - X_{n_\ell}(\alpha_k^\ell)| \ge 4\varepsilon \quad \text{a. s.}$$

by definition of $\tilde\alpha_k^\ell$ and right continuity.

Similarly, $\tilde\beta_k^\ell \ge \alpha_{k+1}^\ell$ a. s. on the event

$$\{H_k < \varepsilon\} \cap \left\{ \sup_{t \in [\beta_k^\ell, \alpha_{k+1}^\ell)} \left| X_\ell(t) - X_\ell(\beta_k^\ell) \right| < \varepsilon \right\}.$$

By Lemma 10.1 below and the Markov inequality,

$$\sum_{\ell \in \mathbb{N}} k_\ell \mathbb{P}(N^{b+1+\ell}(0, \alpha_1^\ell) > \ell\varepsilon) < \infty.$$

By the strong Markov property of $(N^{b+1+\ell}(0,t], t \in \mathbb{R}_+)$ at α_k^ℓ and by the assumption on f, it follows

$$\sum_{\ell \in \mathbb{N}} \mathbb{P}(\max_{k<k_\ell} \sup_{t \in [\alpha_k^\ell, \beta_k^\ell \wedge \alpha_{k+1}^\ell)} \left| X_\ell(t) - X_\ell(\alpha_k^\ell) \right| > \varepsilon) < \infty$$

and

$$\sum_{\ell \in \mathbb{N}} \mathbb{P}(\max_{k<k_\ell} \sup_{t \in [\beta_k^\ell \wedge (T+1), \alpha_{k+1}^\ell \wedge (T+1))} \left| X_\ell(t) - X_\ell(\beta_k^\ell) \right| > \varepsilon, E^\ell) < \infty.$$

Altogether, it follows that there exist δ_ℓ which do not depend on n_ℓ such that $\sum_{\ell=1}^\infty \delta_\ell < \infty$ and

$$\mathbb{P}(\sup_{t \in [0,T]} |X_{n_\ell}(t) - X_\ell(t)| > 4\varepsilon, U(T) \le c, E^\ell) < \delta_\ell$$

for all $\ell \in \mathbb{N}$.

By Corollary 6.5 and the de Finetti Theorem, there exists an event of probability 1 on which the limits $X(t) = \lim_{n\to\infty} X_n(t)$ exist for all $t \in \mathbb{Q}_+$. Hence,

$$\begin{aligned}
&\mathbb{P}(\sup_{t\in[0,T]\cap\mathbb{Q}} |X(t) - X_\ell(t)| > 4\varepsilon, U(T) \le c, E^\ell) \\
&= \mathbb{P}(\sup_{t\in[0,T]\cap\mathbb{Q}} \liminf_{n\to\infty} |X_n(t) - X_\ell(t)| > 4\varepsilon, U(T) \le c, E^\ell) \\
&\le \lim_{n\to\infty} \mathbb{P}(\inf_{j\ge n} \sup_{t\in[0,T]\cap\mathbb{Q}} |X_j(t) - X_\ell(t)| > 4\varepsilon, U(T) \le c, E^\ell) \le \delta_\ell.
\end{aligned}$$

The Borel-Cantelli lemma and right continuity imply the existence of a random integer L such that
$$|X_n(t) - X_\ell(t)| \le 8\varepsilon$$
for all $t \in [0, T]$ and $n \ge \ell \ge L$ a. s. on the event $\{U(T) \le c\}$. This implies

$$\mathbb{P}(\lim_{n,\ell\to\infty} \sup_{t\in[0,T]} |X_n(t) - X_\ell(t)| = 0, U(T) \le c) = 1$$

The assertion follows by letting c tend to infinity. □

Lemma 10.1. *For $\ell \in \mathbb{N}$, let α_1^ℓ be defined as in the proof of Lemma 7.1. Then there exists a constant C such that $\mathbb{E}[(N^{2\ell}(0, \alpha_1^\ell))^4] \le C\ell^{-2}$ for all $\ell \in \mathbb{N}$.*

The proof extends the argument presented on p. 44 in [13] where additional assumptions on Ξ are required to ensure that the process used there instead of $U(t)$ is finite.

Proof. First, let $x \in \Delta$. For $\ell \in \mathbb{N}$, let $(X_1, X_2, \ldots)$ have infinite multinomial distribution with parameters $(\ell, x_1, x_2, \ldots)$, that is, we may consider iid random variables $U_1, U_2, \ldots$ with uniform distribution on $[0, 1]$ and set

$$X_i = \#\{j \in [\ell] : \sum_{k=1}^{i-1} x_k < U_j < \sum_{k=1}^{i} x_k\}$$

for $i \in \mathbb{N}$. The infinite multinomial distribution appears in the context of Ξ-coalescents e. g. in [71]. For a random partition π in $\mathcal{P}$ with distribution $\kappa(x, \cdot)$, Kingman's correspondence implies that $b_\ell(\pi)$ and $\lim_{n\to\infty} \sum_{i=1}^n (X_i - 1)_+$ are equal in distribution. Here we write $(x - 1)_+ = \max\{x - 1, 0\}$. We use the inequalities $[(x - 1)_+]^2 \le x^{(2)}$ and $[(x - 1)_+]^4 \le 3x^{(4)} + 3x^{(2)}$ for $x \in \mathbb{N}_0$, where $x^{(k)} = x!/(x - k)!$. Inserting also mixed factorial moments of multinomial distributions, we obtain a constant C' such that

$$\mathbb{E}\left[\left(\sum_{i=1}^n (X_i - 1)_+\right)^2\right] \le \sum_{\substack{i,j=1\\ i\ne j}}^n \mathbb{E}\left[X_i^{(2)} X_j^{(2)}\right] + \sum_{i=1}^n \mathbb{E}\left[X_i^{(2)}\right] \le \ell^4 |x|_2^4 + \ell^2 |x|_2^2$$

and

$$\mathbb{E}\left[\left(\sum_{i=1}^{n}(X_i-1)_+\right)^4\right] \le C' \left(\ell^8|x|_2^8 + \ell^6|x|_2^6 + \ell^6|x|_4^4\,|x|_2^2 + \ell^4|x|_2^4 + \ell^4|x|_4^4 + \ell^2|x|_2^2\right)$$

for all $n \in \mathbb{N}$. Taking the limit $n \to \infty$ on the left-hand side, we obtain upper bounds for $\kappa(x,\cdot)b_\ell^2$ and $\kappa(x,\cdot)b_\ell^4$.

Let

$$N_0^\ell(I) = \int_{I\times\mathcal{P}} b_\ell(\pi)\eta_0(ds\;d\pi)$$

for each interval $I \subset \mathbb{R}_+$. The random variable $N_0^\ell(I)$ is the number of newborn particles on the first ℓ levels in the large reproduction events in the time interval I. The random variable $N^{2\ell}(0,\alpha_1^\ell) - N_0^{2\ell}(0,\alpha_1^\ell) = \eta_{\mathrm{K}}((0,\alpha_1^\ell)\times\mathcal{P}^{2\ell})$ is stochastically bounded from above by a Poisson random variable with mean $\ell^{-4}\binom{2\ell}{2}$.

Here we only show $\mathbb{E}[(N_0^{2\ell}(0,\alpha_1^\ell))^4] \le C''\ell^{-2}$ for an appropriate constant C'' and all $\ell \in \mathbb{N}$. W.l.o.g. we assume $\Xi_0(\Delta) > 0$. Recall $((t^i,y^i,\pi^i), i \in \mathbb{N})$ from Section 5. We have

$$\begin{aligned}
&\mathbb{E}[(N_0^{2\ell}(0,\alpha_1^\ell))^4|\xi_0]\\
=&\mathbb{E}[\sum_{\substack{i_1,\dots,i_4\in\mathbb{N}:\\ t^{i_1},\dots,t^{i_4}<\alpha_1^\ell}} b_{2\ell}(\pi^{i_1})\cdots b_{2\ell}(\pi^{i_4})|((t^i,y^i), i\in\mathbb{N})]\\
\le& C'''\big[(\sum_{i\in\mathbb{N}:\,t^i<\alpha_1^\ell}\kappa(y^i,\cdot)b_{2\ell})^4\\
&+\sum_{i\in\mathbb{N}:\,t^i<\alpha_1^\ell}\kappa(y^i,\cdot)b_{2\ell}^2(\sum_{j\in\mathbb{N}:\,t^j<\alpha_1^\ell}\kappa(y^j,\cdot)b_{2\ell})^2\\
&+(\sum_{i\in\mathbb{N}:\,t^i<\alpha_1^\ell}\kappa(y^i,\cdot)b_{2\ell}^2)^2\\
&+\sum_{i\in\mathbb{N}:\,t^i<\alpha_1^\ell}\kappa(y^i,\cdot)b_{2\ell}^3\sum_{j\in\mathbb{N}:\,t^j<\alpha_1^\ell}\kappa(y^j,\cdot)b_{2\ell}\\
&+\sum_{i\in\mathbb{N}:\,t^i<\alpha_1^\ell}\kappa(y^i,\cdot)b_{2\ell}^4\big] \quad \text{a.s.}
\end{aligned}$$

for a combinatorial constant C'''. Now we estimate $b_{2\ell} \le b_{2\ell}^2$ and $b_{2\ell}^3 \le b_{2\ell}^4$, and we insert the bounds for $\kappa(y^i,\cdot)b_{2\ell}^2$ and $\kappa(y^i,\cdot)b_{2\ell}^4$. From the definitions of $(U(t), t\in\mathbb{R}_+)$ and α_1^ℓ, we have

$$\sum_{i\in\mathbb{N}:\,t^i<\alpha_1^\ell}\left|y^i\right|_2^2 \le \ell^{-4}$$

which yields the assertion. □

Proof of Lemma 7.2. The proof is analogous to the proofs of Lemmas 7.1 and 10.1. Let us sketch some differences in our argumentation. We work only with the stopping times

α_k^ℓ and $\tilde{\alpha}_k^\ell$ which we define for $\ell \in \mathbb{N}$ and $k \in \mathbb{N}_0$ by $\alpha_0^\ell = 0$,

$$\tilde{\alpha}_k^\ell = \inf\{t > \alpha_k^\ell : \left|X_{n_\ell}(t) - X_{n_\ell}(\alpha_k^\ell)\right| \geq 4\varepsilon\} \wedge (\alpha_k^\ell + 1),$$

and

$$\alpha_{k+1}^\ell = \inf\{t > \alpha_k^\ell : \hat{U}(t) > \hat{U}(\alpha_k^\ell) + \ell^{-1}\}.$$

We replace the distance matrices $(\rho_t, t \in \mathbb{R}_+)$ by the marked distance matrices $((r_t, v_t), t \in \mathbb{R}_+)$. We set $k_\ell = 2\lceil c\,\ell\rceil$. Then we have that $\mathbb{P}(\alpha_{k_\ell} \leq T, \hat{U}(T) \leq c) = 0$. We set

$$E^\ell = \{|\pi|_2^2 > \ell^{-1} \text{ for all } \pi \in \hat{\mathcal{P}}^b \text{ with } \eta_0((0, T+1] \times \{\pi\}) > 0\}$$

and

$$H_k = \left|X_{n_\ell}(\alpha_k^\ell) - X_\ell(\alpha_k^\ell)\right| \vee \left|X_{n_\ell}(\tilde{\alpha}_k^\ell \wedge \alpha_{k+1}^\ell) - X_\ell(\tilde{\alpha}_k^\ell \wedge \alpha_{k+1}^\ell)\right|$$

for $k \in \mathbb{N}_0$. As

$$\{\tilde{\alpha}_k^\ell < \alpha_{k+1}^\ell\} \cap \{\alpha_k^\ell \leq T\} \cap E^\ell \subset \{\eta((\alpha_k^\ell, \tilde{\alpha}_k^\ell] \times \hat{\mathcal{P}}^b) = 0\}$$

for all $k \in \mathbb{N}_0$, we obtain as in the proof of Lemma 7.1

$$\mathbb{P}(\max_{k < k_\ell :\, \alpha_k^\ell \leq T} H_k \geq \varepsilon, E^\ell) \leq 8\lceil c\ell\rceil\, \mathrm{e}^{-\eta_\varepsilon \ell}.$$

It remains to show

$$\sum_{\ell \in \mathbb{N}} k_\ell \mathbb{P}(\hat{N}^{b+\ell}(0, \alpha_1^\ell) > \ell\varepsilon) < \infty.$$

By the Markov inequality, we have

$$\mathbb{P}(\hat{N}^{b+\ell}(0, \alpha_1^\ell) > \ell\varepsilon) \leq \mathrm{e}^{-\ell\varepsilon}\, \mathbb{E}[\mathbb{E}[\exp(\hat{N}^{b+\ell}(0, \alpha_1^\ell))|\xi_0]].$$

We show that $\sup_{\ell \in \mathbb{N}} \mathbb{E}[\exp(\hat{N}^{2\ell}(0, \alpha_1^\ell))|\xi_0]$ is bounded, this will imply the assertion.

For $x \in \Delta$ and a random partition π with distribution $\kappa(x, \cdot)$, the random variable $\hat{b}_{2\ell}(\pi)$ is binomially distributed with parameters 2ℓ and $|x|_1$. Using monotone convergence, conditional independence, and inserting moment generating functions of binomial distributions, we obtain

$$\begin{aligned}
&\mathbb{E}[\exp(\hat{N}^{2\ell}(0, \alpha_1^\ell))|\xi_0] \\
=&\mathbb{E}[\exp(\sum_{i \in \mathbb{N} :\, t^i < \alpha_1^\ell} \hat{b}_{2\ell}(\pi^i))|((t^i, y^i), i \in \mathbb{N})] \\
=&\lim_{n \to \infty} \prod_{i \in [n] :\, t^i < \alpha_1^\ell} \mathbb{E}[\exp(\hat{b}_{2\ell}(\pi^i))|((t^i, y^i), i \in \mathbb{N})] \\
=&\lim_{n \to \infty} \prod_{i \in [n] :\, t^i < \alpha_1^\ell} (1 - |y^i|_1 + |y^i|_1 \mathrm{e})^{2\ell} \\
=&\exp(\sum_{i \in \mathbb{N} :\, t^i < \alpha_1^\ell} 2\ell \log(1 + |y^i|_1(\mathrm{e} - 1))) \\
\leq&\exp(2\ell \sum_{i \in \mathbb{N} :\, t^i < \alpha_1^\ell} |y^i|_1(\mathrm{e} - 1)) \leq \exp(2(\mathrm{e} - 1)) \quad \text{a.s.}
\end{aligned}$$

The last inequality follows from the definitions of $(\hat{U}(t), t \in \mathbb{R}_+)$ and α_1^ℓ. □

List of notation

Here we collect notation that is used globally in this chapter.

Miscellaneous
$\mathbb{R}_+ = [0, \infty)$, $\mathbb{N} = \{1, 2, 3, \ldots\}$, $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$, $[n] = \{1, \ldots, n\}$ for $n \in \mathbb{N}$, $[0] = \emptyset$
d_{P}^X, d_{H}^X: Prohorov metric and Hausdorff distance over X
$\operatorname{supp} \mu$: support of μ
γ_n: restriction map in various contexts (pp. 64,85)
$\mathcal{U}_\varepsilon^X(x) = \{y \in X : d(x,y) < \varepsilon\}$, $\mathcal{B}_\varepsilon^X(x) = \{y \in X : d(x,y) \leq \varepsilon\}$: balls in a metric space (X, d)

(Marked) distance matrices
$\mathfrak{D}$: space of semi-metrics on $\mathbb{N}$ (p. 78)
$\mathfrak{D}_n$: space of semi-metrics on $[n]$ (p. 85)
$\hat{\mathfrak{D}}$: space of decomposed semi-metrics on $\mathbb{N}$ (p. 83)
$\hat{\mathfrak{D}}_n$: space of decomposed semi-metrics on $[n]$ (p. 85)

Partitions and semi-partitions
$\mathcal{P}$: Set of partitions of $\mathbb{N}$ (p. 64)
$K_{i,j}$: partition of $\mathbb{N}$ that contains only $\{i, j\}$ and singleton blocks (p. 70)
$\mathcal{P}_n$: Set of partitions of $[n]$, associated transformations (p. 87)
$\mathcal{P}^n$: Set of partitions of $\mathbb{N}$ in which the first n integers are not all in different blocks, equation (2.2)
$\hat{\mathcal{P}}^n$: Set of partitions of $\mathbb{N}$ in which the first n integers are not all in singleton blocks, (p. 87)
$B(\pi, i)$: block of the partition $\pi \in \mathcal{P}$ that contains $i \in \mathbb{N}$
$\pi(i) = k$ such that i is in the k-th block π (p. 87)
$M(\pi)$: set of the minimal elements of the blocks of $\pi \in \mathcal{P}$ (p. 86)
$|B|_n = n^{-1}(\#B \cap [n])$, $B = \lim_{n\to\infty} |B|_n$ for $B \subset \mathbb{N}$, $n \in \mathbb{N}$: relative and asymptotic frequency
$\#\pi$: number of blocks of a partition π
$\mathcal{S}_n$ set of semi-partitions of $[n]$, associated transformations (p. 88)
$\Delta = \{x = (x_1, x_2, \ldots) : x_1 \geq x_2 \geq \ldots \geq 0, |x|_1 \leq 1\}$
$\kappa(x, \cdot)$: paintbox distribution associated with Δ (p. 70)

Genealogy in the lookdown model
η: point measure on $(0, \infty) \times \mathcal{P}$ that encodes the reproduction events (pp. 65, 70, 86)
η_0, η_{K}: restrictions of η to multiple and binary reproduction events, respectively (p. 86)
ξ_0 point measure on $(0, \infty) \times \Delta$ that encodes the family sizes in the large reproduction events (p. 86)
(t, i): individual on level i at time t (p. 65)
$A_s(t, i)$: level of the ancestor at time s of (t, i) (p. 65)
(Z, ρ): lookdown space with genealogical distance (p. 65)
$\rho_t(i, j) = \rho((t, i), (t, j))$ (equation (2.3))
$(\hat{Z}, \rho)$: extended lookdown space with genealogical distance (p. 67)
X_t closure of $\{t\} \times \mathbb{N}$ in Z (p. 72)
$X_n(t)$: relative frequencies in Lemmas 7.1 and 7.2
$z(t, i)$: parent of (t, i) (p. 66)

$v_t(i) = \rho((t,i), z(t,i))$ (p. 66)
$r_t(i,j) = \rho(z(t,i), z(t,j))$ (equation (2.4))
$D_t(s,i)$: lowest level at time t of a descendant of (s,i) (p. 68)
$\tau_{s,i}$: extinction time of (s,i) (p. 69)
$\Theta^{\mathrm{ext}} = \{\tau_{s,i} : s \in \mathbb{R}_+, i \in \mathbb{N}\}$ (equation (2.5))
Θ_0: set of large reproduction times (equation (2.6))
μ_t^n: uniform measure on (t,i), $i \in [n]$ (p. 72)
m_t^n: uniform measure on $(z(t,i), v_t(i))$, $i \in [n]$ (p. 75)
μ_t, m_t: weak limits of μ_t^n and m_t, respectively (pp. 72, 75)
$(\Pi_{s,t}, 0 \leq s \leq t)$: flow of partitions (p. 96)
Π_t: partition of siblings: (p. 105)
$C_t^{a,\varepsilon,I}$: subsets of individuals at time t (p. 96)
$\Pi_t^{a,\varepsilon}$: like Π_t but individuals that are in a common $C_t^{a,\varepsilon,I}$ are also in the same block (p. 96)

Characteristic measures
$\Xi = \Xi_0 + \Xi\{0\}\delta_0$ (pp. 70, 86)
H_Ξ: characteristic measure of η (p. 70)
$\mathcal{M}_{\mathrm{dust}}$, $\mathcal{M}_{\mathrm{nd}}$, $\mathcal{M}_{\mathrm{CDI}}$: Sets of finite measures on Δ with and without dust, and with the coming down from infinity property, respectively (pp. 71, 72)

Acknowledgments

The author thanks Götz Kersting and Anton Wakolbinger for their valuable advice. He also thanks them and two anonymous referees for very helpful comments. Partial support from the DFG Priority Programme 1590 “Probabilistic Structures in Evolution” is acknowledged.

Chapter 4

Construction of a tree-valued Fleming-Viot process from a Brownian excursion

We construct a tree-valued Fleming-Viot process from a Brownian excursion and its local time. To this aim, we remove the highest leaf and the root from the associated Aldous continuum random tree and we deform the metric by a factor given by the inverse of the local time at each level. The resulting metric space can be endowed with the image of the normalized local time measure at each level. This yields a process with values in the space of metric measure spaces. To identify this process as a tree-valued Fleming-Viot process, we give a lookdown representation for the binary branching forest that is embedded in reflected Brownian motion.

1 Introduction

The Brownian continuum random tree of Aldous [1] is a random compact metric space $(\mathcal{T}, d)$ that can be encoded by a Brownian excursion and arises as the limit of the genealogical tree of a branching population. It is related to the construction of the (historical) measure-valued Dawson-Watanabe process from a Brownian excursion of Le Gall [66]. The measure-valued (historical) Dawson-Watanabe process can be read off from a spatial tree and its local time measures, we refer to Le Gall [68] for an overview. Each state of the historical Dawson-Watanabe process describes the genealogy at the respective time u which corresponds to the genealogy of samples from the continuum random tree that are drawn according to the local time measure at level u.

We are interested in the tree-valued Fleming-Viot process whose states are (isomorphy classes of) metric measure spaces, i. e. complete and separable metric spaces that are endowed with a probability measure on the Borel sigma algebra. A tree-valued Fleming-Viot process describes the evolution of the genealogical trees in a limiting model for a population whose size is constant over time. This process is introduced by Greven, Pfaffelhuber, and Winter [46] as the solution of a well-posed martingale problem. In Chapter

3, it is constructed pathwise from a Donnelly-Kurtz [28] lookdown approach by means of the random lookdown space, which encodes all individuals ever alive, and a family of sampling measures. The construction of these sampling measures in Chapter 3 is based on techniques from [28]. In the present chapter, we construct the lookdown space from the Brownian continuum random tree. Using the uniform downcrossing representation of local times from Chacon et al. [21], we obtain the sampling measures on the lookdown space as image measures of the normalized local time measures on the Brownian continuum random tree simultaneously for all times on an event of probability 1. We thus construct a tree-valued Fleming-Viot process from a Brownian excursion (Theorem 2.2).

To construct the lookdown space, we remove the root and the highest leaf from the Aldous continuum random tree, and we deform the metric by a factor given by the inverse of the local time at each level. This deformation factor corresponds to the time change that relates the measure-valued Dawson-Watanabe process and the measure-valued Fleming-Viot process in Shiga [87]. To identify the resulting metric space as the lookdown space in two-sided time from Chapter 3, we show in Theorem 3.2 that the levels u of those local minima of the Brownian excursion that separate the i-th and the j-th highest subexcursion above u form a Poisson process after the time change, and that these Poisson processes for $1 \leq i < j$ are independent. For the proof, we use the binary branching forest of Neveu and Pitman [74, 75] and Le Gall [65] that is embedded in reflected Brownian motion, and we represent this forest by another instance of the Donnelly-Kurtz [28] lookdown approach.

The embedding of the one-sided lookdown model into reflected Brownian motion is studied by Berestycki and Berestycki [4], see Proposition 2.1 therein. In Theorem 1 of [4], an embedding of Kingman's coalescent into a Brownian excursion is shown. For the proofs, an alternative approach is used in [4], see Remark 4.2.

The relations between measure-valued Dawson-Watanabe and Fleming-Viot processes are discussed e. g. in Chapter 4 of Etheridge [32]. In particular, Etheridge and March [35] relate these processes by conditioning on constant population size, and the conditional distribution of a Dawson-Watanabe process given the total mass process is characterized by Perkins [76] as an inhomogeneous Fleming-Viot process.

The connection of α-stable continuous-state branching processes with measure-valued Beta-Fleming-Viot processes and Beta-coalescents is studied by Birkner et al. [12]. Berestycki, Berestycki, and Schweinsberg [5] embed the lookdown model into a stable continuum random tree. In the present chapter, however, we restrict ourselves to the Brownian case which is associated with the Kingman coalescent.

We also mention the work of Greven, Popovic, and Winter [47] where the metric of genealogical trees is deformed to describe catalytic branching models. A tree-valued Fleming-Viot processes is further studied by Depperschmidt, Greven, and Pfaffelhuber [26]. Glöde [44] studies tree-valued autocatalytic branching processes and describes their relation to tree-valued Fleming-Viot processes in terms of martingale problems. The local time measures on the more general Lévy trees are studied in Duquesne and Le Gall [30]. In the Brownian case, they are shown to coincide with certain Hausdorff measures in Duquesne and Le Gall [31] and Duhalde [29].

2 Statement of the result

The space of excursions is defined as the space of continuous functions $w : \mathbb{R}_+ \to \mathbb{R}_+$ such that $w(0) = 0$ and $w(s) = 0$ for $s \geq L(w) := \inf\{t > 0 : w(t) = 0\}$. We endow the space of excursions with the metric induced by the supremum norm. Let the random variable $B = (B(s), s \in \mathbb{R}_+)$ with values in the space of excursions be a Brownian excursion conditioned on height $H := \sup B$ larger than 1. The law of this process is $n(\cdot \cap \{w : \sup w > 1\})/n\{w : \sup w > 1\}$, where n denotes the Itô excursion measure of Brownian motion. We also recall that the process B can be obtained a. s. by

$$B = (W((\tilde{g} + s) \wedge \tilde{d}), s \in \mathbb{R}_+), \tag{2.1}$$

where W is standard Brownian motion, and with the hitting time $\varsigma = \inf\{s \in \mathbb{R}_+ : W(s) = 1\}$, the time $\tilde{g} = \sup\{s < \varsigma : W(s) = 0\}$ is defined as the time of the last visit in 0 before ς, and $\tilde{d} = \inf\{s > \varsigma : W(s) = 0\}$ as the time of the first visit in 0 after ς. We denote by $L = L(B)$ the random length of the excursion. For excursion theory and further treatment of Brownian excursions, we refer the reader to Chapter XII of Revuz and Yor [81].

Let $(\zeta(s, u))_{s,u\in\mathbb{R}_+}$ be a jointly continuous version of the local time of B, where $\zeta(s, u)$ denotes the local time accumulated at level u until time s. In the present chapter, we use the uniform downcrossing representation of local time from Chacon et al. [21]: For $\varepsilon > 0$ and $s \in \mathbb{R}_+$, let $D(u, \varepsilon, s)$ be the number of downcrossings of the interval $[u, u + \varepsilon]$ that B completes until time s. By [21, Theorem 2],

$$\lim_{\varepsilon\to 0} \sup_{s,u\in\mathbb{R}_+} |2\varepsilon D(u, \varepsilon, s) - \zeta(u, s)| \quad \text{a. s.} \tag{2.2}$$

The uniform downcrossing representation is stated in [21] for standard Brownian motion and holds as in equation (2.2) for the Brownian excursion by the representation (2.1) and as $L < \infty$.

For $u \in \mathbb{R}_+$, we denote by $\zeta(u) = \zeta(L, u)$ the local time accumulated by B at level u. We always work on the a. s. event that $L < \infty$, the local extrema of B in $[0, L]$ are at pairwise distinct levels, the downcrossing representation (2.2) holds, and $\zeta(u) \in (0, \infty)$ for all $u \in (0, H)$.

From B, we define the semi-metric

$$d(x, y) = B(x) + B(y) - 2 \min_{[x,y]} B \tag{2.3}$$

on $\mathbb{R}_+$ and identify points with distance zero to obtain a real tree $(\mathcal{T}, d)$ which is (a rescaling of) the Aldous continuum random tree. For each $x \in \mathcal{T}$, let $g_x \in [0, L]$ be the leftmost element in the equivalence class $x \in [0, L]/d$, let d_x be the rightmost element, and let $e_x = \min\{s > g_x : B(s) = B(g_x)\}$. The definition of the semi-metric d implies $B(s) \geq B(g_x) = B(d_x)$ for all $s \in (g_x, d_x)$, and as the local minima of B have pairwise distinct levels, there exists at most one $s \in (g_x, d_x)$ with $B(s) = B(g_x)$.

In slight abuse of notation, we refer by $x \in \mathbb{R}_+$ also to the corresponding element of $\mathcal{T}$. The vertex $0 \in \mathcal{T}$ is usually called the root. For $x \in \mathcal{T}$, we define the level $B(x)$ by

$B(x) = B(s)$ where $s \in \mathbb{R}_+$ is any representative of the equivalence class x. Then $B(x)$ is the distance $d(0,x)$ to the root. We say $x \in \mathcal{T}$ is an ancestor of $y \in \mathcal{T}$ if $[g_y, d_y] \subset [g_x, d_x]$. In this case, we write $x \prec y$. By construction, the metric d on $\mathcal{T}$ satisfies

$$d(x,y) = B(x) + B(y) - 2\max\{B(v) : v \in \mathcal{T}, v \prec x, v \prec y\}, \quad x, y \in \mathcal{T}.$$

For further information on continuum random trees, see e. g. Le Gall [67, 68].

For $u \in \mathbb{R}_+$, let the probability measure m_u on $\mathbb{R}_+$ be the Stieltjes measure given by

$$m_u[0,s] = \frac{\zeta(s,u)}{\zeta(u)}, \quad s \in \mathbb{R}_+. \tag{2.4}$$

By the definition (2.3) of d, the canonical map from $\mathbb{R}_+$, endowed with the Euclidean metric, to $(\mathcal{T}, d)$ is continuous. We denote the image measure of m_u under this canonical map also by m_u. The probability measure m_u is usually called the normalized local time measure. The support of m_u lies in the set of vertices $\{x \in \mathcal{T} : B(x) = u\}$ at level u.

We remove the root and the highest leaf from $\mathcal{T}$ to obtain the subtree

$$\mathcal{T}' = \{x \in \mathcal{T} : B(x) \in (0, H)\}.$$

Using the time change given by

$$\tau(u) = \int_1^u \frac{4}{\zeta(v)} dv, \quad u \in (0, H),$$

we set

$$d'(x,y) = \tau(B(x)) + \tau(B(y)) - 2\max\{\tau(B(v)) : v \in \mathcal{T}', v \prec x, v \prec y\}, \quad x, y \in \mathcal{T}'. \tag{2.5}$$

As $\zeta(u) \in (0, \infty)$ for all $u \in (0, H)$, the function $u \mapsto \tau(u)$ is continuous and strictly increasing on $(0, H)$. It is not hard to see that a metric on $\mathcal{T}'$ is given also by d', and that the identity map from $(\mathcal{T}', d)$ to $(\mathcal{T}', d')$ is a homeomorphism. We denote the image measures on $(\mathcal{T}', d')$ of the local time measures m_u, $u \in (0, H)$ on $(\mathcal{T}, d)$ under this homeomorphism again by m_u.

We recall that a metric measure space (X, ρ, μ) is a triple that consists of a complete and separable metric space (X, ρ) and a probability measure μ on the Borel sigma algebra on X. Two metric measure spaces (X, ρ, μ) and (X', ρ', μ') are defined to be isomorphic if there exists a measure-preserving isometry between the supports supp μ and supp μ'. We denote the isomorphy class of (X, ρ, μ) by $[\![X, \rho, \mu]\!]$. We endow the space of isomorphy classes of metric measure spaces with the Gromov-Prohorov metric which is complete and separable, see [45].

In the following, we use the time change

$$U(t) = \inf\{u \in (0, H) : \tau(u) > t\}, \quad t \in \mathbb{R}. \tag{2.6}$$

In Theorem 2.2 below, we identify the process $([\![\mathcal{T}', d', m_{U(t)}]\!], t \in \mathbb{R})$ that we read off from the Brownian excursion as a stationary tree-valued Fleming-Viot process.

Remark 2.1. In the course of the proof of Theorem 2.2, we recall from Section 4.1 of Chapter 3 the construction of a tree-valued Fleming-Viot process from the lookdown space: In Remark 3.5, we point to the construction of the lookdown space $(\bar{Z}, \bar{\rho})$ in two-sided time from a Poisson random measure $\bar{\eta}$ which shows via the convergence (3.3) that the weak limits

$$\bar{\mu}_t = \text{w-}\lim_{n\to\infty} n^{-1} \sum_{i=1}^{n} \delta_{(t,i)}$$

on $\bar{Z}$ exist simultaneously for all $t \in \mathbb{R}$ on an event of probability 1. Analogously to Section 4.1 of Chapter 3, a stationary tree-valued Fleming-Viot process is then given a. s. by $([\![\bar{Z}, \bar{\rho}, \bar{\mu}_t]\!], t \in \mathbb{R})$. By Remark 4.4 of Chapter 3, the process $([\![\bar{Z}, \bar{\rho}, \bar{\mu}_t]\!], t \in \mathbb{R}_+)$ coincides with the tree-valued Fleming-Viot process from Section 8.1 of Chapter 2, hence it solves the martingale problem of Greven, Pfaffelhuber, and Winter [46, Theorem 1].

Theorem 2.2. *The process* $([\![\mathcal{T}', d', m_{U(t)}]\!], t \in \mathbb{R})$ *is a stationary tree-valued Fleming-Viot process.*

The following well-known fact ensures that the process $([\![\mathcal{T}', d', m_{U(t)}]\!], t \in \mathbb{R})$ is a. s. well-defined.

Lemma 2.3. *A. s.,* $\tau(u) \uparrow \infty$ *as* $u \uparrow H$ *and* $\tau(u) \downarrow -\infty$ *as* $u \downarrow 0$. *Hence,* $U(t) \in (0, H)$ *for all* $t \in \mathbb{R}$ *a. s.*

Proof. By the Ray-Knight theorem (see e. g. Theorem XI.2.3 in Revuz and Yor [81] and recall that the concatenation of the subexcursions of B above level 1 forms stopped reflected Brownian motion), the process $(\zeta(1 + u), u \in \mathbb{R}_+)$ is a critical Feller branching diffusion. Hence, Theorem 2.1(i) in [87] or Proposition 1.2 in [12] yield the assertion for $u \uparrow H$. By Williams' description of the Itô excursion measure (see e. g. Theorem XII.4.5 in [81]), the law of $(\zeta(u), u \in [0, H])$ is invariant under time reversal. This implies the assertion for $u \downarrow 0$. □

We will deduce Theorem 2.2 from Theorem 3.2 in the next section by a comparison with the lookdown construction from Chapter 3.

3 The Brownian CRT and the lookdown space

3.1 Ranking the subexcursions

In the lookdown model of Donnelly and Kurtz [28], a rank is assigned to each individual of a population such that an individual at time t with a lower rank has descendants until at least the same time in the future as an individual on a higher rank at time t. In a similar spirit, each subexcursion of B can be assigned a rank according to its height.

Remark 3.1 (Terminology). In particular in [28] and in Chapters 2 and 3, the aforementioned rank is called the level. In the present chapter, we follow [4] and always say rank instead of level as we use the term level already for the distance to the root in the continuum random tree.

A subexcursion interval of B above a level $u \in (0, H)$ is a real interval $(\tilde{g}, \tilde{d})$ such that $B(\tilde{g}) = B(\tilde{d}) = u$ and $B(s) > u$ for all $s \in (\tilde{g}, \tilde{d})$. For $i \in \mathbb{N}$ and $u \in (0, H)$, we denote by $(g_{u,i}, d_{u,i})$ the interval of the i-th highest subexcursion above level u, and by $h_{u,i} = \max_{[g_{u,i}, d_{u,i}]} B - u$ the height of this subexcursion. That is, we choose the subexcursion interval $(g_{u,i},\, d_{u,i})$ such that

$$i = \#\{\tilde{h} \geq h_{u,i} : \tilde{h} = \max_{[\tilde{g},\tilde{d}]} B - u \text{ for a subexcursion interval } (\tilde{g}, \tilde{d}) \text{ of } B \text{ above level } u\}.$$

The integer i is called the rank of the subexcursion. As the local maxima of B are at pairwise distinct levels, the subexcursion intervals $(g_{u,i}, d_{u,i})$ are unique. As the number of downcrossings $D(u, \varepsilon, s)$ is the the number of subexcursions above level u that have height at least ε and that are completed until time s, the downcrossing representation (2.2) and $L < \infty$ yield that for each $u \in (0, H)$, the only accumulation point of $(h_{u,i}, i \in \mathbb{N})$ is 0. In particular, the subexcursion interval $(g_{u,i}, d_{u,i})$ exists for each $i \in \mathbb{N}$.

For $1 \leq i < j$, we say that j looks down on i at level u if $\{g_{u,i}, d_{u,i}\} \cap \{g_{u,j}, d_{u,j}\} \neq \emptyset$, that is, if the i-th and the j-th highest subexcursion above level u are adjacent. In this case, the i-th and the j-th highest subexcursions of B above level u merge to become the i-th highest subexcursion above level $u-$. This corresponds to the embedding of the Donnelly-Kurtz [28] lookdown model from [4, p. 249f], cf. also [5, Section 4.3].

We define a point measure $\xi_{i,j}$ on $\mathbb{R}_+$ that has an atom of mass 1 at each $u \in (0, H)$ such that j looks down on i at level u. Let $\eta_{i,j}$ denote the time-changed point measure

$$\eta_{i,j} = \sum_{t \in \mathbb{R}:\, \xi_{i,j}\{U(t)\} > 0} \delta_t \tag{3.1}$$

on $\mathbb{R}$. We use the following theorem as a link to the lookdown construction from Chapter 3.

Theorem 3.2. *The collection* $(\eta_{i,j} : 1 \leq i < j)$ *consists of independent Poisson random measures whose intensity measure is Lebesgue measure on* $\mathbb{R}$.

In Section 4, we prove Theorem 3.2 and discuss relations to [4] in Remark 4.2. In the remainder of this section, we derive Theorem 2.2. Let us note the following consequence from the definition of the lookdown events.

Lemma 3.3. *For* $1 \leq i < j$*, the point measure* $\xi_{i,j}$ *has no accumulation points in* $(0, H)$.

Proof. This follows from the downcrossing representation (2.2), path continuity of B and $u \mapsto \zeta(u)$, and $L < \infty$. □

3.2 The lookdown space

We define a. s. a semi-metric ρ on $\mathbb{R} \times \mathbb{N}$ by

$$\rho((s, i), (t, j)) = d'(g_{U(s),i}, g_{U(t),j}).$$

Recall that we refer by $g_{u,i}$ to an element of $\mathbb{R}_+$ as well as to the corresponding element of $\mathcal{T}$, and that $U(t) \in (0,H)$ for all $t \in \mathbb{R}$ a. s. by Lemma 2.3.

For each $(t,i) \in \mathbb{R} \times \mathbb{N}$ and $s \leq t$, we define the ancestral rank $A_s(t,i)$ from the collection $(\eta_{i,j}, 1 \leq i < j)$ of point processes by the following properties:

(i) $A_t(t,i) = i$,

(ii) the map $f : \mathbb{R}_+ \to \mathbb{N}$, $r \mapsto A_{t-r}(t,i)$ is left-continuous,

(iii) f jumps from $f(r)$ to $f(r+) = f(r)-1$ at the times $r \in \mathbb{R}_+$ such that $\eta_{j,k}\{t-r\} > 0$ for some $1 \leq j < k < f(r)$,

(iv) f jumps from $k := f(r)$ to some j with $1 \leq j < k$ at the times $r \in \mathbb{R}_+$ such that $\eta_{j,k}\{t-r\} > 0$,

(v) between the jumps described in items (iii) and (iv), f is constant.

Here we use that there exists no pair (j,k) for which the points of $\eta_{j,k}$ accumulate. We also use that for each $r \in \mathbb{R}$, there exists at most one pair (j,k) with $\eta_{j,k}\{r\} > 0$. (These two properties can also be assumed which causes, e. g. by Theorem 3.2, no loss of generality.)

The above is the definition of the ancestral rank in the lookdown model ([28, 77], Chapter 3). The name "ancestral rank" will be justified by the proof of Lemma 3.4 below, where equation (3.2) says that a. s. for $u,v \in (0,H)$, $u < v$, $i \in \mathbb{N}$, the rank of the subexcursion above level u that contains the i-th highest subexcursion above level v is given by $A_{\tau(u)}(\tau(v),i)$. See also [5, p. 1860f], [4, p. 249f] for related considerations.

Lemma 3.4. *A. s.,*

$$\rho((s,i),(t,j)) = s + t - 2\sup\{r \leq s \wedge t : A_r(s,i) = A_r(t,j)\}, \quad (s,i),(t,j) \in \mathbb{R} \times \mathbb{N}.$$

We call $x \in \mathcal{T} \setminus \{0\}$ an edge point if $g_x < e_x = d_x$. In this case, $y \in \mathcal{T}$ is a descendant of x ($x \prec y$) if and only if $[g_y, d_y]$ is contained in the (closed) subexcursion interval $[g_x, e_x]$. A point $x \in \mathcal{T}'$ is called a branchpoint if $g_x < e_x < d_x$, and a leaf if $g_x = d_x$. As branchpoints correspond to local minima of B, they are countable. (We recall from the literature that $\mathcal{T} \setminus \{x\}$ has 3 connected components if $x \in \mathcal{T}$ is a branchpoint, 2 connected components if x is an edge point, and one connected component if x is a leaf.)

Proof. For $j \in \mathbb{N}$ and $u,v \in (0,H)$ with $u \leq v$, let $\tilde{A}_u(v,j)$ be the integer i such that $[g_{v,j}, d_{v,j}] \subset [g_{u,i}, d_{u,i}]$. We claim that $\tilde{A}_u(v,j)$ can be read off from the restrictions of the point measures $\xi_{k,\ell}$, $1 \leq k < \ell$ to $(u,v]$ in the same way as $A_u(v,j)$ is defined from the restrictions of the point measures $\eta_{k,\ell}$, $1 \leq k < \ell$ to $(u,v]$. By the definition (3.1) of $\eta_{k,\ell}$, this claim implies

$$A_s(t,i) = \tilde{A}_{U(s)}(U(t),i) \quad \text{for all } (t,i) \in \mathbb{R} \times \mathbb{N}, s < t \quad \text{a. s.} \tag{3.2}$$

Now we show the claim. First, we note that for $v \in (0,H)$ and $j \in \mathbb{N}$, the map $[0,v) \to \mathbb{N}$, $r \mapsto \tilde{A}_{v-r}(v,j)$ is non-increasing. This follows as for each $u \in (0,v)$, the path of B in $(g_{v,j}, d_{v,j})$ forms part of one of the j highest subexcursions above level u. As

$\tilde{A}_{v-r}(v,j)$ is also $\mathbb{N}$-valued, it follows that for $u \in (0,v)$, the map $r \mapsto \tilde{A}_{v-r}(v,j)$ has at most finitely many jumps in $[0, v-u]$.

Now we consider $r \in [0,v)$ and $1 \leq k < \ell \leq \tilde{A}_{v-r}(v,j)$ such that ℓ looks down on k at level $v-r$. By Lemma 3.3, such times r have no accumulation point in $[0,v)$. If $\ell = \tilde{A}_{v-r}(v,j)$, then $\lim_{r' \downarrow r} \tilde{A}_{v-r'}(v,j) = k$ as such a lookdown event means that the ℓ-th highest subexcursion above level $v-r$ merges with the k-th highest subexcursion above level $v-r$ and takes over the lower rank. If $\ell < \tilde{A}_{v-r}(v,j)$, then $\lim_{r' \downarrow r} \tilde{A}_{v-r'}(v,j) = \tilde{A}_{v-r}(v,j) - 1$, the $\tilde{A}_{v-r}(v,j)$-th highest subexcursion above level $v-r$ then decreases its rank by 1 as subexcursions of lower ranks merge.

Conversely, for $r \in [0,v)$ such that $\xi_{i,j}\{v-r\} = 0$ for all $1 \leq k < \ell \leq \tilde{A}_{v-r}(v,j)$, none of the $\tilde{A}_{v-r}(v,j)$ highest subexcursions merge at level $v-r$ whence the map $r' \mapsto \tilde{A}_{v-r'}(v,j)$ has no jump at r. This shows the claim.

For each $(u,i),(v,j) \in (0,H) \times \mathbb{N}$ such that there exists $x \in \mathcal{T}'$ such that x is not a branchpoint and satisfies $x \prec g_{u,i}$ and $x \prec g_{v,j}$, it holds $\tilde{A}_{B(x)}(u,i) = \tilde{A}_{B(x)}(v,j)$ and $x = g_{B(x),k}$ where $k = \tilde{A}_{B(x)}(u,i)$. Conversely, for each $(u,i),(v,j) \in (0,H) \times \mathbb{N}$ and $p \in (0, u \wedge v]$ such that $\tilde{A}_p(u,i) = \tilde{A}_p(v,j)$, it holds $x \prec g_{u,i}$ and $x \prec g_{v,j}$ for $x = g_{p,\tilde{A}_p(u,i)} \in \mathcal{T}'$.

Hence a. s., for each $(s,i),(t,j) \in \mathbb{R} \times \mathbb{N}$,

$$\begin{aligned}
&s + t - 2\sup\{r \leq s \wedge t : A_r(s,i) = A_r(t,j)\} \\
&= s + t - 2\sup\{r \leq s \wedge t : A_r(s,i) = A_r(t,j), g_{U(r),A_r(s,i)} \text{ is not a branchpoint}\} \\
&= s + t - 2\sup\{\tau(B(x)) : x \in \mathcal{T}' \text{ is not a branchpoint}, x \prec g_{U(s),i}, x \prec g_{U(t),j}\} \\
&= d'(g_{U(s),i}, g_{U(t),j}).
\end{aligned}$$

For the first and the last equality, we use that the branchpoints are countable, and that $\tau \circ U$ is a. s. the identity on $\mathbb{R}$. The second equality follows from the above discussion by setting $r = \tau(B(x))$, using (3.2) and that $U \circ \tau$ is a. s. the identity on $(0,H)$. □

We denote by (Z,ρ) the metric completion of the space obtained from $(\mathbb{R} \times \mathbb{N}, \rho)$ after identifying the elements with distance zero. In slight abuse of notation, we refer by each element of $\mathbb{R} \times \mathbb{N}$ also to the corresponding element of Z.

Remark 3.5 (The lookdown space). Let $\mathcal{P}$ denote the space of partitions of $\mathbb{N}$, and for $1 \leq i < j$, let $K_{i,j} \in \mathcal{P}$ be the partition that consists of the block $\{i,j\}$ and apart from that only of singleton blocks. Furthermore, let $\bar{\eta}$ be a point measure on $\mathbb{R} \times \mathcal{P}$ as in Section 4.1 of Chapter 3, and let $\bar{\eta}_{i,j} = \bar{\eta}(\cdot \times \{K_{i,j}\})$

By Lemma 3.4, the space (Z,ρ) is constructed a. s. in the same deterministic way from the collection of point measures $(\eta_{i,j}, 1 \leq i < j)$ as the two-sided lookdown space $(\bar{Z},\bar{\rho})$ in Sections 2 and 4.1 of Chapter 3 is constructed from $(\bar{\eta}_{i,j}, 1 \leq i < j)$. In Section 4.1 of Chapter 3, $(\bar{\eta}_{i,j}, 1 \leq i < j)$ consists by assumption of independent Poisson random measures whose intensity measure is Lebesgue measure on $\mathbb{R}$.

3.3 The sampling measures

For each $u \in (0, H)$ and $n \in \mathbb{N}$, we define a probability measure m_u^n on $(\mathcal{T}, d)$ by

$$m_u^n = \frac{1}{n} \sum_{i=1}^{n} \delta_{g_{u,i}}.$$

This probability measure charges the root of each of the n highest subtrees above level u in $\mathcal{T}$ with mass $1/n$, where we count both subtrees above a branchpoint as separate trees.

Lemma 3.6. *The weak limits*

$$m_u = w\text{-}\lim_{n\to\infty} m_u^n$$

in the space of probability measures on $(\mathcal{T}, d)$ exist for all $u \in (0, H)$.

To prove Lemma 3.6, we use the uniform downcrossing representation of local times.

Proof of Lemma 3.6. For $\varepsilon > 0$ and $u \in \mathbb{R}_+$, we denote by $\tilde{m}_u^\varepsilon$ the Stieltjes measure on $\mathbb{R}_+$ that is given by

$$\tilde{m}_u^\varepsilon[0, s] = \frac{D(u, \varepsilon, s)}{D(u, \varepsilon, L)}, \quad s \in \mathbb{R}_+.$$

By the downcrossing representation (2.2) and the definition (2.4) of m_u,

$$\lim_{\varepsilon\to 0} \tilde{m}_u^\varepsilon[0, s] = m_u[0, s]$$

for all $s, u \in \mathbb{R}_+$. Hence, the probability measures $\tilde{m}_u^\varepsilon$ converge weakly to the probability measures m_u on $\mathbb{R}_+$. As the canonical map φ from $\mathbb{R}_+$, endowed with the Euclidean metric, to $(\mathcal{T}, d)$ is continuous, this weak convergence also holds for the image measures, that is, $m_u = \text{w-}\lim_{\varepsilon\to 0} \varphi(\tilde{m}_u^\varepsilon)$ for all $u \in \mathbb{R}_+$.

We recall that for $u \in (0, H)$ and $n \in \mathbb{N}$, the n-th largest subexcursion height above level u is denoted by $h_{u,n}$. Then,

$$m_u^n = \varphi(\tilde{m}_u^{h_{u,n}})$$

and the assertion follows as $\lim_{n\to\infty} h_{u,n} = 0$. □

As the identity from $(\mathcal{T}', d)$ to $(\mathcal{T}', d')$ is continuous, the assertion of Lemma 3.6 also holds when m_u^n and m_u are interpreted as probability measures on $(\mathcal{T}', d')$ which we do from now on.

By definition of ρ, an isometry θ from (Z, ρ) to $(\mathcal{T}', d')$ is given a. s. by $\theta(t, i) = g_{U(t),i}$ for $(t, i) \in \mathbb{R}\times\mathbb{N} \subset Z$. The set $\{x \in \mathcal{T} : \max_{[g_x, d_x]} B > B(x)\}$ is usually called the skeleton and is dense in $(\mathcal{T}', d')$. As in Subsection 3.1, the subexcursion heights above each level accumulate only in 0 which implies $\theta(\mathbb{R} \times \mathbb{N}) = \mathcal{T}^0$ a. s. This yields a. s. surjectivity of $\theta : Z \to \mathcal{T}'$, whence θ has a. s. an inverse $\bar{\theta} : \mathcal{T}' \to Z$.

For $t \in \mathbb{R}$ and $n \in \mathbb{N}$, we define the probability measure

$$\mu_t^n = \frac{1}{n} \sum_{i=1}^{n} \delta_{(t,i)}$$

on (Z, ρ). As $\bar{\theta} \circ \theta$ is a. s. the identity on Z,

$$\mu_t^n = \bar{\theta}(\frac{1}{n} \sum_{i=1}^{n} \delta_{\theta(t,i)}) = \bar{\theta}(m_{U(t)}^n)$$

for all $t \in \mathbb{R}$ and $n \in \mathbb{N}$ a. s., hence the weak limits

$$\mu_t := \text{w-}\lim_{n\to\infty} \mu_t^n = \lim_{n\to\infty} \bar{\theta}(m_{U(t)}^n) = \bar{\theta}(m_{U(t)}) \tag{3.3}$$

on (Z, ρ) exist for all $t \in \mathbb{R}$ a. s. by Lemma 3.6 and continuity of $\bar{\theta}$.

Proof of Theorem 2.2. By Theorem 3.2, Remark 3.5, the construction of the measures m_t, and Remark 2.1, a stationary tree-valued Fleming-Viot process is given by $([\![Z, \rho, \mu_t]\!], t \in \mathbb{R})$ a. s. A. s. by equation (3.3), θ is a measure-preserving isometry from the metric measure space (Z, ρ, μ_t) to $(\mathcal{T}', d', m_{U(t)})$ for each $t \in \mathbb{R}$. □

4 The Poisson processes of the lookdown events

This section is devoted to the proof of Theorem 3.2. Let $u_0 \in (0, 1)$. We assume that there exists a probability measure $\mathbb{P}^{(u_0)}$ on our underlying probability space under which B is a Brownian excursion conditioned on height larger than u_0. Then under $\mathbb{P} := \mathbb{P}^{(u_0)}(\cdot|H > 1)$, the process B is a Brownian excursion conditioned on height larger than 1 as in Section 2. We denote by $\mathbb{E}^{(u_0)}$ the expectation associated with $\mathbb{P}^{(u_0)}$. From now on, we mean $\mathbb{P}^{(u_0)}$-a. s. when we just write a. s.

Analogously to Section 2, we define a time change by

$$\sigma(u) = \int_{u_0}^{u} \frac{4}{\zeta(v)} dv, \quad u \in (0, H) \tag{4.1}$$

with inverse

$$V(t) = \inf\{u \in (0, H) : \sigma(u) > t\}, \quad t \in \mathbb{R}. \tag{4.2}$$

Analogously to Lemma 2.3, $\sigma(u) \downarrow -\infty$ for $u \downarrow 0$ and $\sigma(u) \uparrow \infty$ for $u \uparrow H$ a. s. Also note that $V(0) = u_0$.

For $1 \leq i < j$ and $t \in \mathbb{R}_+$, we set

$$\eta_{i,j}(t) = \xi_{i,j}(u_0, V(t)]. \tag{4.3}$$

Then the process $(\eta_{i,j}(t), t \in \mathbb{R}_+)$ counts the lookdowns from j to i above level u_0 under the time change.

In this section, we prove the following lemma.

Lemma 4.1. *Under $\mathbb{P}$, the collection of processes $((\eta_{i,j}(t), t \in \mathbb{R}_+), 1 \le i < j)$ consists of independent Poisson processes with rate* 1*, and this collection is independent of $\tau(u_0)$.*

We deduce Theorem 3.2 by considering arbitrarily small u_0 so as to describe η also at negative times.

Proof of Theorem 3.2. Let $T \in \mathbb{R}_+$. For $1 \le i < j$ and $t \in \mathbb{R}_+$, let

$$\tilde{\eta}_{i,j}(t) = \eta_{i,j}(-T, -T + t].$$

It suffices to show that the processes $(\tilde{\eta}_{i,j}(t), t \in \mathbb{R}_+)$ are independent Poisson processes with rate 1 under $\mathbb{P}$.

By the definition (3.1) of $\eta_{i,j}$,

$$\tilde{\eta}_{i,j}(t) = \xi_{i,j}(U(-T), U(-T + t)], \quad t \in \mathbb{R}_+. \tag{4.4}$$

From the definition of $\sigma(u)$ and $\tau(u)$, we obtain $\mathbb{P}$-a. s. the translation formula

$$\tau(u) = \sigma(u) + \tau(u_0), \quad u \in (0, H).$$

Hence, $\mathbb{P}$-a. s. by the definitions (2.6) of $U(t)$ and (4.2) of $V(t)$,

$$U(t) = \inf\{u \in (0, H) : \sigma(u) + \tau(u_0) > t\} = V(t - \tau(u_0)), \quad t \in \mathbb{R}.$$

Inserting this into equation (4.4) and using the definition (4.3) of $\eta_{i,j}(t)$ yields $\mathbb{P}$-a. s.

$$\tilde{\eta}_{i,j}(t) = \eta_{i,j}(-T - \tau(u_0) + t) - \eta_{i,j}(-T - \tau(u_0)), \quad t \in \mathbb{R}_+$$

on the event $\{-T > \tau(u_0)\}$.

By Lemma 4.1, the processes $(\tilde{\eta}_{i,j}(t), t \in \mathbb{R}_+)$ are independent Poisson processes with rate 1 under $\mathbb{P}(\cdot | -T > \tau(u_0))$. As u_0 was arbitrary and as $\lim_{u_0 \downarrow 0} \mathbb{P}(-T > \tau(u_0)) = 1$ by Lemma 2.3, it follows that the processes $(\tilde{\eta}_{i,j}(t), t \in \mathbb{R}_+)$ satisfy the assertion. □

Remark 4.2. In the proof of Proposition 2.1 of [4], it is shown that j looks down on i with rate $4/Z_u$ conditionally given the path of a stopped reflected Brownian motion below level u, where Z_u denotes the local time accumulated in level u. This rate does not change when one conditions also on the lookdown events below level u. Indeed, the pattern of the Brownian path that constitutes a lookdown event from i to $j > i$ depends only on the i-th highest subexcursion above level u and the point process of the heights of all subexcursions above level u. Moreover, the local time profile above u does not depend on the order of the subexcursions above level u. Hence, using the time change, it can be deduced that $((\eta_{i,j}(t), t \in \mathbb{R}_+), 1 \le i < j)$ is a collection of independent Poisson processes with rate 1, cf. equation (3.6) in [4].

In the present work, we use the independence of the shift $\tau(u_0)$ to embed the lookdown model also in two-sided time. The extension to two-sided time could also be achieved by a modification of Theorem 1.1 of [4] where Kingman's coalescent is embedded into a Brownian excursion.

The general approach in [4] is to analyze the patters of the Brownian path that constitute lookdown events. In the present work, we use the binary branching forest that is embedded in stopped reflecting Brownian motion, and we give a lookdown representation for it. This approach also makes clear that the rate of lookdown events from j to i does not depend on i and j. In [4], this independence of i, j follows from a calculation which is made plausible on p. 249 therein: High subexcursions have small Itô measure, but the height of a low excursion typically has to fit into a smaller interval to result in a lookdown from j to i.

4.1 Pruning

In the remainder of this chapter, $u_0 \in (0,1)$ remains fixed. Let $h > 0$.

For $u \in \mathbb{R}_+$, let

$$D^h(u) = \max\{i \in \mathbb{N} : h_{u_0+u,i} \geq h\}$$

be the number of subexcursions above level $u_0 + u$ of height at least h. Then $D^h(u)$ is equal to the number $D(u_0 + u, h, L)$ of downcrossings of the interval $[u_0 + u, u_0 + u + h]$ by the path B. Hence, the uniform downcrossing representation (2.2), which also holds under $\mathbb{P}^{(u_0)}$, now reads

$$\lim_{h \downarrow 0} \sup_{u \in \mathbb{R}_+} |2hD^h(u) - \zeta(u_0 + u)| = 0. \tag{4.5}$$

For $1 \leq i < j$, let $\xi^h_{i,j}(u)$ be the number of lookdown events from j to i at levels in $(u_0, u_0 + u]$ in which both merging subexcursions have height larger than h, that is,

$$\xi^h_{i,j}(u) = \#\{v \in (u_0, u_0 + u] : \{g_{v,i}, d_{v,i}\} \cap \{g_{v,j}, d_{v,j}\} \neq \emptyset, h_{v,i} \wedge h_{v,j} > h\}. \tag{4.6}$$

Moreover, let

$$\xi_{i,j}(u) = \xi_{i,j}(u_0, u_0 + u], \quad u \in \mathbb{R}_+. \tag{4.7}$$

Then, by definition of the point measure $\xi_{i,j}$ in Section 3.1,

$$\xi_{i,j}(u) = \#\{v \in (u_0, u_0 + u] : \{g_{v,i}, d_{v,i}\} \cap \{g_{v,j}, d_{v,j}\} \neq \emptyset\}. \tag{4.8}$$

Corollary 4.3. *For each $u \in [0, H - u_0)$, we have $\xi^h_{i,j}(u) = \xi_{i,j}(u)$ for all sufficiently small $h > 0$.*

Proof. This follows from Lemma 3.3 and equations (4.6), (4.7), and (4.8). □

Remark 4.4. The quantities $D^h(u)$ and $\xi^h_{i,j}(u)$ can be naturally interpreted in terms of the subtree

$$\mathcal{T}^h = \{x \in \mathcal{T} : \max_{[g_x, d_x]} B - B(x) \geq h\}$$

that is obtained by pruning all elements of $\mathcal{T}$ that are closer than h to a leaf. Then $D^h(u)$ is the number of vertices of $\mathcal{T}^h$ at level $u_0 + u$, and $\xi^h_{i,j}(u)$ is the number of branchpoints within $\mathcal{T}^h$ at levels in $(u_0, u_0 + u]$ in which the heights of the subexcursions that correspond to the subtrees above that branchpoint have ranks i and j.

In Subsections 4.2 – 4.4, we consider the forest given by the vertices of $\mathcal{T}^h$ above level u_0 for a fixed $h > 0$. In Subsection 4.5, we perform time changes and let $h \downarrow 0$ to deduce Lemma 4.1.

4.2 Basic notions

We state some basic notions for most of which we refer to e. g. [65, 73, 80]. We restrict ourselves to finite trees and forests. A rooted ordered binary tree is a finite subset $\mathcal{U}$ of the set $\bigcup_{n\in\mathbb{N}_0}\{1,2\}^n$ of finite sequences $u = u_1 \dots u_n$ such that $\mathcal{U}$ satisfies the following properties:

(i) $\emptyset \in \mathcal{U}$

(ii) $u2 \in \mathcal{U}$ implies $u1 \in \mathcal{U}$

(iii) $u1 \in \mathcal{U}$ implies $u \in \mathcal{U}$.

Here the conventions $\mathbb{N}^0 = \{\emptyset\}$ and $\emptyset u = u$ are used. The vertex $\emptyset$ is called the root. A vertex $u1 \in \mathcal{U}$ is called a left child of u, and a vertex $u2 \in \mathcal{U}$ a right child of u. The length $|u|$ of a sequence $u \in \mathcal{U}$ is called the generation of the vertex u. If $ui \in \mathcal{U}$ for $i = 1$ or 2, then u is called the parent of ui. We say $u \in \mathcal{U}$ is an ancestor of $v \in \mathcal{U}$ if the sequence u is a prefix of v (possibly $u = v$). In this case, we write $u \prec v$. We denote by $\mathbb{T}$ the set of rooted ordered binary trees.

For $n \in \mathbb{N}$, we write $[n] = \{i \in \mathbb{N} : i \leq n\}$. We define a forest as a set of the form $\{(u,m) : u \in \mathcal{U}_m, m \in [M]\}$, where $M \in \mathbb{N}_0$ and $\mathcal{U}_1, \dots, \mathcal{U}_M \in \mathbb{T}$. We denote the set of forests by $\mathbb{F}$ and call the elements (u,m) of $\mathcal{F} \in \mathbb{F}$ vertices. For $\mathcal{F} \in \mathbb{F}$, let $\mathcal{H}(\mathcal{F}) = \max\{|u| : (u,m) \in \mathcal{F}\}$ be the number of generations in the forest. For $n \in \mathbb{N}_0$, let $N_n(\mathcal{F}) = \#\{(u,m) \in \mathcal{F} : |u| = n\}$ be the number of vertices of $\mathcal{F}$ in generation n. For $(u,m) \in \mathcal{F}$, let $\mathcal{H}(u,m) = \max\{|v| - |u| : (v,m) \in \mathcal{F}, u \prec v\}$ be height of the subtree above (u,m).

For $\mathcal{F} \in \mathbb{F}$ and $(u,m) \in \mathcal{F}$, we define the rank of (u,m) according to $\mathcal{H}(u,m)$ among the vertices in the same generation by

$$R(u,m) = \#\{(v,p) \in \mathcal{F} : \mathcal{H}(v,p) \geq \mathcal{H}(u,m), |v| = |u|\}.$$

We define a forest with generation levels as a pair $(\mathcal{F}, g)$ where $\mathcal{F} \in \mathbb{F}$, $g \in \mathbb{R}_+^{\mathcal{H}(\mathcal{F})}$ with $g(1) \leq \dots \leq g(\mathcal{H}(\mathcal{F}))$. For $n \in [\mathcal{H}(\mathcal{F})]$, we interpret $g(n) - g(n-1)$ as the length of each edge between vertices of generations n and $n-1$. We always set $g(0) = 0$ and $g(n) = \infty$ for $n > \mathcal{H}(\mathcal{F})$. We denote the set of forests with generation levels by $\mathbb{G}$.

For $n \in \mathbb{N}_0$, we define the restriction to the first n generations

$$\gamma_n : \mathbb{F} \to \mathbb{F}, \quad \mathcal{F} \mapsto \{(u,m) \in \mathcal{F} : |u| \leq n\}.$$

We define a binary branching forest with rate $\lambda > 0$ as a random forest with generation levels $(\mathcal{F}, g)$ that satisfies the following properties: $(\gamma_n(\mathcal{F}), n \in \mathbb{N}_0)$ is a Markov chain. Conditionally given $\gamma_n(\mathcal{F})$ for some $n \in \mathbb{N}_0$, the next generation is drawn in a two-step random experiment. In the first step, we decide with equal probability whether a birth or a death event occurs. In a death event, a uniformly drawn vertex from generation n has no children, each other vertex in generation n has one child. If a birth event occurs, a uniformly drawn vertex from generation n has two children, each other vertex in generation n has one child. Conditionally given $\mathcal{F}$, the random variables $g(n) - g(n-1)$, $n \in [\mathcal{H}(\mathcal{F})]$ are independent and exponentially distributed with parameter $\lambda N_{n-1}(\mathcal{F})$.

4.3 Binary branching forest in reflected Brownian motion

A central object for our proof of Lemma 4.1 is the binary branching forest that is embedded in reflected Brownian motion and for which we refer to Neveu and Pitman [74, 75], Le Gall [65], and Chapter 7.6 in Pitman [80].

Under $\mathbb{P}^{(u_0)}$, conditionally given $\zeta(u_0)$, the concatenation of the subexcursions of B above level u_0 forms a stopped reflected Brownian motion. More precisely, we use the time change $\alpha(s) = \inf\{t \in \mathbb{R}_+ : \int_0^t \mathbf{1}\{B(r) > u_0\}\,dr > s\}$, and for $s \in \mathbb{R}_+$, we set $\beta(s) = B(\alpha(s)) - u_0$ if $s < \int_0^\infty \mathbf{1}\{B(r) > u_0\}\,dr$ and $\beta(s) = 0$ else. Then by excursion theory, the process $(\beta(s), s \in \mathbb{R}_+)$ is a reflected Brownian motion stopped at local time $\zeta(u_0)$.

Let

$$\mathcal{L} = \{x \in \mathcal{T} : \max_{[g_x,d_x]} B - B(x) = h, B(x) \geq u_0\}.$$

The vertices in $\mathcal{L}$ are the leaves within $\mathcal{T}^h$ above level u_0. Let

$$\mathcal{B} = \{x \in \mathcal{T} : \max_{[g_x,e_x]} B \wedge \max_{[e_x,d_x]} B > B(x) + h, B(x) \geq u_0\}.$$

These vertices are the branchpoints within $\mathcal{T}^h$ above level u_0. Moreover, let

$$\mathcal{A}_0 = \{x \in \mathcal{T} : \max_{[g_x,d_x]} B - u_0 \geq h, B(x) = u_0\}.$$

These vertices are the vertices of $\mathcal{T}^h$ at level u_0. We define the set of vertices

$$\mathcal{A} = \{x \in \mathcal{T} : \max_{[g_x,d_x]} B - B(x) \geq h, \text{ there exists } y \in \mathcal{L} \cup \mathcal{B} \text{ with } B(x) = B(y)\} \cup \mathcal{A}_0$$

which is the smallest subset $\mathcal{A}'$ of $\mathcal{T}^h$ that contains all vertices at level u_0 such that for each vertex in $\mathcal{A}'$ above level u_0, all vertices of $\mathcal{T}^h$ at that level are contained in $\mathcal{A}'$. The vertices in $\mathcal{A}$ can be labeled naturally so that they form a forest with generation levels, as we outline in the following.

By path continuity of B and as $L < \infty$, the sets $\mathcal{A}_0$, $\mathcal{B}$ and $\mathcal{L}$ are finite ($\mathcal{B}$ and $\mathcal{L}$ correspond to the h-extrema from [75]). Hence, the set $\{B(x) : x \in \mathcal{A}\}$ of the levels of the vertices in $\mathcal{A}$ is finite. Finiteness of $\mathcal{A}$ now follows by path continuity of B and as $L < \infty$.

For distinct $x, y \in \mathcal{A}$ such that there exists no $v \in \mathcal{A} \setminus \{x, y\}$ with $x \prec v \prec y$, we say that y is is a left child of x if $(g_y, d_y) \subset (g_x, e_x)$, and that y is a right child of x if $(g_y, d_y) \subset (e_x, d_x)$. Let $M^B = \#\mathcal{A}_0$.

The canonical labeling of the vertices in $\mathcal{A}$ is given by the map $\iota : \mathcal{A} \to \bigcup_{n\in\mathbb{N}_0}\{1, 2\}^n \times [M^B]$, which we define as follows. For $x \in \mathcal{A}_0$, we set $\iota(x) = (\emptyset, m)$, where $m = \#\{y \in \mathcal{A}_0 : g_y \leq g_x\}$ is such that (g_x, d_x) is the m-th interval (counted from the left) of a subexcursion of B with height at least h above level u_0. The other labels are assigned recursively: For $x, y \in \mathcal{A}$ such that x has been assigned some label $\iota(x) = (u, m)$, we set $\iota(y) = (u1, m)$ if y is a left child of x, and $\iota(x) = (u2, m)$ if y is a right child of x. By definition of $\prec$ and finiteness of $\mathcal{A}$, an $\mathbb{F}$-valued random variable is defined by $\mathcal{F}^B = \iota(\mathcal{A})$.

We define generation levels $g^B(i), i \in [\mathcal{H}(\mathcal{F}^B)]$ such that $g^B(i)$ is the i-th smallest element of the set $\{B(x) - u_0 : x \in \mathcal{A} \setminus \mathcal{A}_0\}$ of the levels in $\mathcal{A} \setminus \mathcal{A}_0$, shifted by u_0.

Let $z \in (0, \infty)$. By the results in [65, Théorème 6], [80, Theorem 7.15], [75], [74, Theorem 1.1], under $\mathbb{P}^{(u_0)}$, conditioned on $\zeta(u_0) = z$, the $\mathbb{G}$-valued random variable $(\mathcal{F}^B, g^B)$ is a binary branching forest with rate $2/h$ and a Poisson$(z/(2h))$-distributed number M^B of roots.

By the following lemma, the ranks of the two subexcursions that correspond to the subtrees above a branchpoint $x \in \mathcal{B}$ are the ranks of the children of the corresponding vertex $\iota(x)$ in $\mathcal{F}^B$.

Lemma 4.5. *Let $x \in \mathcal{B}$, $u = B(x)$, $(v, m) = \iota(x)$. Then (v, m) has two children. Denoting their ranks by $i = R(v1, m)$ and $j = R(v2, m)$, we obtain*

$$\{(g_{u,i}, d_{u,i}), (g_{u,j}, d_{v,j})\} = \{(g_x, e_x), (e_x, d_x)\}.$$

Proof. Let $h_{x1} = \max_{[g_x, e_x]} B - u$, $h_{x2} = \max_{[e_x, d_x]} B - u$ be the subexcursion heights that correspond to x and that are larger than h by definition of $\mathcal{B}$. By construction, there exist $y_1, y_2 \in \mathcal{A}$ with $(g_{y_1}, d_{y_1}) \subset (g_x, e_x)$, $(g_{y_2}, d_{y_2}) \subset (e_x, d_x)$ such that the levels $B(y_1)$, $B(y_2)$ are maximized. Then these y_1, y_2 satisfy $B(y_1) = h_{x1} - h$, $B(y_2) = h_{x2} - h$. In particular, $\iota(x)$ has two children.

As the local minima of B are pairwise distinct, the vertices $y \in \mathcal{A}$ with $y \neq x$ and $B(y) = u$ are edge points, that is, the (g_y, d_y) are subexcursion intervals above level u. We denote their heights by $h_y = \max_{[g_y, d_y]} B - u$. We denote by $\mathcal{Y}$ the set of these vertices y. For each $y \in \mathcal{Y}$, the maximal level of vertices $y' \in \mathcal{A}$ with $(g_{y'}, d_{y'}) \subset [g_y, d_y]$ is $h_y - h$.

By construction of $\mathcal{A}$, the definition of the rank in Subsection 4.2, and as the local maxima of B are pairwise distinct, it follows that $\{h_{x1}, h_{x2}\}$ consists of the i-th and j-th highest elements of $\{h_y : y \in \mathcal{Y}\} \cup \{h_{x1}, h_{x2}\}$, which yields the assertion. □

By the following lemma, the numbers $\xi^h_{i,j}(u)$ of lookdown events above level u_0 in which subexcursions of height larger than h merge, and the downcrossing numbers $D^h(u)$ can also be read off from $(\mathcal{F}^B, g^B)$.

Lemma 4.6. *Let $u \in \mathbb{R}_+$ and $1 \leq i < j$. Then we have:*

(i) A. s., $D^h(u) = N_n(\mathcal{F}^B)$ for $n \in \mathbb{N}$ such that $g^B(n-1) \leq u < g^B(n)$.

(ii) Furthermore,

$$\xi^h_{i,j}(u) = \#\{(v, m) \in \mathcal{F}^B : g^B(|v|) \leq u, v2 \in \mathcal{F}^B, \{R(v1, m), R(v2, m)\} = \{i, j\}\}.$$

Proof. For (i), we use that $g^B(n) \neq u$ a. s. No subexcursions of height at least h above level $g^B(n)$ merge in $(u, g^B(n))$, as otherwise there would be a branchpoint in $\mathcal{B}$ at a level in $(u, g^B(n))$, in contradiction to the definition of g^B. Also, every subexcursion above level u of height at least h contains a subexcursion of height at least h above level $g^B(n)$, as otherwise there would be a leaf in $\mathcal{L}$ at a level in $[u, g^B(n))$. Hence, there are as many subexcursions of height at least h above level u as there are vertices in $\mathcal{A}$ at level $g^B(n)$, which yields assertion (i).

To prove (ii), note that it follows from the definition in Section 4.1 that $\xi^h_{i,j}(u)$ is the number of branchpoints $x \in \mathcal{B}$ with $B(x) \in (u_0, u_0 + u]$ such that $\{\max_{[g_x,e_x]} B - B(x), \max_{[e_x,d_x]} B - B(x)\}$ consist of the i-th and j-th highest elements of the set of subexcursion heights at level $B(x)$. Assertion (ii) now follows from Lemma 4.5 and as each vertex of $\mathcal{F}^B$ with two children corresponds to a branchpoint in $\mathcal{B}$. □

As we work on the event of probability 1 on which $L < \infty$ and the local extrema of B are pairwise distinct, the jumps of $u \mapsto D^h(u)$ are of size 1. Hence, the convergence (4.5) still holds when we work with càdlàg versions of the processes $(D^h(u), u \in \mathbb{R}_+)$, which we do from now on.

4.4 A lookdown forest

In this subsection, we consider a random forest that corresponds to "Model II" of Donnelly and Kurtz, cf. Sections 1 and 2 in [28].

We define a ranked forest as a pair $(\mathcal{F}, \mathcal{R})$ where $\mathcal{F} \in \mathbb{F}$ and $\mathcal{R} \in \mathbb{N}^{\mathcal{F}}$. We denote the set of ranked forests by $\hat{\mathbb{F}}$. (That is, each vertex of $\mathcal{F}$ is marked with an integer, following the notion of a marked tree from [73].) We call a triple $(\mathcal{F}, g, \mathcal{R})$ a ranked forest with generation levels if $(\mathcal{F}, g) \in \mathbb{G}$ and $(\mathcal{F}, \mathcal{R}) \in \hat{\mathbb{F}}$. We denote the set of ranked forests with generation levels by $\hat{\mathbb{G}}$.

For a forest $\mathcal{F} \in \mathbb{F}$, the set $\{(u, m) : (u, m) \in \mathcal{F}, |u| = n\}$ of the vertices in generation $n \in \mathbb{N}_0$ can be ordered naturally, for instance, such that a vertex (u, m) precedes a vertex (u', m') if and only if either $m < m'$ or both $m = m'$ and u precedes u' lexicographically. In this sense, we view a collection indexed by $\{(u, m) : (u, m) \in \mathcal{F}, |u| = n\}$ as a vector of length $N_n(\mathcal{F})$. We denote by $\mathcal{R}_n = (\mathcal{R}(u, m) : (u, m) \in \mathcal{F}, |u| = n)$ the vector of the ranks of the vertices of a ranked forest $(\mathcal{F}, \mathcal{R}) \in \hat{\mathbb{F}}$ in generation n.

Let M^{ld} be a Poisson distributed random variable with parameter $z/(2h)$. Now we define a $\hat{\mathbb{G}}$-valued random variable $(\mathcal{F}^{\mathrm{ld}}, g^{\mathrm{ld}}, \mathcal{R}^{\mathrm{ld}})$ by the process $(\gamma_n(\mathcal{F}^{\mathrm{ld}}, g^{\mathrm{ld}}, \mathcal{R}^{\mathrm{ld}}), n \in \mathbb{N}_0)$ of its growing restrictions. We start with the roots $(\emptyset, m)$, $m \in [M^{\mathrm{ld}}]$. Let the vector of their ranks $(\mathcal{R}^{\mathrm{ld}}(\emptyset, 1), \dots, \mathcal{R}^{\mathrm{ld}}(\emptyset, M^{\mathrm{ld}}))$ be an independent uniform permutation of $[M^{\mathrm{ld}}]$.

We obtain generation $n+1$ conditionally given $\gamma_n(\mathcal{F}^{\mathrm{ld}}, g^{\mathrm{ld}}, \mathcal{R}^{\mathrm{ld}})$ in a two-step random experiment. It suffices to work on the event $\{N_n(\mathcal{F}^{\mathrm{ld}}) \geq 1\}$. In the first step, we decide with equal probability whether a birth or a death event occurs.

If a death event occurs, then in generation n, the vertex $(u, m) \in \mathcal{F}^{\mathrm{ld}}$ with the maximal rank $\mathcal{R}^{\mathrm{ld}}(u, m) = N_n(\mathcal{F}^{\mathrm{ld}})$ has no children. Each of the other vertices (v, p) in generation n has one child $(v1, p)$ which inherits the rank of its parent, $\mathcal{R}^{\mathrm{ld}}(v1, p) = \mathcal{R}^{\mathrm{ld}}(v, p)$.

If a birth event occurs, then we draw uniformly and independently a pair $i < j$ of integers from $[N_n(\mathcal{F}^{\mathrm{ld}}) + 1]$, and a permutation (a, b) of $\{1, 2\}$. We let the vertex (u, m) with rank $\mathcal{R}^{\mathrm{ld}}(u, m) = i$ in generation n have 2 children with ranks $\mathcal{R}^{\mathrm{ld}}(ua, m) = i$ and $\mathcal{R}^{\mathrm{ld}}(ub, m) = j$. Each of the other vertices (v, p) in generation n has one child and the relative ranks are preserved, i. e. $\mathcal{R}^{\mathrm{ld}}(v1, p) = \mathcal{R}^{\mathrm{ld}}(v, p)$ for $(v, p) \neq (u, m)$ with $\mathcal{R}^{\mathrm{ld}}(v, p) < j$ and $\mathcal{R}^{\mathrm{ld}}(v1, p) = \mathcal{R}^{\mathrm{ld}}(v, p) + 1$ for $(v, p) \neq (u, m)$ with $\mathcal{R}^{\mathrm{ld}}(v, p) \geq j$.

In any case, conditionally given $\gamma_n(\mathcal{F}^{\mathrm{ld}})$ and conditioned on $N_{n+1}(\mathcal{F}^{\mathrm{ld}}) \geq 1$, we let $g^{\mathrm{ld}}(n + 1) - g^{\mathrm{ld}}(n)$ be independent and exponentially distributed with parameter

$2N_n(\mathcal{F}^{\mathrm{ld}})/h$.

The following lemma and its proof can be compared to Sections 2.1 and 2.2 in [28].

Lemma 4.7. *The* $\mathbb{G}$*-valued random variable* $(\mathcal{F}^{\mathrm{ld}}, g^{\mathrm{ld}})$ *is a binary branching forest with rate* $2/h$.

Proof. We have to show that the $\mathbb{G}$-valued Markov chain $(\gamma_n(\mathcal{F}^{\mathrm{ld}}, g^{\mathrm{ld}}), n \in \mathbb{N}_0)$, where the ranks are omitted, has the transition kernel of a binary branching forest. To this aim, we claim for each $n \in \mathbb{N}_0$ that conditionally given $\gamma_n(\mathcal{F}^{\mathrm{ld}})$, the vector $\mathcal{R}_n^{\mathrm{ld}}$ of the ranks of the vertices in generation n is a uniform permutation of $[N_n(\mathcal{F}^{\mathrm{ld}})]$. The claim implies that conditionally given $\gamma_n(\mathcal{F}^{\mathrm{ld}})$, it is a uniformly drawn individual that has no children in a death event from generation n to generation $n+1$, or two children in a birth event. The assertion then follows as $g^{\mathrm{ld}}(n+1) - g^{\mathrm{ld}}(n)$ is conditionally independent and has the required distribution.

The claim holds for $n = 0$. We proceed by induction. First we consider the event that a birth event occurs from generation n to $n+1$. Let $f \in \mathbb{F}$ such that f can occur as a typical realization of $\gamma_n(\mathcal{F}^{\mathrm{ld}})$. Let N be the number of vertices of f in generation n. W. l. o. g., we assume $N \geq 1$. Let $i < j$ be integers in $[N+1]$, let the pair (a, b) be a permutation of $\{1, 2\}$, and let $f' \in \mathbb{F}$ be the realization of $\gamma_{n+1}(\mathcal{F}^{\mathrm{ld}})$ that occurs when $\gamma_n(\mathcal{F}^{\mathrm{ld}}) = f$ and a birth event occurs from generation n to $n+1$ that is characterized as in the above definition of $\mathcal{F}^B$ by our given i, j, a, b. Moreover, let p be a permutation of $[N]$, and let p' be any of the two permutations of $[N+1]$ that can occur as typical realizations of $\mathcal{R}_{n+1}^{\mathrm{ld}}$ when $\mathcal{R}_n^{\mathrm{ld}} = p$ and $\gamma_{n+1}(\mathcal{F}^{\mathrm{ld}}) = f'$.

Then,

$$
\begin{aligned}
&\mathbb{P}(\gamma_{n+1}(\mathcal{F}^{\mathrm{ld}}) = f', \mathcal{R}_{n+1}^{\mathrm{ld}} = p') \\
&= \mathbb{P}(\gamma_n(\mathcal{F}^{\mathrm{ld}}) = f, \mathcal{R}_n^{\mathrm{ld}} = p) \frac{1}{2} \frac{1}{\binom{N+1}{2}} \frac{1}{2} \\
&= \frac{1}{2} \frac{1}{N} \mathbb{P}(\gamma_n(\mathcal{F}^{\mathrm{ld}}) = f) \frac{1}{N+1} \frac{1}{N!} \\
&= \mathbb{P}(\gamma_{n+1}(\mathcal{F}^{\mathrm{ld}}) = f') \frac{1}{(N+1)!}
\end{aligned}
$$

In the second line, the first factor $1/2$ appears as birth and death events occur with equal probability, the next factor from drawing i and j, and the last factor $1/2$ from drawing the permutation (a, b). The induction hypothesis, namely that $\mathcal{R}_n^{\mathrm{ld}}$ is a uniform permutation of $[N_n(\mathcal{F}^{\mathrm{ld}})]$, is used in the second equality. In the second last line, the factor $1/2$ again stands for the probability of a birth event, and the factor $1/N$ can be interpreted as the uniform weight on a vertex of generation n. Then, for the third equality, we use the assertion on the transition probability from generations n to $n+1$ which follows already from the induction hypothesis.

Similarly, to account for a death event from generation n to $n+1$, let f, N, and p be defined as above. Let (f'', p'') be the realization of $(\gamma_{n+1}(\mathcal{F}^{\mathrm{ld}}), \mathcal{R}_{n+1}^{\mathrm{ld}})$ that occurs when

$(\gamma_n(\mathcal{F}^{\mathrm{ld}}), \mathcal{R}_n^{\mathrm{ld}}) = (f, p)$ and a death event occurs. Then,

$$\begin{aligned}
&\mathbb{P}(\gamma_{n+1}(\mathcal{F}^{\mathrm{ld}}) = f'', \mathcal{R}_{n+1}^{\mathrm{ld}} = p'') \\
&= \mathbb{P}(\gamma_n(\mathcal{F}^{\mathrm{ld}}) = f, \mathcal{R}_n^{\mathrm{ld}} = p)\frac{1}{2} \\
&= \frac{1}{2}\frac{1}{N}\mathbb{P}(\gamma_n(\mathcal{F}^{\mathrm{ld}}) = f)\frac{1}{(N-1)!} \\
&= \mathbb{P}(\gamma_{n+1}(\mathcal{F}^{\mathrm{ld}}) = f'')\frac{1}{(N-1)!}.
\end{aligned}$$

The factor 1/2 in the second line is the probability for a death event, and again the induction hypothesis is used in the second and third equalities. □

By the following lemma, the ranks $\mathcal{R}^{\mathrm{ld}}$ are determined by $\mathcal{F}^{\mathrm{ld}}$ as they coincide with the ranks from Section 4.2. (However, the ranks $\mathcal{R}_n^{\mathrm{ld}}$ of generation n give information about the future and cannot be read off from the restriction $\gamma_n(\mathcal{F}^{\mathrm{ld}})$).

Lemma 4.8. *For $(u, m) \in \mathcal{F}^{\mathrm{ld}}$, we have $\mathcal{R}^{\mathrm{ld}}(u, m) = R(u, m)$.*

Proof. The definitions of the birth and death events in $\mathcal{F}^{\mathrm{ld}}$ imply that a vertex of the lookdown forest with a lower $\mathcal{R}^{\mathrm{ld}}$-rank has descendants until a later generation compared to a vertex with a higher $\mathcal{R}^{\mathrm{ld}}$-rank in the same generation. □

For $u \in \mathbb{R}_+$, we set $\tilde{D}^h(u) = N_n(\mathcal{F}^{\mathrm{ld}})$ where $n \in \mathbb{N}$ is such that $g^{\mathrm{ld}}(n-1) \le u < g^{\mathrm{ld}}(n)$. Then $\tilde{D}^h(u)$ can be interpreted as the number of branches at level u. Moreover, for $1 \le i < j$, we define

$$\tilde{\xi}_{i,j}^h(u) = \#\{(v, m) \in \mathcal{F}^{\mathrm{ld}} : g^{\mathrm{ld}}(|v|) \le u, v2 \in \mathcal{F}^{\mathrm{ld}}, \{R(v1, m), R(v2, m)\} = \{i, j\}\}.$$

Then $\tilde{\xi}_{i,j}^h(u)$ is the number of birth events in which the parent generation lives not above level u and the newborn pair obtains the ranks i and j.

We fix $k \in \mathbb{N}$ with $k \ge 2$ and define $\mathbb{N}_0^{\binom{k}{2}}$-valued processes $(\xi^h(u), u \in \mathbb{R}_+)$ and $(\tilde{\xi}^h(u), u \in \mathbb{R}_+)$ by $\xi^h(u) = (\xi_{i,j}^h(u), 1 \le i < j \le k)$ and $\tilde{\xi}^h(u) = (\tilde{\xi}_{i,j}^h(u), 1 \le i < j \le k)$. The following corollary is a consequence of Lemmas 4.6 and 4.7.

Corollary 4.9. *Under $\mathbb{P}^{(u_0)}$, conditioned on $\zeta(u_0) = z$, the process $((D^h(u), \xi^h(u)), u \in \mathbb{R}_+)$ is distributed as $((\tilde{D}_u^h, \tilde{\xi}^h(u)), u \in \mathbb{R}_+)$ in the space of càdlàg paths.*

4.5 Proof of Lemma 4.1

Let $k \ge 2$ as in the end of the previous subsection. We define a. s. an $\mathbb{N}_0^{\binom{k}{2}}$-valued process $(\eta(t), t \in \mathbb{R}_+)$ by

$$\eta(t) = (\eta_{i,j}(t), 1 \le i < j \le k),$$

and we set

$$Y(t) = \zeta(V(t)), \quad t \in \mathbb{R}_+.$$

By the definitions (4.3) and (4.7),

$$(\eta_{i,j}(t), t \in \mathbb{R}_+) = (\xi_{i,j}(V(t) - u_0), t \in \mathbb{R}_+). \tag{4.9}$$

By Lemma 3.3 and as in Lemma 2.3, $\eta_{i,j}(t) < \infty$ for all $t \in \mathbb{R}_+$ a. s.

In Lemma 4.10 below, we describe the process $((Y(t), \eta(t)), t \in \mathbb{R}_+)$ by a martingale problem. We denote by $(\mathcal{F}_t, t \in \mathbb{R}_+)$ the filtration induced by $((Y(t), \eta(t)), t \in \mathbb{R}_+)$.

We denote by $C_c^\infty(0,\infty)$ the set of infinitely often differentiable functions $(0,\infty) \to \mathbb{R}$ with compact support. Let $\mathscr{D}$ be the set of functions $f : (0,\infty) \times \mathbb{N}_0^{\binom{k}{2}} \to \mathbb{R}$ that can be written as $f(x,n) = f_1(x)f_2(n)$ with $f_1 \in C_c^\infty(0,\infty)$ and bounded $f_2 : \mathbb{N}_0^{\binom{k}{2}} \to \mathbb{R}$.

For $1 \le i < j \le k$, let

$$\alpha_{i,j} : \mathbb{N}_0^{\binom{k}{2}} \to \mathbb{N}_0^{\binom{k}{2}}, \quad (n_{i',j'}, 1 \le i' < j' \le k) \mapsto (n_{i',j'} + \mathbf{1}\{i' = i, j' = j\}, 1 \le i' < j' \le k)$$

be the function that increases the component $n_{i,j}$ by one.

Let G be the operator that maps a function $f \in \mathscr{D}$ to the bounded continuous function Gf on $(0,\infty) \times \mathbb{N}_0^{\binom{k}{2}}$ with

$$Gf(x,n) = 2x^2 \partial_x^2 f(x,n) + \sum_{1 \le i < j \le k} (f(x, \alpha_{i,j} n) - f(x,n))$$

for $(x,n) \in (0,\infty) \times \mathbb{N}_0^{\binom{k}{2}}$.

Lemma 4.10. *For each $f \in \mathscr{D}$, the process*

$$f(Y(t), \eta(t)) - \int_0^t Gf(Y(s), \eta(s)) ds, \quad t \in \mathbb{R}_+$$

is an $(\mathcal{F}_t)$-martingale under $\mathbb{P}^{(u_0)}$.

For the proof of Lemma 4.10, we use the lookdown representation of the process $((2hD^h(u), \xi^h(u)), u \in \mathbb{R}_+)$ given by Corollary 4.9. We will let h tend to zero after applying a time-change. We define this time-change by

$$\sigma^h(u) = \int_0^u \frac{4}{2hD^h(v)} dv, \quad u \in \mathbb{R}_+ \tag{4.10}$$

and

$$V^h(t) = \inf\{u \in \mathbb{R}_+ : \sigma^h(u) > t\}, \quad t \in \mathbb{R}_+. \tag{4.11}$$

Then we define the population size (scaled by $2h$) under the time change by

$$Y^h(t) = 2hD^h(V^h(t)), \quad t \in \mathbb{R}_+$$

and the numbers of lookdowns by

$$\eta^h_{i,j}(t) = \xi^h_{i,j}(V^h(t)), \quad t \in \mathbb{R}_+, 1 \le i < j. \tag{4.12}$$

We also set

$$\eta^h(t) = (\eta^h_{i,j}(t), 1 \le i < j \le k), \quad t \in \mathbb{R}_+.$$

We now relate the time changes given by $V^h(t)$ and $V(t)$.

Lemma 4.11. *For each $t \in \mathbb{R}_+$,*

$$u_0 + \lim_{h \downarrow 0} V^h(t) = V(t).$$

Proof. By the consequence (4.5) of the downcrossing representation, and by the definitions (4.1) and (4.10),

$$\lim_{h \downarrow 0} \sup_{u \in [0,v]} |\sigma^h(u) - \sigma(u_0 + u)| = 0$$

for all $v \in [0, H - u_0)$. Here we also use that $\min_{u \in [0,v]} \zeta(u_0 + u) > 0$. By the definition (4.2) of $V(t)$,

$$V(t) = u_0 + \inf\{u \in [0, H - u_0) : \sigma(u_0 + u) > t\}.$$

The assertion now follows from the definition (4.11) of $V^h(t)$ and as $u \mapsto \sigma(u)$ is strictly increasing on $[0, H - u_0)$. □

By the following lemma, the process $(\eta_{i,j}(t), t \in \mathbb{R}_+)$ has a. s. no jumps at fixed times.

Lemma 4.12. *For each $1 \le i < j \le k$ and $t \in (0, \infty)$, we have $\xi_{i,j}\{V(t)\} = 0$ a. s., hence $\xi_{i,j}(V(t) - u_0) = \xi_{i,j}(V(t) - u_0-)$ a. s.*

Proof. We use that $V(t)$ is a stopping time with respect to the filtration $(\mathcal{E}_u, u \in \mathbb{R}_+)$ where $\mathcal{E}_u$ is the sigma field generated by the path of B below level u, which can be defined as the time-changed path $(B(\tilde{\alpha}(s)), s \in \mathbb{R}_+)$, where $\tilde{\alpha}(s) = \inf\{t \in \mathbb{R}_+ : \int_0^t \mathbf{1}\{B(r) \le u\}\, dr > s\}$. This filtration is known as the excursion filtration.

By the strong Markov property from [5, Lemma 17] which also applies in the Brownian case, the concatenation of the subexcursions of B above level $V(t)$ forms stopped reflected Brownian motion under $\mathbb{P}^{(u_0)}$, hence by excursion theory, the i-th and the j-th highest subexcursion of B above level $V(t)$ are a. s. not adjacent, thus j does a. s. not look down on i at $V(t)$. □

From the above lemmas, we now deduce that also under the time changes, the scaled population sizes converge to the local time, and that the numbers $\eta^h(t) = (\eta^h_{i,j}(t), 1 \le i < j \le k)$ of lookdown events are well-behaved as $h \downarrow 0$.

Corollary 4.13. *Let $t \in \mathbb{R}_+$. Then, $\lim_{h \downarrow 0} Y^h(t) = Y(t)$. Furthermore, $\eta^h(t) = \eta(t)$ for sufficiently small $h > 0$ a. s.*

Proof. The first assertion follows from the definitions of $Y^h(t)$ and $Y(t)$, Lemma 4.11, and the convergence (4.5). The second assertion is a consequence of equation (4.9), the definition (4.12) of $\eta^h_{i,j}(t)$, and Lemmas 4.11, 3.3, and 4.12. □

Proof of Lemma 4.10. First we work with the lookdown forest from Section 4.4. We define a time change by

$$\tilde{\sigma}^h(u) = \int_0^u \frac{4}{2h\tilde{D}^h(v)} dv, \quad u \in \mathbb{R}_+$$

and

$$\tilde{V}^h(t) = \inf\{u \in \mathbb{R}_+ : \tilde{\sigma}^h(u) > t\}, \quad t \in \mathbb{R}_+.$$

Then we define the scaled population size under the time change by

$$\tilde{Y}^h(t) = 2h\tilde{D}^h(\tilde{V}^h(t)), \quad t \in \mathbb{R}_+$$

and the number of birth events in which the pair of newborns obtains the ranks $1 \le i < j$ by

$$\tilde{\eta}^h_{i,j}(t) = \tilde{\xi}^h_{i,j}(\tilde{V}^h(t)), \quad t \in \mathbb{R}_+.$$

Recall from Lemma 4.8 that the $\mathcal{R}^{\mathrm{ld}}$-ranks from Subsection 4.4 coincide with the ranks from Subsections 4.2.

Now we define an operator G^h such that $((\tilde{Y}^h(t), \tilde{\eta}^h(t)), t \in \mathbb{R}_+)$ solves the martingale problem for $(G^h, \mathscr{D})$. For $1 \le i < j \le k$ and $x \in 2h\mathbb{N}_0$, let

$$p^h_{i,j}(x) = \binom{\frac{x}{2h}+1}{2}^{-1} \mathbf{1}\{\tfrac{x}{2h} + 1 \ge i \vee j\}.$$

In the context of the definition of the lookdown forest in Subsection 4.4, $p^h_{i,j}(x)$ is the probability that in a birth event, the vertex with rank i has two children with ranks i, j when there are $x/(2h)$ many vertices in its generation. We define the generator G^h by

$$\begin{aligned} G^h f(x,n) &= x \sum_{1 \le i < j \le k} \{f(x+2h, \alpha_{i,j}n) - f(x,n)\} \frac{1}{h}\frac{x}{2h} p^h_{i,j}(x) \\ &\quad + x\{f(x+2h,n) - f(x,n)\}\frac{1}{h}\frac{x}{2h}\left(1 - \sum_{1\le i<j\le k} p^h_{i,j}(x)\right) \\ &\quad + x\{f(x-2h,n) - f(x,n)\}\frac{1}{h}\frac{x}{2h} \end{aligned}$$

for $f \in \mathscr{D}$ and $(x,n) \in (0,\infty) \times \mathbb{N}_0^{\binom{k}{2}}$. Each summand in the first term on the right-hand side of the expression for $G^h f(x,n)$ stands for birth events in which the vertex with rank i has two children with ranks $i, j \in [k]$. When we grow the forest in continuous time as in the definition of $(\tilde{\xi}^h(u), u \in \mathbb{R}_+)$ in Subsection 4.4, such birth events occur with rate $\frac{1}{h}\frac{x}{2h}p^h_{i,j}(x)$ when the population size is $\frac{x}{2h}$, and by the time change we obtain an additional factor x. The second term stands for the birth events in which the ranks i and j are not both in $[k]$, hence these birth events are not counted by $(\tilde{\xi}^h(u), u \in \mathbb{R}_+)$. Nevertheless, these birth events increase the scaled population size $2h\tilde{D}^h(u)$ by $2h$. The third term stands for the death events.

By construction, the process $((\tilde{Y}^h(t), \tilde{\eta}^h(t)), t \in \mathbb{R}_+)$ solves the martingale problem for $(G^h, \mathscr{D})$. By Corollary 4.9, also the process $((Y^h(t), \eta^h(t)), t \in \mathbb{R}_+)$ under $\mathbb{P}^{(u_0)}$ solves the martingale problem for $(G^h, \mathscr{D})$. Hence,

$$\begin{aligned} &\mathbb{E}^{(u_0)}[(f(Y^h(t), \eta^h(t)) - f(Y^h(s), \eta^h(s)) \\ &\qquad - \int_s^t G^h f(Y^h(r), \eta^h(r))dr) \prod_{i=1}^{\ell} g_i(Y^h(s_i), \eta^h(s_i))] = 0. \end{aligned} \tag{4.13}$$

for all $f \in \mathscr{D}$, $\ell \in \mathbb{N}$, $0 \le s_1 \le \ldots \le s_\ell \le s \le t$, and bounded continuous functions $g_1, \ldots, g_\ell$.

A straightforward calculation (using Taylor's formula, the definition of $\mathscr{D}$, and that $p^h_{i,j}(x) \sim 2h^2/x^2$) shows that

$$\lim_{h \downarrow 0} \sup\{|G^h f(x,n) - Gf(x,n)| : (x,n) \in (0,\infty) \times \mathbb{N}_0^{\binom{k}{2}}\} = 0 \tag{4.14}$$

for each $f \in \mathscr{D}$.

Using Corollary 4.13 and dominated convergence, we deduce that the left-hand side of equation (4.13) converges to

$$\begin{aligned} &\mathbb{E}^{(u_0)}[(f(Y(t),\eta(t)) - f(Y(s),\eta(s)) \\ &\qquad - \int_s^t Gf(Y(r),\eta(r))dr) \prod_{i=1}^{\ell} g_i(Y(s_i),\eta(s_i))] \end{aligned} \tag{4.15}$$

as $h \downarrow 0$. Here we use that for $C_1 := \sup|g_1 \cdots g_\ell|$ and $C_2 := \sup|Gf|$,

$$\begin{aligned} &|\mathbb{E}[\int_s^t G^h f(Y^h(r),\eta^h(r))dr \prod_{i=1}^{\ell} g_i(Y^h(s_i),\eta^h(s_i))] \\ &\quad - \mathbb{E}[\int_s^t Gf(Y(r),\eta(r))dr \prod_{i=1}^{\ell} g_i(Y(s_i),\eta(s_i))]| \\ &\le C_1(t-s)\sup\{|G^h f(x,n) - Gf(x,n)| : (x,n) \in (0,\infty) \times \mathbb{N}_0^{\binom{k}{2}}\} \\ &\quad + C_1 \int_s^t \mathbb{E}[|Gf(Y^h(r),\eta^h(r)) - Gf(Y(r),\eta(r))|]dr \\ &\quad + C_2(t-s)\mathbb{E}[|\prod_{i=1}^{\ell} g_i(Y^h(s_i),\eta^h(s_i)) - \prod_{i=1}^{\ell} g_i(Y(s_i),\eta(s_i))|], \end{aligned}$$

and that this expression converges to zero as $h \to 0$ by (4.14), Corollary 4.13, and dominated convergence. It now follows from equation (4.13) that expression (4.15) equals zero, which yields the assertion. □

Now we deduce Lemma 4.1 from Lemma 4.10.

Proof of Lemma 4.1. First we work under $\mathbb{P}^{(u_0)}$. By Lemma 4.10, the process $(Y_t, t \in \mathbb{R}_+)$ solves the martingale problem for $(G_0, C_c^\infty(0,\infty))$ with generator given by

$$G_0 f(x) = 2x^2 \partial_x^2 f(x), \quad x \in (0,\infty).$$

This martingale problem is unique by Theorems 8.2.1, 1.2.6, 4.4.1, and 4.6.2 of [36]. By Lemma 4.10, for $1 \le i < j \le k$, the process $(\eta_{i,j}(t), t \in \mathbb{R}_+)$ solves the martingale problem $(G', \mathscr{D}')$ where $\mathscr{D}'$ is the set of bounded functions $\mathbb{N}_0^{\binom{k}{2}} \to \mathbb{R}$ and

$$G' f(n) = f(n+1) - f(n), \quad n \in \mathbb{N}_0.$$

Hence, $(\eta_{i,j}(t), t \in \mathbb{R}_+)$ is a Poisson process with rate 1. By Theorem 4.10.1 of [36] and Lemma 4.10, a collection of independent processes is given by $(Y(t), t \in \mathbb{R}_+)$ and $(\eta_{i,j}(t), t \in \mathbb{R}_+)$ for $1 \le i < j \le k$.

As $V(0) = u_0$ and $\sigma(V(t)) = t$ for all $t \in \mathbb{R}_+$ a. s., we have

$$V(t) = u_0 + \int_0^t \tfrac{1}{4} Y(s)\, ds$$

for all $t \in \mathbb{R}_+$ a. s., cf. e. g. equation (1.5) in [36, Chapter 6]. Hence, as the right-continuous inverse of V is again σ by continuity of the time change,

$$-\tau(u_0) = \sigma(1) = \inf\{t \in \mathbb{R}_+ : u_0 + \int_0^t \tfrac{1}{4} Y(s) ds > 1\} \quad \text{a. s.}$$

As in Lemma 2.3, $\{H > 1\} = \{\sigma(1) < \infty\}$ up to $\mathbb{P}^{(u_0)}$-null events. Hence, the collection $((\eta(t), t \in \mathbb{R}_+), 1 \le i < j)$ is independent of the event $\{H > 1\}$ under $\mathbb{P}^{(u_0)}$, and thus has the same distribution under $\mathbb{P}^{(u_0)}$ as under $\mathbb{P} = \mathbb{P}^{(u_0)}(\cdot | H > 1)$. It also follows that $((\eta(t), t \in \mathbb{R}_+), 1 \le i < j)$ is independent of $\tau(u_0)$ under $\mathbb{P}$. $\square$

Chapter 5

Invariance principles for tree-valued Cannings chains

We consider sequences of tree-valued Markov chains that describe evolving genealogies in Cannings models, and we show their convergence in distribution to tree-valued Fleming-Viot processes. Under the conditions of Möhle and Sagitov [72, 85], this convergence holds for all tree-valued Fleming-Viot processes from Chapter 2 in the dust-free case, and for the Fleming-Viot processes with values in the space of distance matrix distributions in the case with dust. We show convergence to Fleming-Viot processes with values in the space of marked metric measure spaces in the case with dust under an additional assumption on the probability that a randomly sampled individual belongs to a non-singleton family.

1 Introduction

In population genetics, Cannings models [18, 19] are classical evolutionary models with constant population size N and non-overlapping generations. Reproduction events occur independently between the generations such that the individuals in generation $k+1$ are subdivided into families according to an exchangeable random partition and each family draws its ancestor in generation k independently without replacement. The Wright-Fisher model is a classical example for a Cannings model. In the Wright-Fisher model, each individual in generation $k+1$ draws its ancestor in generation k independently with replacement. We refer the reader also to e. g. [33] for these models.

Coalescents with simultaneous multiple mergers are robust infinite population size limits of partition-valued processes that describe the genealogies in Cannings models at fixed times. Möhle and Sagitov [72] give a criterion for this convergence. Sagitov [85] gives an equivalent criterion in terms of the measure Ξ of Schweinsberg [86]. A first robustness result for the Kingman coalescent is shown in Kingman [60], see also e. g. [11, Theorem 2].

In the present work, we consider evolving genealogies, and we describe the genealogical tree of the individuals in Cannings models at each time in various ways. First, we consider the metric measure space that consists of the set of individuals at that time, their mutual genealogical distances, and the uniform probability measure. Second, we consider the distance matrix distribution of the aforementioned metric measure space, i. e. the distribution of the infinite matrix of the genealogical distances between iid samples. Third, we decompose the genealogical tree into the external branches and the remaining subtree. We then consider the semi-metric space that consists of the starting vertices of the external branches and their mutual distances. On the product space of this semi-metric space with $\mathbb{R}_+$, we define a probability measure such that for each external branch, mass $1/N$ is added to the pair that consists of its starting vertex and its length. We then obtain a marked metric measure space. Fourth, we also decompose the genealogical tree and describe it by a marked metric measure space, pruning not the whole external branch, but only the part from each leaf to the most recent reproducing individual on the ancestral lineage.

We endow the space of distance matrix distributions with the Prohorov metric, and the space of (marked) metric measure spaces with the (marked) Gromov-Prohorov metric [24, 45]. We consider Markov chains whose states describe the evolving genealogy in Cannings models in one of the four aforementioned ways. Our invariance principles show that sequences of such Markov chains converge in distribution, under a rescaling of time and genealogical distances, in the space of càdlàg paths in the respective state space, endowed with the Skorohod topology. The limit processes are the Fleming-Viot processes from Section 8 in Chapter 2. Under the condition of Sagitov [85] and the assumption that the initial states converge, we show the convergence of the prelimiting chains with values in the space of distance matrices. For the convergence of the prelimiting chains with values in the space of metric measure spaces, we have to assume in addition that the limiting genealogy is dust-free as the tree-valued Fleming-Viot process with values in the space of metric measure spaces exists only in the dust-free case. Dust-freeness can be characterized by the property that a randomly drawn external branch has a. s. length zero (cf. Propositions 6.6 and 7.4 in Chapter 2).

In the dust-free case, the sequences of prelimiting chains for the third and the fourth description of the genealogy converge under the condition of Sagitov [85] and an appropriate condition on the initial state. We show their convergence in the case with dust under the additional assumption (3.8) on the probability that a randomly sampled individual from a fixed generation belongs to a non-singleton family. An additional assumption is needed here as the convergence in the marked Gromov-Prohorov metric (other than the weak convergence of the distance matrix distributions) implies weak convergence of the empirical distribution of the external branch lengths or the distances to the most recent reproducing individual, respectively.

In Section 2.1, we recall the decomposition of the genealogical trees at the external branches. In Section 2.2, we recall some notions on metric measure spaces and marked metric measure spaces. We state our convergence results in Section 3. We recall tree-valued Ξ-Fleming-Viot processes in Section 4. In Section 5, we give an example in which assumption (3.8) is not satisfied and the chains with values in the space of marked metric measure spaces do not converge. The proofs of the invariance principles are given in the further sections.

The tree-valued Fleming-Viot process with binary reproduction events is introduced in Greven, Pfaffelhuber, and Winter [46] as the solution of a well-posed martingale problem that is the limit in distribution of tree-valued processes read off from Moran models. In [46, Remark 2.21], it is conjectured that a tree-valued Fleming-Viot process is the robust limit of tree-valued processes read off from Cannings models. In Chapter 2, tree-valued Fleming-Viot processes are studied in the setting with simultaneous multiple reproduction events, the case with dust is included by the decomposition of the genealogical trees into the external branches and the remaining subtree. These decomposed genealogical trees are described by marked metric measure spaces and their distance matrix distributions. Path regularity of tree-valued Fleming-Viot processes follows from the pathwise construction in Chapter 3.

In Section 6, we prove the invariance principles for the Markov chains associated with the first, second, and in the case with dust also for the fourth of the above descriptions of the genealogy. Here we can apply a general convergence result from Ethier and Kurtz [36, Chapter 4.8] as the transition kernels of the prelimiting chains converge on a core (in the sense of [36, Chapter 1.3]) to the generators of the tree-valued Fleming-Viot processes. In Chapter 2, it is shown that the domains of the martingale problems for the tree-valued Fleming-Viot processes are cores, and that the semigroups on these cores are strongly continuous. We use existence and path regularity of the limit processes, and we do not need to show relative compactness of the prelimiting processes here.

We prove the invariance principles for the processes with values in the space of marked metric measure spaces in Sections 7 and 8 by comparison with processes whose convergence is proved in Section 6. In the case without dust, we can compare these process in the supremum metric. In the case with dust, we compare only the finite dimensional distributions whence we also have to check then the relative compactness of the prelimiting processes.

We use also exchangeable random partitions, Kingman's correspondence and its continuity properties, for which we refer to Pitman [80, Chapter 2] and Bertoin [7, Chapter 2]. We mention that Stournaras [88] uses [36, Corollary 4.8.17] to show, by verifying the compact containment condition via [46, Proposition 2.22] with some effort, convergence of tree-valued Wright-Fisher models to the tree-valued Fleming-Viot process from [46].

2 Preliminaries

It is well-known that ultrametric spaces can be viewed as leaf-labeled real trees (cf. e. g. Remark 1.1 in Chapter 2). In Subsection 2.1, we recall a decomposition of semi-ultrametrics that corresponds to the decomposition of the associated trees at the external branches. In Subsection 2.2, we recall isomorphy classes of metric measure spaces and marked metric measure spaces which we can be interpreted as unlabeled genealogical trees. In the finite case, the (marked) metric measure space associated with a (decomposed) ultrametric can be viewed as the equivalence class under permutations of the labels of the leaves.

2.1 Distance matrices

We denote the set of the positive integers by $\mathbb{N}$, the set of the non-negative integers by $\mathbb{N}_0$, and for $N \in \mathbb{N}$, we write $[N] = \{1, \dots, N\}$. Let $\mathfrak{D}$ denote the space of semimetrics on $\mathbb{N}$ and $\mathfrak{U} \subset \mathfrak{D}$ the set of semi-ultrametrics on $\mathbb{N}$. We do not distinguish between a semimetric $\rho \in \mathfrak{D}$ and the distance matrix $(\rho(i,j))_{i,j\in\mathbb{N}}$, and we view $\mathfrak{U}$ and $\mathfrak{D}$ as subspaces of the space $\mathbb{R}^{\mathbb{N}^2}$ which we endow with a complete and separable metric that induces the product topology, where $\mathbb{R}$ is endowed with the Euclidean topology. Analogously, for $N \in \mathbb{N}$, we denote by $\mathfrak{D}_N$ the space of semimetrics on $[N]$ and by $\mathfrak{U}_N \subset \mathfrak{D}_N$ the space of semi-ultrametrics on $[N]$. Again we do not distinguish between semi-metrics and distance matrices and we view $\mathfrak{D}_N$ as a subspace of the finite-dimensional space $\mathbb{R}^{N^2}$ which we endow with a norm and the (induced) Euclidean topology.

We now decompose semi-ultrametrics in $\mathfrak{U}_N$ as in Section 2 of Chapter 2. The continuous map

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

is used to retrieve ultrametric distance matrices from the elements of the space

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

of decomposed semi-ultrametrics which we also call marked distance matrices. Conversely, for $N \geq 2$, we decompose ultrametric distance matrices using the maps

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

and $\beta : \mathfrak{U}_N \to \hat{\mathfrak{U}}_N$, $\rho \mapsto (r, v)$, where $v = \Upsilon(\rho)$ and $r(i,j) = (\rho(i,j) - v(i) - v(j))\, \mathbf{1}\{i \neq j\}$ for $i, j \in [N]$. For $\rho \in \mathfrak{U}_1$, we set $\Upsilon(\rho) = 0$ and $\beta(\rho) = (\rho, 0)$.

As in Remark 2.2 of Chapter 2, the quantity $v(i)$ is the length of the external branch that ends in leaf i of the coalescent tree associated with ρ, and $r(i,j)$ is the distance between the starting vertices of the external branches that end in leaves i and j, respectively. Here an external branch is defined to consist only of the leaf i if there exists $j \in [N] \setminus \{i\}$ with $\rho(i,j) = 0$. (In fact, the finite setting discussed in this subsection can be seen as a special case of Section 2 in Chapter 2 as any semi-ultrametric ρ on $[N]$ can be extended to $\mathbb{N}$ by setting e. g. $\rho(1,k) = 0$ for $k > N$.)

We also use the decomposition of semi-ultrametrics in $\mathfrak{U}$ from Section 2 of Chapter 2. Here we have the continuous map

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

the space $\hat{\mathfrak{U}} = \{(r, v) \in \mathfrak{D} \times \mathbb{R}_+^{\mathbb{N}} : \alpha(r, v) \in \mathfrak{U}\} \subset \mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}}$ of marked distance matrices (or decomposed semi-ultrametrics), and the map

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

which maps a semi-ultrametric ρ to the sequence of the external branch lengths of the associated tree.

2.2 Metric measure spaces and marked metric measure spaces

For the theory of metric measure spaces and marked metric measure spaces, we refer to [24, 45, 48, 69, 91].

A metric measure space is a triple (X, ρ, μ) that consists of a complete and separable metric space (X, ρ) and a probability measure μ on the Borel sigma algebra on X. Two metric measure spaces (X, ρ, μ) and (X', ρ', μ') are defined to be isomorphic if there exists a measure-preserving isometry between the supports supp μ and supp μ'. The distance matrix distribution $\nu^{(X,\rho,\mu)}$ of a metric measure space (X, ρ, μ) is defined as the distribution of the random matrix $(\rho(x(i), x(j)))_{i,j\in\mathbb{N}}$, where $(x(i), i \in \mathbb{N})$ is a μ-iid sequence in X. The Gromov reconstruction theorem ([91, Theorem 4] and [48, Section $3\frac{1}{2}$]) states that metric measure spaces are isomorphic if and only if they have the same distance matrix distribution. Hence, we can define the isomorphy class $[\![X, \rho, \mu]\!]$ of a metric measure space (X, ρ, μ) such that

$$\mathbb{U} = \{[\![X', \rho', \mu']\!] : (X', \rho', \mu') \text{ ultrametric measure space}\}$$

is a set. For $\chi \in \mathbb{U}$, we denote the associated distance matrix distribution by ν^{χ}. A sequence $(\chi_k, k \in \mathbb{N})$ in $\mathbb{U}$ converges to $\chi \in \mathbb{U}$ in the Gromov-weak topology if and only if the distance matrix distributions ν^{χ_k} converge weakly to ν^{χ}. The Gromov-Prohorov metric d_{GP} induces the Gromov-weak topology and is complete and separable, as shown in [45]. As in [46], the elements of $\mathbb{U}$ can be considered as trees.

For $N \in \mathbb{N}$, we also work with the closed subspace

$$\mathbb{U}_N = \{[\![X, \rho, \mu]\!] : (X, \rho, \mu) \text{ ultrametric measure space such that } N\mu \text{ is integer-valued}\}$$

of $\mathbb{U}$ which can be interpreted as the space of semi-ultrametric spaces that contain N elements and that are endowed with the uniform probability measure. When we identify points $x, y \in X$ with $\rho(x, y) = 0$ in a semi-metric space (X, ρ) to obtain a metric space, we refer by x, y also to the corresponding element of the metric space, in slight abuse of notation. We define the isomorphy class of a semi-metric measure space as the isomorphy class of the metric measure space obtained by identifying the elements with distance zero. We define the function

$$\psi_N : \mathfrak{U}_N \to \mathbb{U}_N, \quad \rho \mapsto [\![[N], \rho, N^{-1} \sum_{i=1}^{N} \delta_i]\!]$$

which maps a semi-ultrametric to the isomorphy class of the associated semi-metric measure space with the uniform measure. It is clear that the map ψ_N is continuous, for formal proofs, cf. Remark 11.1 of Chapter 2 or Lemma 4.5 of Chapter 3.

For each $\chi \in \mathbb{U}_N$, there exists $\rho \in \mathfrak{U}_N$ with $\psi_N(\rho) = \chi$, and the N-distance matrix distribution $\nu^{N,\chi}$ is defined as the distribution of the random matrix $(\rho(x(i), x(j)))_{i,j\in[N]}$, where $x(1), \dots, x(N)$ are sampled from $[N]$ according to the uniform measure without replacement. (That is, $(x(i), i \in [N])$ is a uniform permutation of $[N]$.) For every $\mathfrak{U}_N$-valued random variable ρ' with distribution $\nu^{N,\chi}$, it holds $\psi_N(\rho') = \chi$ a.s. Hence, χ is uniquely determined by $\nu^{N,\chi}$, as in [46].

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 space $X \times \mathbb{R}_+$ which is endowed with the product metric $d((x, v), (x', v')) = r(x, x') \vee |v - v'|$. Two marked metric measure spaces (X, r, m) and (X', r', m') are defined to be isomorphic if there exists an isometry φ between the supports $\operatorname{supp} m(\cdot \times \mathbb{R}_+)$ and $\operatorname{supp} m'(\cdot \times \mathbb{R}_+)$ such that the isometry

$$\hat{\varphi} : \operatorname{supp} m \to \operatorname{supp} m', \quad (x, v) \mapsto (\varphi(x), v)$$

satisfies $\hat{\varphi}(m) = m'$. The marked distance matrix distribution $\nu^{(X,r,m)}$ 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 $X \times \mathbb{R}_+$. The Gromov reconstruction theorem for marked metric measure spaces (see Proposition 3.12 of Chapter 2, cf. [24, Theorem 1]) states that marked metric measure spaces are isomorphic if and only if they have the same marked distance matrix distribution. We can now define the isomorphy class $[\![X, r, m]\!]$ of a marked metric measure space (X, r, m) such that

$$\hat{\mathbb{U}} = \{[\![X', r', m']\!] : \ (X', r', m') \text{ marked metric measure space with } \nu^{(X',r',m')}(\hat{\mathfrak{U}}) = 1\}$$

is a set. As in Chapter 2, $\hat{\mathbb{U}}$ is the set of isomorphy classes of marked metric measure spaces that yield ultrametric spaces when marks in the support of the measure are added to the distances of the metric space, and the elements of $\hat{\mathbb{U}}$ can be viewed as trees.

We denote the marked distance matrix distribution associated with any $\chi \in \hat{\mathbb{U}}$ by ν^χ. Using the continuous map α from Subsection 2.1, we associate with χ the probability distribution $\alpha(\nu^\chi)$ on $\mathfrak{U}$ which is called in Chapter 2 the distance matrix distribution of χ. (We denote by $f(\mu)$ the image measure of a measure μ under a map f). A sequence $(\chi_k, k \in \mathbb{N})$ converges to χ in $\hat{\mathbb{U}}$ in the marked Gromov-weak topology if and only if the marked distance matrix distributions ν^{χ_k} converge weakly to ν^χ. The marked Gromov-weak topology is metrized by the marked Gromov-Prohorov metric d_{mGP} which is complete and separable, see [24].

By Proposition 3.3 of Chapter 2, each element of $\hat{\mathbb{U}}$ is uniquely characterized in $\hat{\mathbb{U}}$ by its distance matrix distribution. Hence, the set of distance matrix distributions of marked metric measure spaces, denoted by

$$\mathcal{U}^{\mathrm{erg}} := \{\alpha(\nu^\chi) : \chi \in \hat{\mathbb{U}}\},$$

is in one-to-one correspondence with $\hat{\mathbb{U}}$, and its elements can likewise be viewed as trees. We endow $\mathcal{U}^{\mathrm{erg}}$ with the Prohorov metric d_{P}. Then $\mathcal{U}^{\mathrm{erg}}$ is separable and by Corollary 3.25 of Chapter 2 complete.

We say a a marked metric measure (X, r, m) space supports only the zero mark if $m = \mu \otimes \delta_0$ for some probability measure μ on X. Clearly, this property depends only on the isomorphy class of the marked metric measure space. Also note that the distance matrix distribution of a marked metric measure space $(X, r, \mu \otimes \delta_0)$ that supports only the zero mark equals the distance matrix distribution of the metric measure space (X, r, μ).

For $N \in \mathbb{N}$, we also define the closed subspace

$$\hat{\mathbb{U}}_N = \{[\![X, r, m]\!] \in \hat{\mathbb{U}} : (X, r, m) \text{ marked metric measure space}, Nm \text{ integer-valued}\}$$

of $\hat{\mathbb{U}}$ that stands for finite marked metric measure spaces. To obtain marked metric measure spaces from marked distance matrices, we use the map

$$\hat{\psi}_N : \hat{\mathfrak{U}}_N \to \hat{\mathbb{U}}_N, (r, v) \mapsto [\![[N], r, N^{-1} \sum_{i=1}^{N} \delta_{(i,v(i))}]\!],$$

where we understand the isomorphy class of a marked semi-metric measure space as the isomorphy class of the marked metric measure space obtained by identifying elements of the metric space with distance zero. Clearly, the map $\hat{\psi}_N$ is continuous, a formal proof is given in Remark 11.1 of Chapter 2.

As in [25], the N-marked distance matrix distribution of $\chi \in \hat{\mathbb{U}}_N$ is defined as the distribution of $((r(x(i), x(j)))_{i,j\in[N]}, (v(x(i))_{i\in[N]})$, where (r, v) is any element of $\hat{\mathfrak{U}}_N$ with $\hat{\psi}_N(r, v) = \chi$, and $x(1), \ldots, x(N)$ is sampled uniformly from $[N]$ without replacement. Clearly, $\hat{\psi}_N(r', v') = \chi$ a. s. for any random variable (r', v') that has the marked distance matrix distribution of χ.

3 Invariance principles

In Subsection 3.1, we define for each $N \in \mathbb{N}$ the Cannings population model of population size N. From this construction, we read off tree-valued processes in Subsections 3.2 – 3.5.

3.1 The Cannings model and the process of the genealogical distances

The population model has discrete generations, enumerated by $\mathbb{N}_0$, and N individuals in each generation, labeled by $1, \ldots, N$. The dynamics is characterized by a probability measure Ξ^N on the subspace

$$\Delta^N = \{x \in \Delta : |x|_1 = 1, Nx(i) \in \mathbb{N}_0 \text{ for all } i \in \mathbb{N}\}$$

of the simplex

$$\Delta = \{x = (x(1), x(2), \ldots) : x(1) \geq x(2) \geq \ldots \geq 0, |x|_1 \leq 1\},$$

where we write $|x|_p = (\sum_{i\in\mathbb{N}} |x(i)|^p)^{1/p}$.

First, we sample a Ξ^N-iid sequence $(x_k^N, k \in \mathbb{N})$ in Δ^N. Then, conditionally given $(x_k^N, k \in \mathbb{N})$, let $(\pi_k^N, k \in \mathbb{N})$ be a sequence of independent random partitions of $[N]$ such that for each $k \in \mathbb{N}$, the partition π_k^N is uniformly distributed on the set of partitions of $[N]$ whose block sizes are given by (any reordering of) Nx_k^N. In each generation $k \in \mathbb{N}$ of the population model, we partition the individuals into families, saying that individuals are in the same family if their labels are in the same block of π_k^N. Each family draws its common ancestor in generation $k - 1$ uniformly without replacement. Tracing back the ancestral lineage, we denote by $A_j(k, i)$ the label of the ancestor in generation j of the individual i of generation k, for $j \in \mathbb{N}_0$ with $j \leq k$.

We are interested in the genealogical distances between the individuals in each generation. Given a distance matrix $\rho_0^N \in \mathfrak{U}_N$, we define $\rho_0^N(i,j)$ as the genealogical distance between the individuals i and j in generation 0, for $i,j \in [N]$. Then we define the genealogical distance $\rho_\ell^N(i,j)$ between individuals i and j in a later generation $\ell \in \mathbb{N}$ by

$$\rho_\ell^N(i,j) = \begin{cases} 2c_N(\ell - \max\{k = 0,\dots,\ell : A_k(\ell,i) = A_k(\ell,j)\}) & \text{if } A_0(\ell,i) = A_0(\ell,j) \\ 2c_N\ell + \rho_0^N(A_0(\ell,i), A_0(\ell,j)) & \text{else,} \end{cases}$$

where we choose the scaling factor

$$c_N = \int \sum_{i=1}^N x(i) \frac{Nx(i)-1}{N-1} \Xi^N(dx). \tag{3.1}$$

We always assume $c_N > 0$. Analogously to e. g. [46] and Chapter 2, the genealogical distance $\rho_\ell^N(i,j)$ between the individuals i and j in generation ℓ is, up to the scaling factor, the number of generations backwards until these individuals have the same ancestor if they have the same ancestor in generation 0, else $\rho_\ell^N(i,j)$ is given by the genealogical distance of their ancestors in generation 0. The quantity c_N is known as the pairwise coalescence probability, i. e. the probability that two individuals that are sampled uniformly without replacement from some generation $k \in \mathbb{N}$ have the same ancestor in generation $k-1$. Indeed, conditionally given x_k^N, the first sample is in any family with probability $x_k^N(i)$ when $Nx_k^N(i)$ is the size of that family. Conditionally given x_k^N and the first sample, the second sample is in the same family with probability $(Nx_k^N(i)-1)/(N-1)$.

In the next subsections, we state four invariance principles for processes that we read off from this population model. The invariance principles define the limit processes, but we recall the limit processes independently in Section 4. All our limit processes are characterized by the probability measures on the simplex Δ, we denote the space of these measures by $\mathcal{M}_1(\Delta)$. If $\Xi \in \mathcal{M}_1(\Delta)$ satisfies the condition

$$\Xi\{0\} > 0 \text{ or } \int |x|_1 \, |x|_2^{-2} \, \Xi(dx) = \infty, \tag{3.2}$$

then we speak of the dust-free case and we write $\Xi \in \mathcal{M}_{\text{nd}}$. The converse case is called the case with dust. We set $\mathcal{M}_{\text{dust}} = \mathcal{M}_1(\Delta) \setminus \mathcal{M}_{\text{nd}}$. For $\Xi \in \mathcal{M}_1(\Delta)$, we denote by Ξ_0 the measure on Δ with

$$\Xi = \Xi_0 + \Xi\{0\}\delta_0. \tag{3.3}$$

3.2 Processes with values in the space of metric measure spaces

Let $\chi_0^N \in \mathbb{U}_N$, and let ρ_0^N be a random variable with distribution ν^{N,χ_0^N} that is independent of the sequence $(\pi_k^N, k \in \mathbb{N})$ from Subsection 3.1. We define the process $(\rho_k^N, k \in \mathbb{N}_0)$ as in Subsection 3.1 from $(\pi_k^N, k \in \mathbb{N})$ and the initial state ρ_0^N. For $k \in \mathbb{N}_0$, we set

$$\chi_k^N = [\![[N], \rho_k^N, N^{-1} \sum_{i=1}^N \delta_i]\!] = \psi_N(\rho_k^N). \tag{3.4}$$

While ρ_k^N describes the genealogy of generation k as a leaf-labeled tree, the unlabeled tree is given by χ_k^N. We call the process $(\chi_k^N, k \in \mathbb{N}_0)$ a $\mathbb{U}_N$-valued Ξ^N-Cannings chain with initial state χ_0^N. We call this process also a tree-valued Ξ^N-Cannings chain.

Remark 3.1 (Markov property and transition kernel). We denote by $\mathcal{P}_N$ the set of partitions of $[N]$. For $\pi \in \mathcal{P}_N$ and $i \in [N]$, we denote by $\pi(i)$ the integer k such that the k-th block of π contains i when the blocks are ordered increasingly according to their respective smallest element. As in Chapter 2, we associate with each element π of $\mathcal{P}_N$ a transformation $\mathfrak{U}_N \to \mathfrak{U}_N$, which we also denote by π, by

$$\pi(\rho) = (\rho(\pi(i), \pi(j)))_{i,j\in[n]}. \tag{3.5}$$

We write $\underline{\underline{2}}_N = (\mathbf{1}\{i \neq j\})_{i,j\in[N]}$. The Markov property of $(\chi_k^N, k \in \mathbb{N}_0)$ follows as for each $k \in \mathbb{N}_0$, the distance matrix $\rho_{k+1}^N - c_N \underline{\underline{2}}_N$ has conditional distribution $\pi_{k+1}^N(\nu^{N,\chi_k^N})$ given π_{k+1}^N and $(\rho_j^N, j \leq k)$ by the construction in Section 3.1. We denote by p_N the transition kernel of $(\chi_k^N, k \in \mathbb{N}_0)$ which can be stated as

$$p_N(\chi, B) = \sum_{\pi\in\mathcal{P}_N} \mathbb{P}(\pi_1^N = \pi) \int \nu^{N,\chi}(d\rho)\, \mathbf{1}\Big\{\psi_N(\pi(\rho) + c_N \underline{\underline{2}}_N) \in B\Big\}$$

for all $\chi \in \mathbb{U}_N$ and measurable $B \subset \mathbb{U}$. This transition kernel generalizes the one of a tree-valued Moran model from [46] or a tree-valued Λ-Cannings process that is discussed in [61].

For $c \in \mathbb{R}_+$, we set $\Delta_c = \{x \in \Delta : x(1) > c\}$.

Theorem 3.2. *Let $\Xi \in \mathcal{M}_{\mathrm{nd}}$. Assume that χ_0^N converges to some χ_0 in $(\mathbb{U}, d_{\mathrm{GP}})$ as N tends to infinity. Furthermore, assume that*

$$\lim_{N\to\infty} c_N = 0 \tag{3.6}$$

and that

$$\begin{aligned}&\text{for arbitrarily small } \varepsilon > 0, \text{ on } \Delta_\varepsilon, \text{ the finite measures } c_N^{-1}\Xi^N(dx)\\ &\text{converge weakly to } |x|_2^{-2}\,\Xi(dx) \text{ as } N \text{ tends to infinity.}\end{aligned} \tag{3.7}$$

Then the processes $(\chi^N_{\lfloor c_N^{-1} t\rfloor}, t \in \mathbb{R}_+)$ converge in distribution to a $\mathbb{U}$-valued Ξ-Fleming-Viot process with initial state χ_0 in the space of càdlàg paths in $(\mathbb{U}, d_{\mathrm{GP}})$, endowed with the Skorohod metric, as N tends to infinity.

Condition (3.6) is Condition (1.6) in [85] and ensures that the limit process is a process in continuous time with no fixed jumps. Condition (3.7) is Condition (2.9) of [85] and yields the convergence of the transition kernels to the generator of the limit process. The assumption $\Xi \in \mathcal{M}_{\mathrm{nd}}$ in the theorem above is necessary as the $\mathbb{U}$-valued Ξ-Fleming-Viot process exists only for these Ξ. We include the case with dust in two ways: In Section 3.3, we consider the distance matrix distributions. In Section 3.4, we decompose the genealogical trees.

3.3 Processes of distance matrix distributions

We associate with the $\mathbb{U}_N$-valued Ξ^N-Canning chain from the last subsection the process

$$(\xi_k^N, k \in \mathbb{N}_0) = (\nu^{\chi_k^N}, k \in \mathbb{N}_0)$$

with values in the space $(\mathcal{U}^{\mathrm{erg}}, d_{\mathrm{P}})$ which we recalled in Section 2.2. Also $(\xi_k^N, k \in \mathbb{N}_0)$ is a Markov process. This follows from the Markov property of $(\chi_k^N, k \in \mathbb{N}_0)$ as ξ_k^N uniquely determines χ_k^N by the Gromov reconstruction theorem. As $(\xi_k^N, k \in \mathbb{N}_0)$ takes values in the space of distance matrix distributions

$$\mathcal{U}_N = \{\nu^\chi : \chi \in \mathbb{U}_N\} \subset \mathcal{U}^{\mathrm{erg}},$$

we call this process a $\mathcal{U}_N$-valued Ξ^N-Cannings chain. When we do not want to specify the state space, we call also this process a tree-valued Ξ^N-Cannings chain.

Theorem 3.3. *Let $\Xi \in \mathcal{M}_1(\Delta)$. Assume that conditions (3.6) and (3.7) hold, and that ξ_0^N converges weakly to some probability measure ξ_0 on $\mathfrak{U}$ as N tends to infinity. Then the processes $(\xi^N_{\lfloor c_N^{-1} t\rfloor}, t \in \mathbb{R}_+)$ converge in distribution to a $\mathcal{U}^{\mathrm{erg}}$-valued Ξ-Fleming-Viot process with initial state ξ_0 in the space of càdlàg paths in $(\mathcal{U}^{\mathrm{erg}}, d_{\mathrm{P}})$, endowed with the Skorohod metric, as N tends to infinity.*

Remark 3.4. Consider the case that $\Xi \in \mathcal{M}_{\mathrm{nd}}$ and that $\xi_0 = \nu^{\chi_0}$ for some $\chi_0 \in \mathbb{U}$. Let $(\chi_t, t \in \mathbb{R}_+)$ be a $\mathbb{U}$-valued Ξ-Fleming-Viot process with initial state χ_0. Then, by the definition in Section 8 of Chapter 2, the process $(\nu^{\chi_t}, t \in \mathbb{R}_+)$ is a $\mathcal{U}^{\mathrm{erg}}$-valued Ξ-Fleming-Viot process and the assertion of Theorem 3.3 follows from Theorem 3.2 as the map $\mathbb{U} \to \mathcal{U}^{\mathrm{erg}}$, $\chi \mapsto \nu^\chi$ is continuous.

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

Recall the decomposition $\beta : \mathfrak{U}_N \to \hat{\mathfrak{U}}_N$ of ultrametric distance matrices from Section 2.1 and the construction $\hat{\psi}_N : \hat{\mathfrak{U}}_N \to \hat{\mathbb{U}}_N$ of marked metric measure spaces from marked distance matrices from Section 2.2. We define a process $(\hat{\chi}_k^N, k \in \mathbb{N})$ which we call a $\hat{\mathbb{U}}_N$-valued Ξ^N-Cannings chain by

$$\hat{\chi}_k^N = \hat{\psi}_N \circ \beta(\rho_k^N)$$

for $k \in \mathbb{N}_0$, where $(\rho_k^N, k \in \mathbb{N}_0)$ is defined as in Section 3.2. More loosely, we also call the process $(\hat{\chi}_k^N, k \in \mathbb{N})$ a tree-valued Ξ^N-Cannings chain.

We denote by b_N the probability that an individual that is sampled uniformly from a fixed generation belongs to a family with more than one member. By construction,

$$b_N = \mathbb{E}\left[\sum_{i=1}^{\infty} x_1^N(i)\, \mathbf{1}\{x_1^N(i) > 1/N\}\right].$$

Theorem 3.5. *Let $\Xi \in \mathcal{M}_1(\Delta)$. Assume that conditions (3.6) and (3.7) hold, and that $\hat{\chi}_0^N$ converges to some $\hat{\chi}_0$ in $(\hat{\mathbb{U}}, d_{\mathrm{mGP}})$ as N tends to infinity.*

(i) If $\Xi \in \mathcal{M}_{\mathrm{dust}}$, assume in addition that

$$\lim_{N\to\infty} b_N/c_N = \int |x|_1 \, |x|_2^{-2} \, \Xi(dx). \tag{3.8}$$

(ii) If $\Xi \in \mathcal{M}_{\mathrm{nd}}$, assume in addition that the marked metric measure space $\hat{\chi}_0$ supports only the zero mark.

Then the processes $(\hat{\chi}^N_{\lfloor c_N^{-1} t\rfloor}, t \in \mathbb{R}_+)$ converge in distribution to a $\hat{\mathbb{U}}$-valued Ξ-Fleming-Viot process with initial state $\hat{\chi}_0$ in the space of càdlàg paths in $(\hat{\mathbb{U}}, d_{\mathrm{mGP}})$, endowed with the Skorohod metric, as N tends to infinity.

In the dust-free case, assumption (ii) is necessary for right continuity of the limit process at time 0. The expression on the right-hand side of (3.8) is the rate $\lambda_{1,\{\{1\}\}}$ in the martingale problem for the $\hat{\mathbb{U}}$-valued Ξ-Fleming-Viot process which we recall in Section 4.2. In the lookdown construction from Chapter 2, $\lambda_{1,\{\{1\}\}}$ is the rate of reproduction events in which the individual on a fixed level belongs to a non-singleton family, which we will use in Remark 3.9. The expression on the right-hand side of (3.8) also occurs as the parameter of the limiting exponential distribution of the length of an external branch that is drawn randomly from a Ξ-coalescent as the sample size tends to infinity, see [71].

We also decompose some metric measure spaces analogously to Section 2.1: We denote also by β the function $\mathbb{U}_N \to \hat{\mathbb{U}}_N$ that maps χ to $\psi_N \circ \beta(\rho)$ where ρ is any element of $\mathfrak{U}_N$ with $\psi_N(\rho) = \chi$. The map $\beta : \mathbb{U}_N \to \hat{\mathbb{U}}_N$ decomposes the (unlabeled) tree given by (the isomorphy class of) a metric measure space χ such that in $\beta(\chi)$, the lengths of the external branches are encoded by the marks, and the distances between their starting points are given by the metric. Note that $\hat{\chi}_k^N = \beta(\chi_k^N)$ for all $k \in \mathbb{N}_0$. The Markov property of the process $(\hat{\chi}_k^N, k \in \mathbb{N}_0)$ now follows from the Markov property of $(\chi_k^N, k \in \mathbb{N}_0)$ as χ_k^N is uniquely determined by $\beta(\chi_k^N)$.

3.5 Another decomposition of the evolving genealogical trees

In Theorem 3.5, the external branches of genealogical trees play a crucial role. However, for the proof in the case with dust, it seems more convenient to work with a different decomposition of the genealogical trees that does not bring about the freeing phenomenon which we mention below. We will therefore use Proposition 3.8 below in the proof of Theorem 3.5 in the case with dust. The process of these differently decomposed trees corresponds to the construction in Section 7 of Chapter 2 where an infinite population is considered.

First we define a process $((r_k^N, v_k^N), k \in \mathbb{N}_0)$ of marked distance matrices. Let $\tilde{\chi}_0^N \in \hat{\mathbb{U}}_N$, let (r_0^N, v_0^N) be a $\hat{\mathfrak{U}}_N$-valued random variable with distribution $\nu^{N,\tilde{\chi}_0^N}$, and let $\rho_0^N = \alpha(r_0^N, v_0^N)$, with α defined in Section 2.1. We define the process $(\rho_k^N, k \in \mathbb{N})$ as in Subsection 3.1 from the initial state ρ_0 and the sequence $(\pi_k^N, k \in \mathbb{N})$, assuming that $(\pi_k^N, k \in \mathbb{N})$ is independent of (r_0^N, v_0^N).

For $\ell \in \mathbb{N}$, we define (r_ℓ^N, v_ℓ^N) as follows. For $i \in [N]$, if there exists a latest generation $k \in [\ell]$ in which the individual $A_k(\ell, i)$ is in a non-singleton block of π_k^N, we set $v_\ell^N(i) = c_N(\ell - k + 1)$. Else, that is, if $A_k(\ell, i)$ forms a singleton block of π_k^N for each $k \in [\ell]$, then we set $v_\ell^N(i) = c_N \ell + v_0^N(A_0(\ell, i))$. For $i, j \in [N]$, we set

$$r_\ell^N(i,j) = (\rho_\ell^N(i,j) - v_\ell^N(i) - v_\ell^N(j)) \mathbf{1}\{i \neq j\}.$$

Lemma 3.6. *It holds $(r_k^N, v_k^N) \in \hat{\mathfrak{U}}_N$ for all $k \in \mathbb{N}_0$.*

Proof. For $k \in \mathbb{N}$, we denote by $\cup\sigma_k^N$ the union of the non-singleton blocks of π_k^N (anticipating the notation of Remark 3.7 below). By construction,

$$\begin{aligned} r_k^N(i,j) =& v_{k-1}^N(A_{k-1}(k,i)) \mathbf{1}\{i \in \cup\sigma_k^N\} + r_{k-1}^N(A_{k-1}(k,i), A_{k-1}(k,j)) \\ &+ v_{k-1}^N(A_{k-1}(k,j)) \mathbf{1}\{j \in \cup\sigma_k^N\} \end{aligned}$$

for all $k \in \mathbb{N}$ and all distinct $i, j \in [N]$. Under the assumption that r_{k-1}^N satisfies the triangle inequality, it is easily checked that also r_k^N satisfies the triangle inequality. The assertion follows by induction. □

Analogously to Remark 7.1 of Chapter 2, the quantity $v_k^N(i)$ can be interpreted as the age of the individual i of generation k. The quantity $v_k^N(i)$ needs not coincide with the (scaled) length of the external branch that ends in individual i in the (scaled) genealogical tree of generation k. Indeed, if $v_k(i) < k$ and all other members in the family of $A_{k-v_k^N(i)}(k,i)$ in generation $k - v_k^N(i)$ have no descendants in generation k, then this external branch is longer than $v_k^N(i)$. A related phenomenon in evolving coalescents is the so-called freeing where internal branches become part of external branches (see Dahmer and Kersting [22], in particular Figure 2 therein). The quantity $r_k^N(i,j)$ is the distance between the individuals that correspond to the parents in Chapter 3.

We obtain a $\hat{\mathbb{U}}_N$-valued process $(\tilde{\chi}_k^N, k \in \mathbb{N}_0)$ by setting $\tilde{\chi}_k^N = \hat{\psi}_N(r_k^N, v_k^N)$ for $k \in \mathbb{N}$. The definition of (r_0^N, v_0^N) yields $\tilde{\chi}_0^N = \hat{\psi}_N(r_0^N, v_0^N)$.

Remark 3.7 (Markov property and transition kernel). We denote by $\mathcal{S}_N$ the set of semi-partitions of $[N]$. A semi-partition σ of $[N]$ is a system of nonempty disjoint subsets of $[N]$ which we call blocks. We denote by $\cup\sigma$ the union of the blocks of σ (which can be a proper subset of $[N]$). For $\sigma \in \mathcal{S}_N$ and $i \in [N]$, we define $\sigma(i) = \pi(i)$ where π is the partition of $[n]$ that has the same non-singleton blocks as σ, that is, $\{B \in \sigma : \#B \geq 2\} = \{B \in \pi : \#B \geq 2\}$, and $\pi(i)$ is defined as in Remark 3.1. As in Chapter 2, we associate with each element σ of $\mathcal{S}_N$ a transformation $\hat{\mathfrak{U}}_N \to \hat{\mathfrak{U}}_N$, which we also denote by σ, by $\sigma(r, v) = (r', v')$, where

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

and

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

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

For $k \in \mathbb{N}$, let σ_k^N be the semi-partition that consists of the non-singleton blocks of the partition π_k^N from Section 3.1, $\sigma_k^N = \{B \in \pi_k^N : \#B \geq 2\}$. We write $\underline{1}_N = (1)_{i\in[N]}$. The Markov property of $(\tilde{\chi}_k^N, k \in \mathbb{N}_0)$ follows as by construction, for each $k \in \mathbb{N}_0$, the marked distance matrix $(r_{k+1}^N, v_{k+1}^N - c_N\underline{1}_N)$ has conditional distribution $\sigma_{k+1}^N(\nu^{N,\tilde{\chi}_k^N})$ given σ_{k+1}^N and $((r_j^N, v_j^N), j \leq k)$. We denote by $\tilde{p}_N$ the transition kernel of $(\tilde{\chi}_k^N, k \in \mathbb{N}_0)$ which can be stated as

$$\tilde{p}_N(\chi, B) = \sum_{\sigma\in\mathcal{S}_N} \mathbb{P}(\sigma_1^N = \sigma) \int \nu^{N,\chi}(d(r,v))\, \mathbf{1}\Big\{\hat{\psi}_N(\sigma(r,v) + c_N(0,\underline{1}_N)) \in B\Big\}$$

for all $\chi \in \hat{\mathbb{U}}_N$ and measurable $B \subset \hat{\mathbb{U}}$.

Proposition 3.8. *Assume that assumptions of Theorem 3.5 hold and that $\tilde{\chi}_0^N$ converges to $\hat{\chi}_0$ in $(\hat{\mathbb{U}}, d_{\mathrm{mGP}})$ as N tends to infinity. Then the processes $(\tilde{\chi}^N_{\lfloor c_N^{-1}t\rfloor}, t \in \mathbb{R}_+)$ converge in distribution to a $\hat{\mathbb{U}}$-valued Ξ-Fleming-Viot process with initial state $\hat{\chi}_0$ in the space of càdlàg paths in $(\hat{\mathbb{U}}, d_{\mathrm{mGP}})$, endowed with the Skorohod metric, as N tends to infinity.*

Remark 3.9. In Proposition 3.8, the assumption (3.8) is necessary in case $\Xi \in \mathcal{M}_{\mathrm{dust}}$. To see this, let $t \in (0,\infty)$, and let $v_t = (v_t(i), i \in \mathbb{N})$ be defined as in Section 7.2 of Chapter 2. By the lookdown construction in Chapter 2, the truncated first entry $v_t(1) \wedge t$ is equal in distribution to $T \wedge t$ for an exponentially distributed random variable T with parameter $\int |x|_1\, |x|_2^{-2}\, \Xi_0(dx)$. Moreover, by the construction in Subsection 3.1, the random variable $v^N_{\lfloor c_N^{-1}t\rfloor}(1) \wedge (c_N \lfloor c_N^{-1}t\rfloor)$ is distributed as $c_N(T_N \wedge \lfloor c_N^{-1}t\rfloor)$ for a geometrically distributed random variable T_N with parameter b_N. Let $(\hat{\chi}_s, s \in \mathbb{R}_+)$ be a $\hat{\mathbb{U}}$-valued Ξ-Fleming-Viot process with initial state $\hat{\chi}_0$ (as defined in Chapters 2 and 3) and let $\varpi : \hat{\mathfrak{U}} \to \mathbb{R}_+$, $(r,v) \mapsto v(1)$. Then, as in Proposition 4.7 of Chapter 2, the random variable $v_t(1)$ has distribution $\mathbb{E}[\varpi(\nu^{\hat{\chi}_t})]$. Moreover, as $v^N_{\lfloor c_N^{-1}t\rfloor}$ is exchangeable, the random variable $v^N_{\lfloor c_N^{-1}t\rfloor}(1)$ has distribution $\mathbb{E}[\varpi(\nu^{\tilde{\chi}^N_{\lfloor c_N^{-1}t\rfloor}})]$.

Now assume that $(\tilde{\chi}^N_{\lfloor c_N^{-1}s\rfloor}, s \in \mathbb{R}_+)$ converges as asserted in Proposition 3.8. Then, as the path $(\hat{\chi}_s, s \in \mathbb{R}_+)$ is a.s. continuous at fixed times, also $\tilde{\chi}^N_{\lfloor c_N^{-1}t\rfloor}$ converges in distribution to $\hat{\chi}_t$. By continuity of the map ϖ, also $v^N_{\lfloor c_N^{-1}t\rfloor}(1)$ converges in distribution to $v_t(1)$. As t can be chosen arbitrarily large, it follows that $c_N T_N$ converges in distribution to T which implies condition (3.8).

4 Tree-valued Fleming-Viot processes

In this section, we recall from Chapter 2 the martingale problems for tree-valued Ξ-Fleming-Viot processes. Path regularity is shown in Chapter 3. (In Chapters 2 and 3, tree-valued Ξ-Fleming-Viot processes are considered for finite measures Ξ on Δ.) For $n \in \mathbb{N}$, we define the restrictions

$$\gamma_n : \mathfrak{U} \cup \bigcup_{\ell\geq n} \mathfrak{U}_\ell \to \mathfrak{U}_n, \quad \rho \mapsto (\rho(i,j))_{i,j\in[n]}$$

and

$$\gamma_n : \hat{\mathfrak{U}} \cup \bigcup_{\ell \geq n} \hat{\mathfrak{U}}_\ell \to \hat{\mathfrak{U}}_n, \quad (r, v) \mapsto ((r(i,j))_{i,j \in [n]}, (v(i))_{i \in [n]}).$$

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 write also ϕ for the function $\phi \circ \gamma_n$. For $\phi \in \mathcal{C}_n$, we call the function $\mathbb{U} \to \mathbb{R}$, $\chi \mapsto \nu^\chi \phi$ the polynomial associated with ϕ. As in [69, Corollary 2.8], the algebra of polynomials

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

is convergence determining in $\mathbb{U}$.

Analogously, let $\hat{\mathcal{C}}_n$ be the set of bounded differentiable functions $\phi : \mathbb{R}^{n^2} \times \mathbb{R}^n \to \mathbb{R}$ with bounded uniformly continuous derivative. For $\phi \in \hat{\mathcal{C}}_n$, we denote also the function $\phi \circ \gamma_n$ by ϕ, and we associate with ϕ the marked polynomial $\hat{\mathbb{U}} \to \mathbb{R}$, $\chi \mapsto \nu^\chi \phi$. The algebra of marked polynomials

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

is convergence determining in $\hat{\mathbb{U}}$, see [69, Corollary 2.8]. The algebra

$$\mathscr{C} = \{\mathcal{U}^{\text{erg}} \to \mathbb{R}, \nu \mapsto \nu\phi : n \in \mathbb{N}, \phi \in \mathcal{C}_n\}$$

is convergence determining in $\mathcal{U}^{\text{erg}}$, see Remark 4.5 in Chapter 2.

4.1 The $\mathbb{U}$-valued Ξ-Fleming-Viot process

A $\mathbb{U}$-valued Ξ-Fleming-Viot process exists for $\Xi \in \mathcal{M}_{\text{nd}}$ and any initial state in $\mathbb{U}$. It is a Markov process and has a version with càdlàg paths and no discontinuities at fixed times. This process is the unique solution of the martingale problem (B, Π), we recall the generator B in this subsection. For $n \in \mathbb{N}$, let $\mathcal{P}_n$ be the set of partitions of $[n]$, and let κ_n^∞ be the probability kernel from Δ to $\mathcal{P}_n$ given by Kingman's correspondence, which we also recall here as we need it in Section 6.2. Consider independent uniform $[0,1]$-valued random variables $U_1, \dots, U_n$. For $x \in \Delta$, let $\kappa_n^\infty(x, \cdot)$ be the distribution of the random partition in $\mathcal{P}_n$ such that two different integers i and j are in the same block if and only if there exists $\ell \in \mathbb{N}$ with

$$U_i, U_j \in \left(\sum_{k=1}^{\ell-1} x(k), \sum_{k=1}^{\ell} x(k) \right).$$

For $\phi \in \mathcal{C}_n$, we define the function

$$\langle \nabla \phi, \underline{\underline{2}} \rangle : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}, \quad \rho \mapsto 2 \sum_{\substack{i,j \in \mathbb{N} \\ i \neq j}} \frac{\partial}{\partial \rho(i,j)} \phi(\rho).$$

Let $\mathbf{0}_n = \{\{1\},\ldots,\{n\}\}$, and recall Ξ_0 from equation (3.3). For $\pi \in \mathcal{P}_n \setminus \{\mathbf{0}_n\}$, we define

$$\lambda_\pi = \int \kappa_n^\infty(x,\pi)\,|x|_2^{-2}\,\Xi_0(dx) \\ + \Xi\{0\}\,\mathbf{1}\{\pi \text{ contains one doubleton and apart from that only singletons}\}\,.$$

The rates λ_π are equal to those of Schweinsberg [86], see Remark 6.1 of Chapter 2. Moreover, we associate with each element of $\mathcal{P}_n$ a transformation on $\mathfrak{U}_n$ as in Remark 3.1.

Now we set $B = B_{\text{grow}} + B_{\text{res}}$ with

$$B_{\text{grow}}\Phi(\chi) = \int \nu^\chi(d\rho)\langle\nabla\phi, \underline{\underline{2}}\rangle(\rho)$$

and

$$B_{\text{res}}\Phi(\chi) = \sum_{\pi\in\mathcal{P}_n\setminus\{\mathbf{0}_n\}} \lambda_\pi \int \nu^\chi(d\rho)(\phi(\pi(\gamma_n(\rho))) - \phi(\rho))$$

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

4.2 The $\hat{\mathbb{U}}$-valued Ξ-Fleming-Viot process

We first consider the case $\Xi \in \mathcal{M}_{\text{nd}}$. Using the isometry

$$\beta_0 : \mathbb{U} \to \hat{\mathbb{U}}, \quad [\![X,\rho,\mu]\!] \mapsto [\![X,\rho,\mu\otimes\delta_0]\!]$$

which adds a mark component that is concentrated in zero, we define for $\chi \in \mathbb{U}$ a $\hat{\mathbb{U}}$-valued Ξ-Fleming-Viot process with initial state $\beta_0(\chi)$ by $(\beta_0(\chi_t), t \in \mathbb{R}_+)$, where $(\chi_t, t \in \mathbb{R}_+)$ is a $\mathbb{U}$-valued Ξ-Fleming-Viot process with initial state χ and càdlàg paths. Also this process is Markov and has càdlàg paths with no discontinuities at fixed times.

Now we consider the case $\Xi \in \mathcal{M}_{\text{dust}}$. Then for each initial state in $\hat{\mathbb{U}}$, there exists a $\hat{\mathbb{U}}$-valued Ξ-Fleming-Viot process which is a Markov process and has a version with càdlàg paths and no discontinuities at fixed times. This process is the unique solution of the martingale problem $(\hat{B}, \hat{\Pi})$, we recall the generator $\hat{B}$ now. For $n \in \mathbb{N}$, we define the set $\mathcal{S}_n$ of semi-partitions of $[n]$ and the transformation on $\hat{\mathfrak{U}}_n$ associated with each element of $\mathcal{S}_n$ as in Remark 3.7.

Let K_n^∞ be the probability kernel from Δ to $\mathcal{S}_n$ defined as follows: Consider independent uniform $[0,1]$-valued random variables $U_1,\ldots,U_n$. For $x \in \Delta$, let $K_n^\infty(x,\cdot)$ be the distribution of the random element σ of $\mathcal{S}_n$ such that any two integers i and j are in a common subset $B \in \sigma$ if and only if there exists $\ell \in \mathbb{N}$ with

$$U_i, U_j \in \left(\sum_{k=1}^{\ell-1} x(k), \sum_{k=1}^{\ell} x(k)\right).$$

For $\phi \in \hat{\mathcal{C}}_n$, we define the function

$$\langle\nabla^v\phi, \underline{1}\rangle : \mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}} \to \mathbb{R}, \quad (r,v) \mapsto \sum_{i\in\mathbb{N}} \frac{\partial}{\partial v(i)}\phi(r,v).$$

For $\sigma \in \mathcal{S}_n \setminus \{\emptyset\}$, we define the rates

$$\lambda_{n,\sigma} = \int K_n^\infty(x,\sigma)\,|x|_2^{-2}\,\Xi_0(dx).$$

These rates are finite by the assumption that $\Xi \in \mathcal{M}_{\text{dust}}$ and as

$$K_n^\infty(x,\sigma) \leq K_1^\infty(x,\{\{1\}\}) = |x|_1$$

for all $x \in \Delta$. Now we set

$$\hat{B} = \hat{B}_{\text{grow}} + \hat{B}_{\text{res}},$$

$$\hat{B}_{\text{grow}}\Phi(\chi) = \int \nu^\chi(dr\,dv)\langle\nabla^v\phi,\underline{1}\rangle(r,v)$$

and

$$\hat{B}_{\text{res}}\Phi(\chi) = \sum_{\sigma\in\mathcal{S}_n\setminus\{\emptyset\}} \lambda_{n,\sigma}\int \nu^\chi(dr\,dv)(\phi(\sigma(\gamma_n(r,v))) - \phi(r,v))$$

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

4.3 The $\mathcal{U}^{\text{erg}}$-valued Ξ-Fleming-Viot process

A $\mathcal{U}^{\text{erg}}$-valued Ξ-Fleming-Viot process exists for every $\Xi \in \mathcal{M}_1(\Delta)$ and every initial state in $\mathcal{U}^{\text{erg}}$. It is a Markov process and has a version with càdlàg paths and no discontinuities at fixed times. This process is the unique solution of the martingale problem $(C, \mathscr{C})$ which is defined as follows. Let the rates λ_π for $\pi \in \mathcal{P}_n \setminus \{\mathbf{0}_n\}$, $n \in \mathbb{N}$ be defined from the measure $\Xi \in \mathcal{M}_1(\Delta)$ as in Subsection 4.1. We define the generator C by

$$C = C_{\text{grow}} + C_{\text{res}}$$

$$C_{\text{grow}}\Psi(\xi) = \int \xi(d\rho)\langle\nabla\phi,\underline{\underline{2}}\rangle(\rho)$$

$$C_{\text{res}}\Psi(\xi) = \sum_{\pi\in\mathcal{P}_n\setminus\{\mathbf{0}_n\}} \lambda_\pi \int \xi(d\rho)(\phi(\pi(\gamma_n(\rho))) - \phi(\rho))$$

for $\xi \in \mathcal{U}^{\text{erg}}$, $n \in \mathbb{N}$, $\phi \in \mathcal{C}_n$, and $\Psi \in \mathscr{C}$, $\Psi : \xi' \mapsto \xi'\phi$.

If $(\hat{\chi}_t, t \in \mathbb{R}_+)$ is a $\hat{\mathbb{U}}$-valued Ξ-Fleming-Viot process, then the process $(\alpha(\nu^{\hat{\chi}_t}), t \in \mathbb{R}_+)$ of the associated distance matrix distributions is a $\mathcal{U}^{\text{erg}}$-valued Ξ-Fleming-Viot process.

5 An example

In this section, we consider a sequence of Cannings models that does not satisfy the assertion of Theorem 3.5, but all its assumptions for $\Xi \in \mathcal{M}_{\text{dust}}$ except condition (3.8). The convergence of the $\hat{\mathbb{U}}_N$-valued Cannings chains to the $\hat{\mathbb{U}}$-valued Ξ-Fleming-Viot process in Theorem 3.5 is excluded as the length of a randomly sampled external branch converges to zero in distribution, while this quantity is a. s. non-zero at a fixed time for a $\hat{\mathbb{U}}$-valued

Fleming-Viot process in the case with dust. The convergence in Theorem 3.3 nevertheless holds. Here it comes to bear that the length of the first external branch is not a continuous functional of the infinite ultrametric that describes the genealogy. In the example of the present section, the length of a randomly sampled external branch converges to zero due to "perturbative" reproduction events that occur at high rate. However, these reproduction events are not visible in the limiting genealogy as each of them affects only a small part of the population.

We now work in the context of Section 3.1. We assume in this section that for N sufficiently large, the law Ξ^N of the sequence of the family sizes $x = (x(i), i \in \mathbb{N})$ in a reproduction event is given by the outcome of a two-step random experiment that can be described as follows. In the first step, we draw the type of the reproduction event from the set $\{\text{ordinary}, \text{perturbative}, \text{trivial}\}$. With probability N^{-1}, a reproduction event shall be ordinary, with probability $N^{-1/2}$ perturbative, and with probability $1 - N^{-1} - N^{-1/2}$ trivial. In the second step, the sequence x of the family sizes is drawn depending on the type of the reproduction event:

(i) If the reproduction event is ordinary, sample the largest family size $Nx(1)$ according to the binomial distribution with number of trials N and success probability $1/2$. That is, each individual belongs to the largest family independently with probability $1/2$. Let the other families be singletons, so that $Nx(i) \in \{0, 1\}$ for $i \geq 2$.

(ii) If the reproduction event is perturbative, sample the largest family size $Nx(1)$ according to the binomial distribution with parameters N and $N^{-1/3}$. Let the other families be singletons, so that $Nx(i) \in \{0, 1\}$ for $i \geq 2$.

(iii) If the reproduction event is trivial, let all families be singletons, so that $Nx(i) = 1$ for $i \in [N]$, and $Nx(i) = 0$ for $i > N$.

The pairwise coalescence probability for this reproduction law equals

$$c_N = N^{-1}(\tfrac{1}{2})^2 + N^{-1/2}(N^{-1/3})^2$$

and is asymptotically equivalent to $\frac{1}{4}N^{-1}$ as N tends to infinity. The probability b_N is bounded from below by

$$N^{-1/2}N^{-1/3}(1 - N^{-(1/3)\cdot(N-1)}).$$

This is the probability that a perturbative reproduction event occurs in which the individual labeled by 1 belongs to a non-singleton family. Condition (3.7) is satisfied with $\Xi = \delta_{(1/2,0,0,\ldots)} \in \mathcal{M}_{\text{dust}}$. As b_N/c_N tends to infinity as N tends to infinity, condition (3.8) is not satisfied.

In this section, we fix $t \in (0, \infty)$.

Proposition 5.1. *For $N \in \mathbb{N}$, let $(\hat{\chi}^N_k, k \in \mathbb{N}_0)$ be a $\hat{\mathbb{U}}$-valued Cannings chain defined as in Section 3.4 from the measure Ξ^N. Let $(\hat{\chi}_s, s \in \mathbb{R}_+)$ be a $\hat{\mathbb{U}}$-valued $\delta_{(1/2,0,0,\ldots)}$-Fleming-Viot process. Then $\hat{\chi}^N_{\lfloor c_N^{-1} t\rfloor}$ does not converge in distribution to $\hat{\chi}_t$ as N tends to infinity.*

Remark 5.2. Proposition 5.1 implies that the processes $(\hat{\chi}^N_{\lfloor c_N^{-1}s\rfloor}, s \in \mathbb{R}_+)$ do not converge in distribution in the Skorohod metric to $(\hat{\chi}_s, s \in \mathbb{R}_+)$ as N tends to infinity. This follows as $(\hat{\chi}_s, s \in \mathbb{R}_+)$ does a. s. not jump at fixed times.

To prove Proposition 5.1, we use the following lemma which states that for large N, in the leaf-labeled genealogical tree of generation $\lfloor c_N^{-1}t\rfloor$, the length of the external branch that ends in individual 1 is typically small.

Lemma 5.3. *Let $(\rho^N_k, k \in \mathbb{N}_0)$ be defined as in Section 3 from the measure Ξ^N. Let $\varepsilon \in (0, t)$ and*

$$v^N = (v^N(1), \dots, v^N(N)) = \Upsilon(\rho^N_{\lfloor c_N^{-1}t\rfloor}).$$

Then, $\mathbb{P}(v^N(1) > \varepsilon) < 4\varepsilon$ for sufficiently large N.

Proof of Proposition 5.1. Let $\varpi : \hat{\mathfrak{U}} \to \mathbb{R}_+$, $(r, v) \mapsto v(1)$. By exchangeability, the first entry $v^N(1)$ of the vector v^N in Lemma 5.3 has distribution $\mathbb{E}[\varpi(\nu^{\hat{\chi}^N_{\lfloor c_N^{-1}t\rfloor}})]$. By Lemma 5.3, the random variable $v^N(1)$ converges to zero in probability.

To obtain a contradiction, we assume that $\hat{\chi}^N_{\lfloor c_N^{-1}t\rfloor}$ converges in distribution to $\hat{\chi}_t$. Then as in Remark 3.9, also the probability distributions $\mathbb{E}[\varpi(\nu^{\hat{\chi}^N_{\lfloor c_N^{-1}t\rfloor}})]$ converge weakly to $\mathbb{E}[\varpi(\nu^{\hat{\chi}_t})]$. But as in Remark 3.9, the measure $\mathbb{E}[\varpi(\nu^{\hat{\chi}_t})]$ is not the Dirac measure in zero. □

Proof of Lemma 5.3. Let A_N be the number of ancestors in generation $\lfloor c_N^{-1}t\rfloor - \lfloor c_N^{-1}\varepsilon\rfloor$ of the individuals of generation $\lfloor c_N^{-1}t\rfloor$. Then,

$$\mathbb{P}(v^N(1) > \varepsilon, A_N \geq N/2) \leq (1 - N^{-1/2}N^{-1/3}(1 - (1 - N^{-1/3})^{N/2-1}))^{\lfloor c_N^{-1}\varepsilon\rfloor}. \tag{5.1}$$

We sketch a proof of the bound (5.1). The number of ancestors in generation k of the individuals of generation $\lfloor c_N^{-1}t\rfloor$ is non-decreasing in k for $k \leq \lfloor c_N^{-1}t\rfloor$. On the event $\{v^N(1) > \varepsilon, A_N \geq N/2\}$, no reproduction event in which the individual labeled by $A_k(\lfloor c_N^{-1}t\rfloor, 1)$ and another one of these ancestors are in the same block occurs in any generation k with $\lfloor c_N^{-1}t\rfloor - \lfloor c_N^{-1}\varepsilon\rfloor < k \leq \lfloor c_N^{-1}t\rfloor$. The right-hand side of inequality (5.1) is the probability that in none of the generations k with $\lfloor c_N^{-1}t\rfloor - \lfloor c_N^{-1}\varepsilon\rfloor < k \leq \lfloor c_N^{-1}t\rfloor$, the reproduction event is perturbative and individual $A_k(\lfloor c_N^{-1}t\rfloor, 1)$ is in the same family as any other of the at least $N/2$ many ancestors of the individuals of generation $\lfloor c_N^{-1}t\rfloor$.

For N sufficiently large, $\lfloor c_N^{-1}\varepsilon\rfloor \geq 3\varepsilon N$. As

$$\log(1 - N^{-1/3})^{N/2-1} \leq -(N/2 - 1)N^{-1/3} \to -\infty \quad (N \to \infty),$$

the right hand side of inequality (5.1) is bounded from above by $(1 - N^{-5/6}/2)^{3\varepsilon N}$ for N sufficiently large. This expression converges to zero as N tends to infinity.

It now suffices to show $\limsup_{N\to\infty} \mathbb{P}(A_N < N/2) < 4\varepsilon$. The event that no ordinary reproduction events occur between generations $\lfloor c_N^{-1}t\rfloor - \lfloor c_N^{-1}\varepsilon\rfloor$ and $\lfloor c_N^{-1}t\rfloor$ has probability at least $(1 - N^{-1})^{4\varepsilon N}$. Note that on this event, the random variable $N - A_N$ is stochastically bounded from above by $\sum_{i=1}^{X} Y_i$ where $X, Y_1, Y_2, \dots$ are independent random variables, X is binomially distributed with parameters $\lfloor 4\varepsilon N\rfloor$ and $N^{-1/2}$, and all

Y_i are binomially distributed with parameters N and $N^{-1/3}$. Here X corresponds to the number of perturbative reproduction events, and Y_i corresponds to the decrease in the number of ancestors, backwards in time, in the i-th of these reproduction events. Then we have

$$\mathbb{P}(A_N < N/2) \leq 1 - (1 - N^{-1})^{4\varepsilon N} + \mathbb{P}(\sum_{i=1}^{X} Y_i > N/2). \tag{5.2}$$

We estimate

$$\mathbb{P}(\sum_{i=1}^{X} Y_i > N/2) \leq \mathbb{P}(X > 8\varepsilon N^{1/2}) + 8\varepsilon N^{1/2}\mathbb{P}(Y_1 > (8\varepsilon N^{1/2})^{-1}N/2). \tag{5.3}$$

By the Chebychev inequality, the first summand is bounded from above by

$$(8\varepsilon N^{1/2} - 4\varepsilon N^{1/2})^{-2} 4\varepsilon N^{1/2},$$

and the second summand is bounded from above by

$$8\varepsilon N^{1/2}((8\varepsilon N^{1/2})^{-1}N/2 - N^{2/3})^{-2}N^{2/3}.$$

Hence, the expression on the right-hand side of (5.3) tends to zero, and the expression on the right-hand side of (5.2) converges to $1 - \mathrm{e}^{-4\varepsilon} < 4\varepsilon$ as $N \to \infty$. □

6 Convergence of the transition kernels

This section contains the proofs of Theorems 3.2 and 3.3, and the proof for the case $\Xi \in \mathcal{M}_{\mathrm{nd}}$ in Proposition 3.8. We write γ_n also for the restriction maps $\mathcal{S}_N \to \mathcal{S}_n$ and $\mathcal{P}_N \to \mathcal{P}_n$ for $n \leq N$ (that is, $\gamma_n(\sigma) = \{B \cap [n] : B \in \sigma\} \setminus \{\emptyset\}$). We define the rates λ_π and $\lambda_{n,\sigma}$ from the measure Ξ as in Section 4. We will need the following lemmas.

Lemma 6.1. *Let $\Xi \in \mathcal{M}_1(\Delta)$ and assume that condition (3.7) holds. Then,*

$$\lim_{N\to\infty} c_N^{-1}\mathbb{P}(\gamma_n(\pi_1^N) = \pi) = \lambda_\pi$$

for all $n \in \mathbb{N}$ and $\pi \in \mathcal{P}_n \setminus \{\mathbf{0}_n\}$.

Lemma 6.2. *Let $\Xi \in \mathcal{M}_{\mathrm{dust}}$. Assume that conditions (3.7) and (3.8) hold. Then,*

$$\lim_{N\to\infty} c_N^{-1}\mathbb{P}(\gamma_n(\sigma_1^N) = \sigma) = \lambda_{n,\sigma}$$

for all $n \in \mathbb{N}$ and $\sigma \in \mathcal{S}_n \setminus \{\emptyset\}$.

Remark 6.3. Note that the assumption of Lemma 6.1 is condition (2.9) in [85, Theorem 2.1], and that the limits in the assertion of Lemma 6.1 are the limits (16) in [72, Theorem 2.1]. In Subsection 6.2, we prove Lemmas 6.1 and 6.2 directly, using continuity properties in particular of the probability kernel associated with Kingman's correspondence.

6.1 Proofs of invariance principles

Proof of Theorem 3.2. Let $n \in \mathbb{N}$, $N \geq n$, and $\phi \in \mathcal{C}_n$. As in [46], we associate with ϕ not only the polynomial $\Phi : \mathbb{U} \to \mathbb{R}$, but also the N-polynomial

$$\Phi_N : \mathbb{U}_N \to \mathbb{R}, \quad \chi \mapsto \nu^{N,\chi}\phi.$$

As in Remark 3.1, let $\mathcal{P}_N$ denote the transition kernel of the Markov chain $(\chi_k^N, k \in \mathbb{N}_0)$. Then,

$$\begin{aligned} p_N\Phi_N(\chi_0^N) &= \mathbb{E}[\Phi_N(\chi_1^N)] = \mathbb{E}[\phi(\rho_1^N)] \\ &= \sum_{\pi\in\mathcal{P}_n\setminus\{\mathbf{0}_n\}} \mathbb{P}(\gamma_n(\pi_1^N) = \pi) \int \nu^{N,\chi_0^N}(d\rho)\phi(\pi(\gamma_n(\rho)) + c_N \underline{\underline{2}}_n). \end{aligned}$$

For the second equality, we use that ν^{N,χ_k^N} is the conditional distribution of ρ_1^N given χ_1^N which follows as ρ_1^N is exchangeable. By the construction in Section 3.1, the conditional distribution of $\gamma_n(\rho_1^N) - c_N\underline{\underline{2}}_n$ given $\gamma_n(\pi_1^N)$ equals $\gamma_n(\pi_1^N)(\gamma_n(\nu^{N,\chi_0^N}))$. This yields the third equality. Now we have

$$\begin{aligned} &c_N^{-1}(p_N - I)\Phi_N(\chi) \\ &= \mathbb{P}(\gamma_n(\pi_1^N) = \mathbf{0}_n) \int \nu^{N,\chi}(d\rho)c_N^{-1}(\phi(\rho + c_N\underline{\underline{2}}_N) - \phi(\rho)) \\ &\quad + \sum_{\pi\in\mathcal{P}_n\setminus\{\mathbf{0}_n\}} c_N^{-1}\mathbb{P}(\gamma_n(\pi_1^N) = \pi) \int \nu^{N,\chi}(d\rho)(\phi(\pi(\gamma_n(\rho)) + c_N\underline{\underline{2}}_n) - \phi(\rho)) \end{aligned} \tag{6.1}$$

for all $\chi \in \mathbb{U}_N$. As $c_N = \mathbb{P}(\gamma_2(\pi_1^N) = \{\{1, 2\}\})$, by exchangeability of π_1^N, and by assumption (3.6),

$$1 - \mathbb{P}(\gamma_n(\pi_1^N) = \mathbf{0}_n) \leq \binom{n}{2} c_N \to 0 \quad (N \to \infty).$$

As $\phi \in \mathcal{C}_n$,

$$\lim_{N\to\infty} \sup_{\rho\in\mathfrak{U}_N} \left| c_N^{-1}(\phi(\rho + c_N\underline{\underline{2}}_N) - \phi(\rho)) - \langle\nabla\phi, \underline{\underline{2}}\rangle(\rho)\right| = 0,$$

this follows from the mean value theorem and uniform continuity of the derivative. Furthermore, for every bounded measurable function $f : \mathbb{R}^{n^2} \to \mathbb{R}$,

$$\int \nu^{N,\chi}(d\rho)f(\gamma_n(\rho)) - \int \nu^{\chi}(d\rho)f(\gamma_n(\rho)) \leq 2\sup|f|\, n^2/N. \tag{6.2}$$

This follows as for every semi-metric measure space $\chi = [\![[N], \rho, N^{-1}\sum_{i=1}^N \delta_i]\!] \in \mathbb{U}_N$, we can couple $\gamma_n(\nu^{N,\chi})$ and $\gamma_n(\nu^\chi)$ by sampling n times from $[N]$ uniformly with replacement and accepting this as a sample without replacement on the event that no element of $[N]$ is drawn more than once. The probability of the complementary event is bounded from above by

$$1 - (1 - n/N)^n \leq n^2/N.$$

Taking $f(\rho) = \langle \nabla\phi, \underline{\underline{2}}\rangle(\rho)$, we obtain from the bound (6.2)

$$\begin{aligned}&\left|\int \nu^{N,\chi}(d\rho) c_N^{-1}(\phi(\rho + c_N \underline{\underline{2}}_N) - \phi(\rho)) - \int \nu^{\chi}(d\rho)\langle \nabla\phi, \underline{\underline{2}}\rangle(\rho)\right| \\ &\leq \int \nu^{N,\chi}(d\rho)\left|c_N^{-1}(\phi(\rho + c_N \underline{\underline{2}}_N) - \phi(\rho)) - \langle \nabla\phi, \underline{\underline{2}}\rangle(\rho)\right| + 2\sup|\langle \nabla\phi, \underline{\underline{2}}\rangle|\, n^2/N.\end{aligned}$$

With $f(\rho) = \phi(\pi(\rho)) - \phi(\rho)$, we obtain from (6.2)

$$\begin{aligned}&\left|\int \nu^{N,\chi}(d\rho)(\phi(\pi(\gamma_n(\rho)) + c_N \underline{\underline{2}}_N) - \phi(\rho)) - \int \nu^{\chi}(d\rho)(\phi(\pi(\gamma_n(\rho))) - \phi(\rho))\right| \\ &\leq \sup\{|\phi(\rho') - \phi(\rho)| : \rho, \rho' \in \mathfrak{U}_n, \|\rho - \rho'\| \leq c_N\} \\ &\quad + 4\sup|\phi|\, n^2/N.\end{aligned}$$

Using also Lemma 6.1, we now obtain from equation (6.1) and the definition of B in Section 4.1 the convergence

$$\lim_{N\to\infty} \sup_{\chi \in \mathbb{U}_N} \left|c_N^{-1}(p_N - I)\Phi_N(\chi) - B\Phi(\chi)\right| = 0.$$

The algebra Π of polynomials strongly separates points in $\mathbb{U}$ by [14, Lemma 4] and as Π generates the topology on $\mathbb{U}$. Let L denote the closure of Π for the supremum norm in the space of bounded continuous functions on $\mathbb{U}$. Analogously to Corollaries 9.3 and 9.4 in Chapter 2, the closure of B generates the semigroup on L of a $\mathbb{U}$-valued Ξ-Fleming-Viot process, which is strongly continuous. The assertion now follows from Corollary 4.8.9 in [36], condition (h) therein is satisfied. To see that the limit on the left-hand side of equation (8.48) on p. 234 of [36] equals zero, we set $f = \phi$ in equation (6.2). □

Proof of Theorem 3.3. Let $n \in \mathbb{N}$ and $N \geq n$. Let p'_N be the transition kernel of the Markov chain $(\xi_k^N, k \in \mathbb{N}_0)$. Let $\phi \in \mathcal{C}_n$,

$$\Psi : \mathcal{U}^{\text{erg}} \to \mathbb{R}, \quad \xi \mapsto \xi\phi,$$

and

$$\Psi_N : \mathcal{U}_N \to \mathbb{R}, \quad \nu^{\chi} \mapsto \nu^{N,\chi}\phi.$$

Then, as in the proof of Theorem 3.2,

$$p'_N \Psi_N(\xi_0^N) = \mathbb{E}[\Psi_N(\xi_1^N)] = \mathbb{E}[\nu^{N,\chi_1^N}\phi] = \mathbb{E}[\phi(\rho_1^N)]$$

and

$$\begin{aligned}&c_N^{-1}(p_N - I)\Psi_N(\nu^{\chi}) \\ &= \mathbb{P}(\gamma_n(\pi_1^N) = \mathbf{0}_n)\int \nu^{N,\chi}(d\rho) c_N^{-1}(\phi(\rho + c_N \underline{\underline{2}}_N) - \phi(\rho)) \\ &\quad + \sum_{\pi \in \mathcal{P}_n \setminus \{\mathbf{0}_n\}} c_N^{-1}\mathbb{P}(\gamma_n(\pi_1^N) = \pi)\int \nu^{N,\chi}(d\rho)(\phi(\pi(\gamma_n(\rho)) + c_N \underline{\underline{2}}_n) - \phi(\rho))\end{aligned}$$

for all $\chi \in \mathbb{U}_N$. As in the proof of Theorem 3.2, we have

$$\lim_{N\to\infty} \sup_{\chi\in\mathbb{U}_N} \left| c_N^{-1}(p_N - I)\Psi_N(\nu^\chi) - C\Psi(\nu^\chi) \right| = 0.$$

The algebra $\mathscr{C}$ strongly separates points in $\mathcal{U}^{\text{erg}}$ by [14, Lemma 4] and as $\mathscr{C}$ generates the weak topology on $\mathcal{U}^{\text{erg}}$. We conclude in the same way as in the proof of Theorem 3.2, applying [36, Corollary 4.8.9] and the analogs of Corollaries 9.3 and 9.4 in Chapter 2 for $\mathcal{U}^{\text{erg}}$-valued Ξ-Fleming-Viot processes. □

Proof of Proposition 3.8 (beginning). In the the first part of the proof, we assume $\Xi \in \mathcal{M}_{\text{dust}}$. Let $n \in \mathbb{N}$, $N \geq n$, $\phi \in \hat{\mathcal{C}}_n$. As in [25], we associate with ϕ not only the marked polynomial Φ but also the marked N-polynomial

$$\Phi_N : \hat{\mathbb{U}}_N \to \mathbb{R}, \quad \chi \mapsto \nu^{N,\chi}\phi.$$

Let $\tilde{p}_N$ be the transition kernel of the Markov chain $(\tilde{\chi}_k^N, k \in \mathbb{N}_0)$. Similarly to the proof of Theorem 3.2, we have

$$\tilde{p}_N\Phi_N(\tilde{\chi}_k^N) = \sum_{\sigma\in\mathcal{S}_n} \mathbb{P}(\gamma_n(\sigma_1^N) = \sigma) \int \nu^{N,\chi}(dr\, dv)\phi(\sigma(\gamma_n(r,v)) + c_N(0, \underline{1}_n)).$$

We obtain

$$\begin{aligned}
&c_N^{-1}(\tilde{p}_N - I)\Phi_N(\chi) \\
&= \mathbb{P}(\gamma_n(\sigma_1^N) = \emptyset) \int \nu^{N,\chi}(dr\, dv) c_N^{-1}(\phi(r, v + c_N\underline{1}_N)) - \phi(r,v)) \\
&\quad + \sum_{\sigma\in\mathcal{S}_n\setminus\{\emptyset\}} c_N^{-1}\mathbb{P}(\gamma_n(\sigma_1^N) = \sigma) \int \nu^{N,\chi}(dr\, dv)(\phi(\sigma(\gamma_n(r,v)) + c_N(0, \underline{1}_n)) - \phi(r,v))
\end{aligned}$$

for all $\chi \in \hat{\mathbb{U}}_N$. By condition (3.8) and as $\Xi \in \mathcal{M}_{\text{dust}}$,

$$\lim_{N\to\infty} c_N^{-1}\mathbb{P}(\gamma_1(\sigma_1^N) = \{\{1\}\}) = \lambda_{1,\{\{1\}\}} < \infty.$$

Here we can use Lemma 6.2 or, more simply, that $b_N = \mathbb{P}(\gamma_1(\sigma_1^N) = \{\{1\}\})$ by exchangeability. Exchangeability and the assumption $\lim_{N\to\infty} c_N = 0$ now imply

$$\mathbb{P}(\gamma_n(\sigma_1^N) = \emptyset) \geq 1 - n\mathbb{P}(\gamma_1(\sigma_1^N) = \{\{1\}\}) \to 1 \quad (N \to \infty).$$

Under our assumption that $\Xi \in \mathcal{M}_{\text{dust}}$, we conclude analogously to the proof of Theorem 3.2. Here we use Lemma 6.2 from the present chapter, Corollaries 9.3 and 9.4 from Chapter 2, and we apply Lemma 4 of [14] to $\hat{\Pi}$. □

6.2 Proofs of Lemmas 6.1 and 6.2

For $N \in \mathbb{N}$ and $n \in [N]$, we define a probability kernel κ_n^N from Δ^N to $\mathcal{P}_n$, and a probability kernel K_n^N from Δ^N to $\mathcal{S}_n$.

For $x \in \Delta^N$, let $\kappa_N^N(x, \cdot)$ be the uniform distribution on those partitions in $\mathcal{P}_N$ whose block sizes are given by (any reordering of) the nonzero elements of the sequence $(Nx(\ell), \ell \in \mathbb{N})$. Then we define $\kappa_n^N(x, \cdot)$ as the restriction $\kappa_n^N(x, \cdot) = \gamma_n(\kappa_N^N(x, \cdot))$.

The distribution of $\kappa_n^N(x, \cdot)$ can also be described by the following urn scheme: Consider an urn that contains $Nx(\ell)$ balls of color ℓ for each $\ell \in \mathbb{N}$. Sample n balls without replacement. Then the random partition of $[n]$ where $i, j \in [n]$ are in the same block if and only if the i-th and j-th ball have the same color has distribution $\kappa_n^N(x, \cdot)$. If we modify this urn scheme such that the balls are sampled with replacement, then we obtain the distribution $\kappa_n^\infty(x, \cdot)$, as a comparison with the definition of $\kappa_n^\infty(x, \cdot)$ in Section 4.1 shows.

For $x \in \Delta^N$, let $(y(1), y(2), \ldots)$ be the (possibly empty) finite subsequence of $(Nx(i), i \in \mathbb{N})$ that consists of the elements that are greater or equal to 2. Let $K_N^N(x, \cdot)$ be the uniform distribution on those elements of $\mathcal{S}_N$ that consist of disjoint subsets of $[N]$ whose sizes are given by (an arbitrary reordering of) $(y(1), y(2), \ldots)$. If $(Nx(i), i \in \mathbb{N})$ contains no elements that are greater or equal to 2, then $K_N^N(x, \cdot)$ is the Dirac measure in $\emptyset \in \mathcal{S}_N$. We set $K_n^N(x, \cdot) = \gamma_n(K_N^N(x, \cdot))$.

In other words, consider an urn that contains $Nx(\ell)$ balls of color ℓ for each $\ell \in \mathbb{N}$. Recolor each ball whose color occurs only once with a new color 0. Then sample n balls without replacement. Consider the random element σ of $\mathcal{S}_n$ where any $i, j \in [n]$ are in a common block if and only if the i-th and j-th ball have the same color that is not 0. Then σ has distribution $K_n^N(x, \cdot)$. When we sample with replacement, this distribution is $K_n^\infty(x, \cdot)$ instead, as a comparison with the definition of $K_n^\infty(x, \cdot)$ in Section 4.2 shows.

In the following lemma and its proof, we recall and slightly extend some well-known continuity properties from e. g. Proposition 2.9 in Bertoin [7]. We endow $\mathbb{N}$ with the discrete topology and consider the one-point compactification $\bar{\mathbb{N}} = \mathbb{N} \cup \{\infty\}$. Let $\bar{\mathbb{N}} = \mathbb{N} \cup \{\infty\}$ be the one-point compactification of the space $\mathbb{N}$, endowed with the discrete topology. We write $\Delta^\infty = \Delta$ and for $\varepsilon > 0$, we define

$$S_\varepsilon = \{(x, N) \in \Delta_\varepsilon \times \bar{\mathbb{N}} : x \in \Delta^N\}.$$

Lemma 6.4. *Let $n \in \mathbb{N}$, $\pi \in \mathcal{P}_n$, $\sigma \in \mathcal{S}_n$, and $\varepsilon > 0$. Then the maps*

$$S_\varepsilon \to [0,1], \quad (x, N) \mapsto \kappa_n^N(x, \pi)$$

and

$$S_\varepsilon \to [0,1], \quad (x, N) \mapsto K_n^N(x, \sigma)$$

are continuous.

Proof. Let $N \geq n$ and $x \in \Delta^N$. We couple the probability distributions $\kappa_n^\infty(x, \cdot)$ and $\kappa_n^N(x, \cdot)$ by starting with the urn scheme for $\kappa_n^\infty(x, \infty)$ given above and accepting the sample with replacement as a sample without replacement on the event that all sampled balls are different. The probability of this event is bounded from below by $(1-n/N)^n$, and we obtain the bound

$$|\kappa_n^N(x,\pi)-\kappa_n^\infty(x,\pi)| \le 1-(1-n/N)^n \le n^2/N.$$

Using the analogous coupling of the probability distributions $K_n^\infty(x,\cdot)$ and $K_n^N(x,\cdot)$, we obtain the bound

$$|K_n^N(x,\sigma)-K_n^\infty(x,\sigma)| \le 1-(1-n/N)^n \le n^2/N.$$

Furthermore, for $x,y\in\Delta$, we can couple the probability distributions $\kappa_n^\infty(x,\cdot)$ and $\kappa_n^\infty(y,\cdot)$ by using the same uniform random variables in Kingman's correspondence which we recalled in Section 4.1. This yields

$$|\kappa_n^\infty(x,\pi)-\kappa_n^\infty(y,\pi)| \le n|x-y|_1.$$

Similarly, using the definition of K_n^∞ from Section 4.2, we obtain

$$|K_n^\infty(x,\sigma)-K_n^\infty(y,\sigma)| \le n|x-y|_1.$$

W. l. o. g., let $((x_k,N_k),k\in\mathbb{N})$ be a sequence in S_ε that converges to some $(x,\infty)\in S_\varepsilon$. Then,

$$\begin{aligned}
&|K_n^{N_k}(x_k,\sigma)-K_n^\infty(x,\sigma)|\\
&\le |K_n^{N_k}(x_k,\sigma)-K_n^\infty(x_k,\sigma)|+|K_n^\infty(x_k,\sigma)-K_n^\infty(x,\sigma)|\\
&\le n^2/N_k+n|x_k-x|_1,
\end{aligned}$$

and the right-hand side converges to zero as k tends to infinity. The argument for $|\kappa_n^{N_k}(x_k,\pi)-\kappa_n^\infty(x,\pi)|$ is the same. □

Proof of Lemma 6.1. Let $n\in\mathbb{N}$. By construction, we have

$$\mathbb{P}(\gamma_n(\pi_1^N)=\pi)=\int \kappa_n^N(x,\pi)\Xi^N(dx)$$

for all $\pi\in\mathcal{P}_n$ and $N\ge n$. For arbitrarily small $\varepsilon>0$, assumption (3.7) implies the weak convergence

$$c_N^{-1}\Xi^N(dx)\delta_N(dN') \xrightarrow{w} |x|_2^{-2}\,\Xi(dx)\delta_\infty(dN') \quad \text{on } \Delta_\varepsilon\times\bar{\mathbb{N}} \quad (N\to\infty)$$

which also holds on S_ε as none of these measures have mass on the complement of S_ε in $\Delta_\varepsilon\times\bar{\mathbb{N}}$. Using Lemma 6.4, we obtain for each $\pi\in\mathcal{P}_n$

$$\begin{aligned}
&\lim_{N\to\infty} c_N^{-1}\int_{\Delta_\varepsilon}\kappa_n^N(x,\pi)\Xi^N(dx)=\lim_{N\to\infty}\int_{S_\varepsilon}\kappa_n^{N'}(x,\pi)c_N^{-1}\Xi^N(dx)\delta_N(dN')\\
&=\int_{S_\varepsilon}\kappa_n^{N'}(x,\pi)\,|x|_2^{-2}\,\Xi(dx)\delta_\infty(dN')=\int_{\Delta_\varepsilon}\kappa_n^\infty(x,\pi)\,|x|_2^{-2}\,\Xi(dx).
\end{aligned}\tag{6.3}$$

For every sufficiently small $\varepsilon > 0$, for $N \geq n$, and $x \in \Delta^N \setminus \Delta_\varepsilon$, the urn scheme for the probability distribution $\kappa_n^N(x,\cdot)$ yields that

$$\begin{aligned}
&\kappa_2^N(x,\{\{1,2\}\})(1-n\varepsilon)^n \\
&\leq \kappa_2^N(x,\{\{1,2\}\})\frac{N-\varepsilon N}{N-2}\cdots\frac{N-(n-2)\varepsilon N}{N-(n-1)} \\
&\leq \kappa_n^N(x,\{\{1,2\},\{3\},\ldots,\{n\}\}) \\
&\leq \kappa_2^N(x,\{\{1,2\}\}).
\end{aligned} \tag{6.4}$$

Indeed, we can draw a partition of $\{1,\ldots,n\}$ according to the distribution $\kappa_n^N(x,\cdot)$ by drawing n times without replacement from N individuals that are subdivided into families of sizes $(Nx(\ell),\ell\in\mathbb{N})$, and letting i and j be in the same block if and only if the i-th and j-th drawn individual are in the same family. The second inequality follows as the family sizes are at most εN by our assumption on x.

For $x \in \Delta \setminus \Delta_\varepsilon$ and every partition $\pi \in \mathcal{P}_n$ that contains one doubleton and apart from that only singletons, we have

$$\kappa_n^N(x,\pi) = \kappa_n^N(x,\{\{1,2\},\{3\},\ldots,\{n\}\}) \tag{6.5}$$

by exchangeability. Moreover,

$$c_N = \int \kappa_2^N(x,\{\{1,2\}\})\Xi^N(dx) \tag{6.6}$$

by exchangeability.

Consequently, for arbitrarily small $\varepsilon > 0$ as in condition (3.7), and any partition $\pi \in \mathcal{P}_n$ that contains one doubleton and apart from that only singletons,

$$\begin{aligned}
&\limsup_{N\to\infty} c_N^{-1}\int_{\Delta\setminus\Delta_\varepsilon}\kappa_n^N(x,\pi)\Xi^N(dx) \leq \limsup_{N\to\infty} c_N^{-1}\int_{\Delta\setminus\Delta_\varepsilon}\kappa_2^N(x,\{\{1,2\}\})\Xi^N(dx) \\
&= 1-\liminf_{N\to\infty} c_N^{-1}\int_{\Delta_\varepsilon}\kappa_2^N(x,\{\{1,2\}\})\Xi^N(dx) = 1-\Xi(\Delta_\varepsilon),
\end{aligned}$$

where we use (6.4) and (6.5) for the first, equation (6.6) for the second, and the convergence (6.3) and $\kappa_2^\infty(x,\{\{1,2\}\}) = |x|_2^2$ for the third step. Analogously, we obtain for the same partitions π that

$$\liminf_{N\to\infty} c_N^{-1}\int_{\Delta\setminus\Delta_\varepsilon}\kappa_n^N(x,\pi)\Xi^N(dx) \geq (1-n\varepsilon)^n(1-\Xi(\Delta_\varepsilon)).$$

For every partition $\pi \in \mathcal{P}_n$ that contains a block of size greater than 2, $N \geq n$, and $x \in \Delta^N \setminus \Delta_\varepsilon$ we have

$$\kappa_n^N(x,\pi) \leq \kappa_3^N(x,\{\{1,2,3\}\}) \leq \kappa_2^N(x,\{\{1,2\}\})\frac{N\varepsilon-2}{N-2}$$

which can be seen from the urn scheme for $\kappa_n^N(x,\cdot)$ and the exchangeability therein. We obtain for these partitions π that

$$\limsup_{N\to\infty} c_N^{-1}\int_{\Delta\setminus\Delta_\varepsilon}\kappa_n^N(x,\pi)\Xi^N(dx)\le\varepsilon(1-\Xi(\Delta_\varepsilon)).$$

Similarly, for every partition $\pi\in\mathcal{P}_n$ that contains more than one non-singleton block, $N\ge n$, and $x\in\Delta^N\setminus\Delta_\varepsilon$, we have

$$\kappa_n^N(x,\pi)\le\kappa_4^N(x,\{\{1,2\},\{3,4\}\})\le\kappa_2^N(x,\{\{1,2\}\})\frac{N\varepsilon-1}{N-3},$$

hence

$$\limsup_{N\to\infty} c_N^{-1}\int_{\Delta\setminus\Delta_\varepsilon}\kappa_n^N(x,\pi)\Xi^N(dx)\le\varepsilon(1-\Xi(\Delta_\varepsilon)).$$

For $\pi\in\mathcal{P}_n\setminus\{\mathbf{0}_n\}$, we write

$$c_N^{-1}\mathbb{P}(\gamma_n(\pi_1^N)=\pi)=c_N^{-1}\int_{\Delta_\varepsilon}\kappa_n^N(x,\pi)\Xi^N(dx)+c_N^{-1}\int_{\Delta\setminus\Delta_\varepsilon}\kappa_n^N(x,\pi)\Xi^N(dx).$$

Now we let first $N\to\infty$. For the first integral, we use the convergence (6.3). For the second integral, we use the bounds for the limes superior and if necessary also for the limes inferior. Then we let ε tend to zero along a sequence such that for each element of this sequence, the weak convergence in condition (3.7) holds. Thus we obtain convergence to the rates λ_π as asserted in Lemma 6.1. □

Proof of Lemma 6.2. The proof is analogous to Lemma 6.1. By construction, we have

$$\mathbb{P}(\gamma_n(\sigma_1^N)=\sigma)=\int K_n^N(x,\sigma)\Xi^N(dx)$$

for all $\sigma\in\mathcal{S}_n$ and $N\ge n$. Analogously to the convergence (6.3), we obtain from Lemma 6.4 and assumption (3.7) that

$$\lim_{N\to\infty} c_N^{-1}\int_{\Delta_\varepsilon}K_n^N(x,\sigma)\Xi^N(dx)=\int_{\Delta_\varepsilon}K_n^\infty(x,\sigma)\,|x|_2^{-2}\,\Xi(dx)\qquad(6.7)$$

for all $\sigma\in\mathcal{S}_n$ and arbitrarily small $\varepsilon>0$.

For every $\sigma\in\mathcal{S}_n\setminus\{\emptyset\}$, $N\ge n$, and $x\in\Delta^N$, it holds

$$K_n^N(x,\sigma)\le K_1^N(x,\{\{1\}\}),$$

by exchangeability. Furthermore,

$$b_N=\int K_1^N(x,\{\{1\}\})\Xi^N(dx).$$

Now we obtain for these σ and arbitrarily small $\varepsilon > 0$

$$\begin{aligned}&\limsup_{N\to\infty} c_N^{-1}\int_{\Delta\setminus\Delta_\varepsilon} K_n^N(x,\sigma)\Xi^N(dx) \le \limsup_{N\to\infty} c_N^{-1}\int_{\Delta\setminus\Delta_\varepsilon} K_1^N(x,\{\{1\}\})\Xi^N(dx)\\ &= \lim_{N\to\infty} b_N/c_N - \lim_{N\to\infty} c_N^{-1}\int_{\Delta_\varepsilon} K_1^N(x,\{\{1\}\})\Xi^N(dx) = \int_{\Delta\setminus\Delta_\varepsilon} |x|_1\,|x|_2^{-2}\,\Xi(dx), \qquad (6.8)\end{aligned}$$

where we use assumption (3.8), the convergence (6.7), and $K_1^\infty(x,\{\{1\}\}) = |x|_1$ in the last equality. We conclude by combining the convergence (6.7) and the bound (6.8). We let ε tend to zero by dominated convergence, using that the integrands on the right-hand sides of (6.7) and (6.8) are bounded by $|x|_1|x|_2^{-2}$, and that

$$\int_\Delta |x|_1\,|x|_2^{-2}\,\Xi(dx) < \infty.$$

This yields the assertion. □

7 Convergence of marked metric measure spaces in the dust-free case

In this section, we prove Theorem 3.5 and Proposition 3.8 for $\Xi \in \mathcal{M}_{\mathrm{nd}}$. We use the isometric embedding $\beta_0 : \mathbb{U}_N \to \hat{\mathbb{U}}_N$, $[\![X,\rho,\mu]\!] \mapsto [\![X,\rho,\mu\otimes\delta_0]\!]$ which maps a metric measure space to the associated marked metric measure space that supports only the zero mark. We compare the process $(\beta_0(\chi^N_{\lfloor c_N^{-1}t\rfloor}), t\in\mathbb{R}_+)$ to the processes $(\hat{\chi}^N_{\lfloor c_N^{-1}t\rfloor}, t\in\mathbb{R}_+)$ and $(\tilde{\chi}^N_{\lfloor c_N^{-1}t\rfloor}, t\in\mathbb{R}_+)$.

Recall from Section 3.4 the map $\beta : \mathbb{U}_N \to \hat{\mathbb{U}}_N$ which decomposes a tree at the external branches. Also recall from Section 2.1 the functions α and Υ. We denote also by α the function from $\hat{\mathbb{U}}_N$ to $\mathbb{U}_N$ that maps χ to $\psi_N(\alpha(r,v))$, where (r,v) is any element of $\hat{\mathfrak{U}}_N$ with $\hat{\psi}_N(r,v) = \chi$. The function $\alpha : \hat{\mathbb{U}}_N \to \mathbb{U}_N$ retrieves a metric measure space from a marked metric measure space by adding the marks to the metric distances. For $\ell \ge 2$, we define the map

$$\Upsilon_1^\ell : \mathfrak{U} \to \mathbb{R}_+, \quad \Upsilon_1^\ell = \gamma_1\circ\Upsilon\circ\gamma_\ell.$$

In the subtree spanned by the first ℓ leaves of the tree associated with some $\rho\in\mathfrak{U}$, the length of the external branch that ends in the first leaf is given by $\Upsilon_1^\ell(\rho)$. We also define the restriction $\varpi : \hat{\mathfrak{U}} \to \mathbb{R}_+$, $(r,v)\mapsto v(1)$.

We recall that the Prohorov metric d_{P} on the space of probability measures on the Borel sigma algebra of a metric space (S,d) is given by

$$d_{\mathrm{P}}(\mu,\mu') = \inf\{\varepsilon>0 : \mu'(F) \le \mu(F^\varepsilon)+\varepsilon \text{ for all closed } F\subset S\}, \qquad (7.1)$$

where $F^\varepsilon = \{x\in S : d(x,F) < \varepsilon\}$. If (S,d) is separable, then we also have the coupling characterization

$$d_{\mathrm{P}}(\mu,\mu') = \inf_\nu \inf\{\varepsilon>0 : \nu\{(x,y)\in S^2 : d(x,y)>\varepsilon\} < \varepsilon\} \qquad (7.2)$$

where the first infimum is over all couplings ν of μ and μ', see e. g. [36, Theorem 3.1.2].

Lemma 7.1. *Let $\chi \in \hat{\mathbb{U}}_N$. Then,*

$$d_{\mathrm{mGP}}(\chi, \beta_0 \circ \alpha(\chi)) \leq d_{\mathrm{P}}(\varpi(\nu^{\chi}), \delta_0).$$

Corollary 7.2. *For every $\chi \in \mathbb{U}_N$,*

$$d_{\mathrm{mGP}}(\beta(\chi), \beta_0(\chi)) \leq d_{\mathrm{P}}(\varpi(\nu^{\beta(\chi)}), \delta_0).$$

Proof. This is immediate from Lemma 7.1 as $\alpha \circ \beta$ is the identity on $\mathbb{U}_N$. □

We prove Lemma 7.1 using a characterization of the marked Gromov-Prohorov metric d_{mGP} that we will apply also in Section 8. The distortion dis $\mathfrak{R}$ of a relation $\mathfrak{R} \subset X \times X'$ between two metric spaces (X, r) and (X', r') is defined by

$$\mathrm{dis}\, \mathfrak{R} = \sup\{|r(x, y) - r'(x', y')| : (x, x'), (y, y') \in \mathfrak{R}\}.$$

Proposition 7.3. *Let (X, r, m) and (X', r', m') be marked metric measure spaces. Then $d_{\mathrm{mGP}}([\![X, r, m]\!], [\![X', r', m']\!])$ is the infimum of all $c > 0$ such that there exist a relation $\mathfrak{R} \subset X \times X'$ and a coupling ν of m and m' with $\frac{1}{2}\mathrm{dis}\, \mathfrak{R} \leq c$ and $\nu(\tilde{\mathfrak{R}}) \geq 1 - c$, where $\hat{\mathfrak{R}} \subset (X \times \mathbb{R}_+) \times (X' \times \mathbb{R}_+)$ is defined by*

$$\hat{\mathfrak{R}} = \{((x, v), (x', v')) : (x, x') \in \mathfrak{R}, |v - v'| \leq c\}.$$

Proof. This can be seen as an adaptation of Proposition 6 in [70]. Here we sketch the proof of the upper bound for $d_{\mathrm{mGP}}([\![X, r, m]\!], [\![X', r', m']\!])$. Let $c > 0$ and assume $\mathfrak{R}$, $\hat{\mathfrak{R}}$, and ν with $\frac{1}{2}\mathrm{dis}\, \mathfrak{R} \leq c$ and $\nu(\hat{\mathfrak{R}}) \geq 1 - c$ are given as in the proposition. A metric d^Z on the disjoint union $Z = X \sqcup X'$ can be defined by $d^Z(x, y) = r(x, y)$ for $(x, y) \in X$, $d^Z(x, y) = r'(x, y)$ for $(x, y) \in X'$, and

$$d^Z(x, x') = \inf\{r(x, y) + c + r'(y', x') : (y, y') \in \mathfrak{R}\},$$

for $x \in X, x' \in X'$, as in Remark 5.5 of [45]. We endow $Z \times \mathbb{R}_+$ with the product metric $d^{Z \times \mathbb{R}_+}((z, v), (z', v')) = d^Z(z, z') \vee |v - v'|$. Let $\varphi : X \to Z$ and $\varphi' : X' \to Z$ be the canonical embeddings. Moreover, let $\hat{\varphi}(x, v) = (\varphi(x), v)$ and $\hat{\varphi}'(x', v) = (\varphi'(x'), v)$ for $x \in X$, $x' \in X'$, and $v \in \mathbb{R}_+$. Then the coupling ν induces a coupling $\hat{\nu}$ of $\hat{\varphi}(m)$ and $\hat{\varphi}'(m')$ on $Z \times \mathbb{R}_+$ with

$$\hat{\nu}\{((z, v), (z', v')) : d^Z(z, z') \vee |v - v'| \leq c\} \geq 1 - c.$$

The coupling characterization (7.2) of the Prohorov metric implies $d_{\mathrm{P}}(\hat{\varphi}(m), \hat{\varphi}'(m')) \leq c$. The definition of the marked Gromov-Prohorov metric, see [25, Definition 3.1], implies $d_{\mathrm{mGP}}([\![X, r, m]\!], [\![X', r', m']\!]) \leq c$. □

Proof of Lemma 7.1. Let (r, v) be any element of $\hat{\mathfrak{U}}_N$ with $\hat{\psi}_N(r, v) = \chi$, and let $\rho = \alpha(r, v)$. Then,

$$[\![[N], r, N^{-1} \sum_{i=1}^{N} \delta_{(i, v(i))}]\!] = \chi$$

and

$$[\![[N], \rho, N^{-1} \sum_{i=1}^{N} \delta_{(i,0)}]\!] = \beta_0 \circ \alpha(\chi).$$

Let $c > d_{\mathrm{P}}(\varpi(\nu^{\chi}), \delta_0)$. We conclude by Proposition 7.3 which also holds for marked semi-metric measure spaces. To this aim, we define the relation

$$\mathfrak{R} = \{(i,i) \in [N] \times [N] : v(i) \leq c\}$$

between the semi-metric spaces $([N], r)$ and $([N], \rho)$. As $|r(i,j) - \rho(i,j)| \leq v(i) + v(j)$ by definition of the map $\alpha : \hat{\mathfrak{U}}_N \to \mathfrak{U}_N$, we can bound the distortion by

$$\operatorname{dis} \mathfrak{R} = \max\{|r(i,j) - \rho(i,j)| : i,j \in [N], v(i), v(j) \leq c\} \leq 2c.$$

We set

$$\hat{\mathfrak{R}} = \{((i, v(i)), (i, 0)) : i \in [N], v(i) \leq c\} \subset ([N] \times \mathbb{R}_+) \times ([N] \times \mathbb{R}_+).$$

A coupling ν of the probability measures $N^{-1} \sum_{i=1}^{N} \delta_{(i,v(i))}$ and $N^{-1} \sum_{i=1}^{N} \delta_{(i,0)}$ is given by

$$\nu = N^{-1} \sum_{i=1}^{N} \delta_{((i,v(i)),(i,0))}.$$

Finally,

$$\nu(\hat{\mathfrak{R}}) = N^{-1} \sum_{i=1}^{N} \mathbf{1}\{v(i) \leq c\} = \varpi(\nu^{\chi})[0,c] \geq \delta_0\{0\} - c = 1 - c,$$

where the inequality follows from the choice of c and the usual definition (7.1) of the Prohorov metric. □

Lemma 7.4. *Let $\Xi \in \mathcal{M}_{\mathrm{nd}}$, and let $(\chi_t, t \in \mathbb{R}_+)$ be a $\mathbb{U}$-valued Ξ-Fleming-Viot process with càdlàg paths. Then,*

$$\lim_{\ell \to \infty} \sup_{t \in [0,T]} d_{\mathrm{P}}(\Upsilon_1^{\ell}(\nu^{\chi_t}), \delta_0) = 0 \quad \text{a. s.}$$

for all $T \in \mathbb{R}_+$.

Proof. Let $\varepsilon > 0$. For $\ell \geq 2$, we define the random time

$$\vartheta_{\varepsilon,\ell} = \inf\{t \in \mathbb{R}_+ : d_{\mathrm{P}}(\Upsilon_1^{\ell}(\nu^{\chi_t}), \delta_0) > \varepsilon\}.$$

For all $\rho \in \mathfrak{U}$, the map $\ell \mapsto \Upsilon_1^{\ell}(\rho)$ is non-increasing. We set

$$\vartheta_{\varepsilon} = \sup_{\ell \in \mathbb{N}} \vartheta_{\varepsilon,\ell} = \lim_{\ell \to \infty} \vartheta_{\varepsilon,\ell}.$$

Let $t \in (0, \infty)$ be arbitrary. On an event of probability 1, let (X, ρ, μ) be a representative of $\chi_{\vartheta_{\varepsilon} \wedge t}$, and let $(x_i, i \in \mathbb{N})$ be a μ-iid sequence in X. Then,

$$\inf_{j \in \mathbb{N} \setminus \{1\}} \rho(x_1, x_j) = 0 \quad \text{a. s.}$$

as an iid sequence with respect to a probability measure on the Borel sigma algebra on a separable metric space has a. s. no isolated elements. Consequently,

$$\lim_{\ell\to\infty} \Upsilon_1^\ell((\rho(x_i,x_j))_{i,j\in\mathbb{N}}) = 0 \quad \text{a. s.}$$

As the random matrix $(\rho(x_i,x_j))_{i,j\in\mathbb{N}}$ has conditional distribution $\nu^{\chi_{\vartheta_\varepsilon\wedge t}}$ given $\chi_{\vartheta_\varepsilon\wedge t}$, it follows

$$\lim_{\ell\to\infty} d_{\mathrm{P}}(\Upsilon_1^\ell(\nu^{\chi_{\vartheta_\varepsilon\wedge t}}),\delta_0) = 0 \quad \text{a. s.}$$

Analogously, it can be shown that

$$\lim_{\ell\to\infty} d_{\mathrm{P}}(\Upsilon_1^\ell(\nu^{\chi_{(\vartheta_\varepsilon\wedge t)-}}),\delta_0) = 0 \quad \text{a. s.}$$

As $(\chi_s, s\in\mathbb{R}_+)$ has càdlàg paths and as the maps $\chi\mapsto\nu^\chi$ and Υ_1^ℓ are continuous, it follows that on an event of probability 1, there exists $\ell\in\mathbb{N}$ such that

$$d_{\mathrm{P}}(\Upsilon_1^\ell(\nu^{\chi_s}),\delta_0) < \varepsilon/2$$

for all s in a neighborhood of $\vartheta_\varepsilon\wedge t$. By monotonicity, it also holds

$$d_{\mathrm{P}}(\Upsilon_1^{\ell'}(\nu^{\chi_s}),\delta_0) < \varepsilon/2$$

for all $\ell'\geq\ell$ and s in the same neighborhood, on the same event of probability 1. This implies $\vartheta_{\ell,\varepsilon} > t$ for ℓ sufficiently large a. s., hence $\{\vartheta_\varepsilon < t\}$ is a null event. As t was arbitrary, it follows $\vartheta_\varepsilon = \infty$ a. s. which yields the assertion. □

Proof of Theorem 3.5 (beginning). First we assume $\Xi\in\mathcal{M}_{\mathrm{nd}}$. Let $T\in\mathbb{R}_+$. By the assumption that $\hat{\chi}_0$ supports only the zero mark, there exists $\chi_0\in\mathbb{U}$ such that $\hat{\chi}_0^N$ converges to $\beta_0(\chi_0)$ in the marked Gromov-weak topology as N tends to infinity. Hence, $\nu^{\hat{\chi}_0^N}$ converges weakly to $\nu^{\chi_0}\otimes\delta_0$. Recall the chain $(\chi_k^N, k\in\mathbb{N}_0)$ from Section 3.2. As $\nu^{\chi_0^N} = \alpha(\nu^{\hat{\chi}_0^N})$ converges weakly to ν^{χ_0}, also χ_0^N converges to χ_0 in the Gromov-weak topology. Hence, Theorem 3.2 is applicable and the processes $(\chi^N_{\lfloor c_N^{-1}t\rfloor}, t\in\mathbb{R}_+)$ converge in distribution to a $\mathbb{U}$-valued Ξ-Fleming-Viot process $(\chi_t, t\in\mathbb{R}_+)$ with initial state χ_0 in the space of càdlàg paths in $(\mathbb{U}, d_{\mathrm{GP}})$, endowed with the Skorohod metric. For every $\ell\geq 2$, by continuity of the maps $\chi\mapsto\nu^\chi$ and Υ_1^ℓ, also $(\Upsilon_1^\ell(\nu^{\chi^N_{\lfloor c_N^{-1}t\rfloor}}), t\in\mathbb{R}_+)$ converges in distribution to $(\Upsilon_1^\ell(\nu^{\chi_t}), t\in\mathbb{R}_+)$ in the space of càdlàg paths in $(\mathcal{M}_1(\mathbb{R}_+), d_{\mathrm{P}})$, endowed with the Skorohod metric, where $\mathcal{M}_1(\mathbb{R}_+)$ denotes the space of probability measures on $\mathbb{R}_+$.

For every $(r,v)\in\hat{\mathfrak{U}}$ and $\rho=\alpha(r,v)$, it holds $\rho(1,j)\geq v(1)$ for all $j\geq 2$ by definition of the map $\alpha:\hat{\mathfrak{U}}_N\to\mathfrak{U}_N$, hence $v(1)\leq 2\Upsilon_1^\ell(\alpha(r,v))$ for all $\ell\geq 2$. This implies

$$\begin{aligned}
&\sup_{t\in[0,T]} d_{\mathrm{P}}(\varpi(\nu^{\hat{\chi}^N_{\lfloor c_N^{-1}t\rfloor}}),\delta_0) \qquad (7.3)\\
&\leq \sup_{t\in[0,T]} d_{\mathrm{P}}(2\Upsilon_1^\ell(\alpha(\nu^{\hat{\chi}^N_{\lfloor c_N^{-1}t\rfloor}})),\delta_0)\\
&\leq 2\sup_{t\in[0,T]} d_{\mathrm{P}}(\Upsilon_1^\ell(\nu^{\chi^N_{\lfloor c_N^{-1}t\rfloor}}),\delta_0).
\end{aligned}$$

The expression in the last line converges in distribution to

$$2 \sup_{t\in[0,T]} d_{\mathrm{P}}(\Upsilon_1^\ell(\nu^{\chi_t}), \delta_0) \tag{7.4}$$

as N tends to infinity. This follows from the discussion in the beginning of this proof, as the maps Υ_1^ℓ and $d_{\mathrm{P}}(\cdot, \delta_0)$ are continuous, and as the process $(\Upsilon_1^\ell(\nu^{\chi_t}), t \in \mathbb{R}_+)$ has a. s. no discontinuity at the fixed time T. Finally, we let ℓ tend to infinity. By Lemma 7.4, expression (7.4) then converges to zero a. s. Consequently, also the left-hand side of (7.3) converges to zero in probability as N tends to infinity.

As $\hat{\chi}_k^N = \beta(\chi_k^N)$ for all $k \in \mathbb{N}_0$, Corollary 7.2 implies that

$$\sup_{t\in[0,T]} d_{\mathrm{mGP}}(\hat{\chi}^N_{\lfloor c_N^{-1} t\rfloor}, \beta_0(\chi^N_{\lfloor c_N^{-1} t\rfloor}))$$

converges to zero in probability. By another application of Theorem 3.2, the processes $(\beta_0(\chi^N_{\lfloor c_N^{-1} t\rfloor}), t \in \mathbb{R}_+)$ converge in distribution to $(\beta_0(\chi_t), t \in \mathbb{R}_+)$. The assertion for $\Xi \in \mathcal{M}_{\mathrm{nd}}$ now follows from Slutzky's theorem. □

To prove Proposition 3.8 in case $\Xi \in \mathcal{M}_{\mathrm{nd}}$, we will use the following coupling of $(\tilde{\chi}_k^N, k \in \mathbb{N}_0)$ and $(\chi_k^N, k \in \mathbb{N}_0)$.

Remark 7.5. If $\chi_0^N = \alpha(\tilde{\chi}_0^N)$, then we can define (r_0^N, v_0^N) and ρ_0^N in Section 3 such that

$$\rho_0^N = \alpha(r_0^N, v_0^N).$$

We can then also assume that the processes $((r_k^N, v_k^N), k \in \mathbb{N}_0)$, $(\tilde{\chi}_k^N, k \in \mathbb{N}_0)$, and $(\rho_k^N, k \in \mathbb{N}_0)$ are defined as in Section 3.5. Then we obtain $\rho_k^N = \alpha(r_k^N, v_k^N)$ for all $k \in \mathbb{N}$. When we define $\chi_k^N = \psi_N(\rho_k^N)$ and $\hat{\chi}_k^N = \hat{\psi}_N \circ \beta(\rho_k^N)$ as in Section 3, then we also have $\chi_k^N = \alpha(\tilde{\chi}_k^N)$ and $\hat{\chi}_k^N = \beta \circ \alpha(\tilde{\chi}_k^N)$ for all $k \in \mathbb{N}_0$.

Proof of Proposition 3.8 (end). In case $\Xi \in \mathcal{M}_{\mathrm{nd}}$, the proof is almost identical with the proof of Theorem 3.5. We replace $\hat{\chi}^N_\cdot$ with $\tilde{\chi}^N_\cdot$, and we set $\chi_0^N = \alpha(\tilde{\chi}_0^N)$. Then we use the coupling from Remark 7.5 and apply Theorem 3.2. □

8 Convergence of marked metric measure spaces in the case with dust

In this section, we complete the proof of Theorem 3.5: In the case with dust, we compare the finite dimensional distributions of the processes $(\hat{\chi}^N_{\lfloor c_N^{-1} t\rfloor}, t \in \mathbb{R}_+)$ and $(\tilde{\chi}^N_{\lfloor c_N^{-1} t\rfloor}, t \in \mathbb{R}_+)$, and we show relative compactness of the sequence of processes $((\hat{\chi}^N_{\lfloor c_N^{-1} t\rfloor}, t \in \mathbb{R}_+), N \geq 2)$.

A metric $\tilde{d}$ on $\hat{\mathfrak{U}}_N$ is defined by

$$\tilde{d}((r,v),(r',v')) = \max_{i,j\in[N]} |r(i,j) - r'(i,j)| \vee \max_{i\in[N]} |v(i) - v'(i)|.$$

W. l. o. g., we endow $\hat{\mathfrak{U}}_N$ with the metric

$$d((r,v),(r',v')) = \max_{n\in[N]}(\tilde{d}(\gamma_n(r,v),\gamma_n(r',v')) \wedge (2^{-n}))$$

and $\hat{\mathfrak{U}}$ with the metric

$$d((r,v),(r',v')) = \sup_{n\in\mathbb{N}}(\tilde{d}(\gamma_n(r,v),\gamma_n(r',v')) \wedge (2^{-n})).$$

These metrics induce the topologies on $\hat{\mathfrak{U}}_N$ and $\mathfrak{U}$ from Section 2.1. We have

$$d(\gamma_n(r,v),\gamma_n(r',v')) \leq d((r,v),(r',v')) \leq d(\gamma_n(r,v),\gamma_n(r',v')) + 2^{-n} \tag{8.1}$$

for all $(r,v),(r',v') \in \hat{\mathfrak{U}}$.

We use definitions in particular from Section 2.1 and we need the following lemmas.

Lemma 8.1. *Let $n \in \mathbb{N}$ and $(r,v),(r',v') \in \hat{\mathfrak{U}}_n$ with $\alpha(r,v) = \alpha(r',v')$. Then,*

$$d((r,v),(r',v')) \leq 2\max_{i\in[n]} |v(i) - v'(i)| .$$

Proof. Let $\rho = \alpha(r,v) = \alpha(r',v')$. By definition of α,

$$|r(i,j)-r'(i,j)| = |\rho(i,j)-v(i)-v(j)-\rho(i,j)+v'(i)+v'(j)| \leq |v(i)-v'(i)|+|v(j)-v'(j)|$$

for all distinct $i,j \in [n]$. It follows

$$d((r,v),(r',v')) \leq \tilde{d}((r,v),(r',v')) \leq 2\max_{i\in[n]} |v(i) - v'(i)| .$$

□

In the next lemma, we consider the chain $(\tilde{\chi}^N_k, k \in \mathbb{N}_0)$ from Section 3.5. We show that if the initial state corresponds to the decomposition at the external branches, then in all generations, the mark of a sampled individual is not larger than the length of the corresponding external branch.

Lemma 8.2. *Let the $\hat{\mathbb{U}}_N$-valued chain $(\tilde{\chi}^N_k, k \in \mathbb{N}_0)$ be defined from the $\hat{\mathfrak{U}}_N$-valued chain $((r^N_k, v^N_k), k \in \mathbb{N}_0)$ as in Section 3.5, and assume that $(r^N_0, v^N_0) = \beta \circ \alpha(r^N_0, v^N_0)$. Then,*

$$v \leq \Upsilon \circ \alpha(r,v) \quad \textit{(component-wise)}$$

for $\nu^{N,\tilde{\chi}^N_k}$-a. a. $(r,v) \in \hat{\mathfrak{U}}_N$ and all $k \in \mathbb{N}_0$ a. s.

Proof. Let $(\rho^N_k, k \in \mathbb{N}_0) = (\alpha(r^N_k, v^N_k), k \in \mathbb{N}_0)$ as in Section 3.5. By definition of Υ,

$$\Upsilon(\rho^N_k)(i) = \tfrac{1}{2} \min_{j\in[N]\setminus\{i\}} \rho^N_k(i,j) \tag{8.2}$$

for all $i \in [N]$ and $k \in \mathbb{N}_0$. We show $v^N_k(i) \leq \Upsilon(\rho^N_k)(i)$. The assertion follows as for $\nu^{N,\tilde{\chi}^N_k}$-a. a. $(r,v) \in \hat{\mathfrak{U}}_N$, there exists a. s. a bijection p on $[N]$ with $v(i) = v^N_k(p(i))$ for all $i \in [N]$ by definition of the N-marked distance matrix distribution.

Let $i \in [N]$, $k \in \mathbb{N}_0$, and $\bar{v}(i) = \Upsilon(\rho_k^N(i))$. To show $v_k^N(i) \le \bar{v}(i)$, we consider the cases $v_k^N(i) \le c_N k$ and $v_k^N(i) \ge c_N k$ separately.

In the case $v_k^N(i) \le c_N k$, there are, by definition of v_k^N, no reproduction events from generations $k - c_N^{-1} v_k^N(i) + 1$ to k due to which the ancestral lineage of the individual i of generation k can merge with the ancestral lineage of a different individual. Hence, by definition of ρ_k^N in Section 3.1,

$$\min_{j \in [N] \setminus \{i\}} \rho_k^N(i,j) \ge 2 v_k^N(i).$$

Equation (8.2) yields $v_k^N(i) \le \bar{v}(i)$ in case $v_k^N(i) \le c_N k$.

The statements in the remainder of this proof hold in the case $v_k^N(i) \ge c_N k$. There are no reproduction events in which the ancestral lineage of the individual i of generation k can merge with the ancestral lineage of a different individual. Hence, $A_0(k,i) \ne A_0(k,j)$ for all $j \in [N] \setminus \{i\}$. By definition of ρ_k^N,

$$\rho_k^N(i,j) = 2c_N k + \rho_0^N(A_0(k,i), A_0(k,j)) \tag{8.3}$$

for all $j \in [N] \setminus \{i\}$. By definition of v_k^N,

$$v_k^N(i) = c_N k + v_0^N(A_0(k,i)).$$

Also, $v_0^N = \Upsilon(\rho_0^N)$ by our assumption, hence

$$v_0^N(A_0(k,i)) \le \tfrac{1}{2} \rho_0^N(A_0(k,i), \ell) \tag{8.4}$$

for all $\ell \in [N] \setminus \{A_0(k,i)\}$. Using equations (8.2), (8.3), and (8.4), we obtain

$$\bar{v}(i) = \tfrac{1}{2} \min_{j \in [N] \setminus \{p(i)\}} \rho_k^N(i,j) \ge c_N k + v_0^N(A_0(k,i)) = v_k^N(i).$$

□

For $\ell \ge n \ge 2$, we introduce the map

$$\Upsilon_n^\ell : \mathfrak{U} \cup \bigcup_{N \ge \ell} \mathfrak{U}_N \to \mathbb{R}_+^n, \quad \Upsilon_n^\ell = \gamma_n \circ \Upsilon \circ \gamma_\ell.$$

The vector $\Upsilon_n^\ell(\rho)$ gives the lengths of the first n external branches in the subtree spanned by the first ℓ leaves of the tree associated with some $\rho \in \mathfrak{U}_N$ or $\rho \in \mathfrak{U}$. We also define the restriction $\varpi_{\mathbb{R}_+^{\mathbb{N}}} : \hat{\mathfrak{U}} \to \mathbb{R}_+^{\mathbb{N}}$, $(r,v) \mapsto v$. We endow $\mathbb{R}_+^n$ with the maximum norm and the induced metric. Let $\gamma_n : \mathbb{R}_+^{\mathbb{N}} \to \mathbb{R}_+^n$ be the restriction $v \mapsto (v(i))_{i \in [n]}$. Recall also the maps α and β from Sections 2.1 and 7.

Proof of Theorem 3.5 (continuation). We set $\tilde{\chi}_0^N = \hat{\chi}_0^N$. Then, $\chi_0^N = \alpha(\tilde{\chi}_0^N)$ and we can assume that the processes $(\tilde{\chi}_k^N, k \in \mathbb{N}_0)$, $(\hat{\chi}_k^N, k \in \mathbb{N}_0)$, and $((r_k^N, v_k^N), k \in \mathbb{N}_0)$ are defined as in Remark 7.5. In particular, we have $\hat{\chi}_k^N = \beta \circ \alpha(\tilde{\chi}_k^N)$ for all $k \in \mathbb{N}_0$ a. s.

First we consider finite-dimensional distributions. Let $t \in \mathbb{R}_+$. For $N \geq \ell \geq n \geq 2$,

$$\begin{aligned}
&d_{\mathrm{P}}(\nu^{\hat{\chi}^N_{\lfloor c_N^{-1} t\rfloor}}, \nu^{\tilde{\chi}^N_{\lfloor c_N^{-1} t\rfloor}}) \\
&\leq d_{\mathrm{P}}(\gamma_n(\nu^{\beta\circ\alpha(\tilde{\chi}^N_{\lfloor c_N^{-1} t\rfloor})}), \gamma_n(\nu^{\tilde{\chi}^N_{\lfloor c_N^{-1} t\rfloor}})) + 2^{-n+1} \\
&\leq d_{\mathrm{P}}(\gamma_n(\nu^{N,\beta\circ\alpha(\tilde{\chi}^N_{\lfloor c_N^{-1} t\rfloor})}), \gamma_n(\nu^{N,\tilde{\chi}^N_{\lfloor c_N^{-1} t\rfloor}})) + 2^{-n+1} + 2n^2/N \\
&= d_{\mathrm{P}}(\gamma_n \circ \beta \circ \alpha(\nu^{N,\tilde{\chi}^N_{\lfloor c_N^{-1} t\rfloor}}), \gamma_n(\nu^{N,\tilde{\chi}^N_{\lfloor c_N^{-1} t\rfloor}})) + 2^{-n+1} + 2n^2/N \\
&\leq 2d_{\mathrm{P}}(\gamma_n \circ \Upsilon \circ \alpha(\nu^{N,\tilde{\chi}^N_{\lfloor c_N^{-1} t\rfloor}}), \gamma_n \circ \varpi_{\mathbb{R}_+^{\mathbb{N}}}(\nu^{N,\tilde{\chi}^N_{\lfloor c_N^{-1} t\rfloor}})) + 2^{-n+1} + 2n^2/N \quad \text{a.s.}
\end{aligned}$$

For the first inequality, we use relation (8.1) and either (7.1) or (7.2). For the last inequality, we use Lemma 8.1, the definitions of Υ and β, and (7.1) or (7.2). For the second inequality, we use the bounds

$$d_{\mathrm{P}}(\gamma_n(\nu^{\beta\circ\alpha(\tilde{\chi}^N_{\lfloor c_N^{-1} t\rfloor})}), \gamma_n(\nu^{N,\beta\circ\alpha(\tilde{\chi}^N_{\lfloor c_N^{-1} t\rfloor})})) \leq n^2/N$$

and

$$d_{\mathrm{P}}(\gamma_n(\nu^{\tilde{\chi}^N_{\lfloor c_N^{-1} t\rfloor}}), \gamma_n(\nu^{N,\tilde{\chi}^N_{\lfloor c_N^{-1} t\rfloor}})) \leq n^2/N$$

which can be seen from the coupling characterization (7.2) of the Prohorov metric. Here we couple sampling with and without replacement as in the proofs of e. g. (6.2) or of Lemma 6.4.

By definition of Υ_n^ℓ and Υ,

$$\gamma_n \circ \Upsilon(\rho) \leq \Upsilon_n^\ell(\rho)$$

for all $\rho \in \mathfrak{U}_N$. Using Lemma 8.2, we obtain

$$|\gamma_n \circ \Upsilon \circ \alpha(r, v) - \gamma_n(v)| \leq |\Upsilon_n^\ell \circ \alpha(r, v) - \gamma_n(v)|$$

for $\nu^{\tilde{\chi}^N_{\lfloor c_N^{-1} t\rfloor}}$-a. a. $(r, v) \in \hat{\mathfrak{U}}_N$ a. s. Again using the definition of the Prohorov metric, we obtain the first inequality in the following display.

$$\begin{aligned}
&d_{\mathrm{P}}(\gamma_n \circ \Upsilon \circ \alpha(\nu^{N,\tilde{\chi}^N_{\lfloor c_N^{-1} t\rfloor}}), \gamma_n \circ \varpi_{\mathbb{R}_+^{\mathbb{N}}}(\nu^{N,\tilde{\chi}^N_{\lfloor c_N^{-1} t\rfloor}})) \\
&\leq d_{\mathrm{P}}(\Upsilon_n^\ell \circ \alpha(\nu^{N,\tilde{\chi}^N_{\lfloor c_N^{-1} t\rfloor}}), \gamma_n \circ \varpi_{\mathbb{R}_+^{\mathbb{N}}}(\nu^{N,\tilde{\chi}^N_{\lfloor c_N^{-1} t\rfloor}})) \\
&\leq d_{\mathrm{P}}(\Upsilon_n^\ell \circ \alpha(\nu^{\tilde{\chi}^N_{\lfloor c_N^{-1} t\rfloor}}), \gamma_n \circ \varpi_{\mathbb{R}_+^{\mathbb{N}}}(\nu^{\tilde{\chi}^N_{\lfloor c_N^{-1} t\rfloor}})) + \ell^2/N + n^2/N \quad \text{a.s.}
\end{aligned} \tag{8.5}$$

For the second inequality, we again couple sampling with and without replacement, and we use the triangle inequality twice. By continuity of the maps $\chi \mapsto \nu^\chi$, α, Υ_n^ℓ, $\varpi_{\mathbb{R}_+^{\mathbb{N}}}$, γ_n, and $d_{\mathrm{P}}(\cdot, \cdot)$, and by Proposition 3.8, the right-hand side of (8.5) converges in distribution to

$$d_{\mathrm{P}}(\Upsilon_n^\ell \circ \alpha(\nu^{\hat{\chi}_t}), \gamma_n \circ \varpi_{\mathbb{R}_+^{\mathbb{N}}}(\nu^{\hat{\chi}_t})) \tag{8.6}$$

as $N \to \infty$, where $(\hat{\chi}_s, s \in \mathbb{R}_+)$ is a $\hat{\mathbb{U}}$-valued Ξ-Fleming-Viot process with initial state $\hat{\chi}_0$. Here we also use that $(\hat{\chi}_s, s \in \mathbb{R}_+)$ has a. s. no discontinuity at the fixed time t.

For every $\rho \in \mathfrak{U}$,

$$\lim_{\ell\to\infty} \Upsilon_n^\ell(\rho) = \gamma_n \circ \Upsilon(\rho)$$

by definition of Υ_n^ℓ, and by the definition of $\Upsilon : \mathfrak{U} \to \mathbb{R}_+^{\mathbb{N}}$ in Section 2.1. Hence, for every $\chi \in \hat{\mathbb{U}}$,

$$\lim_{\ell\to\infty} d_{\mathrm{P}}(\Upsilon_n^\ell \circ \alpha(\nu^\chi), \gamma_n \circ \Upsilon \circ \alpha(\nu^\chi)) = 0.$$

By Proposition 3.3 of Chapter 2,

$$\Upsilon \circ \alpha(\nu^\chi) = \varpi_{\mathbb{R}_+^{\mathbb{N}}}(\nu^\chi).$$

Therefore, expression (8.6) converges to zero as ℓ tends to infinity. Now we let also n tend to infinity. Then we can deduce that

$$d_{\mathrm{P}}(\nu^{\hat{\chi}^N_{\lfloor c_N^{-1} t\rfloor}}, \nu^{\tilde{\chi}^N_{\lfloor c_N^{-1} t\rfloor}})$$

converges to zero in probability as N tends to infinity.

As an immediate consequence, for $t_1, \ldots, t_k \in \mathbb{R}_+$, the vector

$$(d_{\mathrm{P}}(\nu^{\hat{\chi}^N_{\lfloor c_N^{-1} t_i\rfloor}}, \nu^{\tilde{\chi}^N_{\lfloor c_N^{-1} t_i\rfloor}}), i \in [k])$$

converges to zero in probability as N tends to infinity. Using Proposition 3.8 again, Slutzky's theorem, and that $(\hat{\chi}_s, s \in \mathbb{R}_s)$ has a. s. no fixed times of discontinuity, we deduce the convergence in distribution

$$(\nu^{\hat{\chi}^N_{\lfloor c_N^{-1} t_i\rfloor}}, i \in [k]) \xrightarrow{d} (\nu^{\hat{\chi}_{t_i}}, i \in [k]) \quad (N \to \infty).$$

Hence,

$$(\hat{\chi}^N_{\lfloor c_N^{-1} t_i\rfloor}, i \in [k]) \xrightarrow{d} (\hat{\chi}_{t_i}, i \in [k]) \quad (N \to \infty)$$

in the marked Gromov-weak topology. As this convergence determines the finite-dimensional distributions of possible limit processes, it now suffices to show relative compactness of the sequence of processes $((\hat{\chi}^N_{\lfloor c_N^{-1} t\rfloor}, t \in \mathbb{R}_+), N \in \mathbb{N})$, see [36, Theorem 3.7.8]. □

To show the desired relative compactness in the proof of Theorem 3.5, we use the following lemma.

Lemma 8.3. *Let $k \in \mathbb{N}_0$ and $N \geq 2$. Then,*

$$d_{\mathrm{mGP}}(\hat{\chi}_k^N, \hat{\chi}_{k+1}^N) \leq 2N^{-1}(N - \#\pi_{k+1}^N) + c_N.$$

The bound in Lemma 8.3 has the following meaning. There are at most $N - \#\pi_{k+1}^N$ many individuals in generation k that have more than one offspring in generation $k+1$. For each such offspring, the associated external branch has length c_N which needs not coincide with the external branch length of the ancestor in generation k. There are also $N - \#\pi_{k+1}^N$ individuals in generation k that die, and each such death can drastically increase an external branch length in generation $k+1$ (this is the freeing phenomenon mentioned in Section 3.5). For the other individuals in generation k, the external branch lengths increase by c_N from generation k to $k+1$. As each individual has weight N^{-1}, the bound is a consequence of the definition of the marked Gromov-Prohorov metric and the coupling characterization of the Prohorov metric.

Proof. Let $L \subset [N]$ denote the set of the labels of the individuals of generation k that have offspring in generation $k+1$, that is,

$$L = \{i \in [N] : \exists j \in [N] \text{ with } A_k(k+1, j) = i\}.$$

By definition of the population model in Section 3,

$$\#L = \#\pi_{k+1}^N. \tag{8.7}$$

For all $j_1, j_2 \in [N]$ with $j_1 \neq j_2$ and $i_1 = A_k(k+1, j_1)$, $i_2 = A_k(k+1, j_2)$, by definition of the population model in Section 3,

$$\rho_{k+1}^N(j_1, j_2) = \rho_k^N(i_1, i_2) + 2c_N. \tag{8.8}$$

For $i \in [N]$, we define the set

$$C_i = \{j \in [N] \setminus \{i\} : \rho_k^N(i, j) = \min\{\rho_k^N(i, \ell) : \ell \in [N] \setminus \{i\}\}\}.$$

In words, C_i consists of the individuals other than i with minimal distance to the individual i. That is, the set $C_i \cup \{i\}$ is the minimal clade of the individual i in the sense of [15]. Moreover, we define

$$M = \{i \in [N] : C_i \cap L \neq \emptyset, \exists! j \in [N] \text{ with } A_k(k+1, j) = i\}.$$

For $i \in M$, the individual i of generation k has exactly one offspring j in generation $k+1$, and at least one other member of the minimal clade of i has offspring in generation $k+1$. Hence, the minimal clade of i in generation k and the minimal clade of j in generation $k+1$ have the same most recent common ancestor. This implies, for i and j as above,

$$\min_{\ell \in [N] \setminus \{j\}} \rho_{k+1}^N(j, \ell) = \min_{\ell \in [N] \setminus \{i\}} \rho_k^N(i, \ell) + 2c_N. \tag{8.9}$$

We write $(r, v) = \beta(\rho_k^N)$ and $(r', v') = \beta(\rho_{k+1}^N)$. For $i \in M$, let $d(i)$ denote the label of the unique descendant in generation $k+1$ of the individual i of generation k. For all $i \in M$ and $j = d(i)$,

$$v'(j) = \tfrac{1}{2} \min_{\ell \in [N] \setminus \{j\}} \rho_{k+1}^N(j, \ell) = \tfrac{1}{2} \min_{\ell \in [N] \setminus \{i\}} \rho_k^N(i, \ell) + c_N = v(i) + c_N \tag{8.10}$$

by equation (8.9). For $i_1, i_2 \in M$ with $i_1 \neq i_2$ and $j_1 = d(i_1)$, $j_2 = d(i_2)$, it holds $j_1 \neq j_2$, and by equations (8.8) and (8.10)

$$r'(j_1, j_2) = \rho^N_{k+1}(j_1, j_2) - v'(j_1) - v'(j_2) = \rho^N_k(i_1, i_2) - v(i_1) - v(i_2) = r(i_1, i_2). \quad (8.11)$$

We define a relation $\mathfrak{R}$ between the semi-metric spaces $([N], r)$ and $([N], r')$ by

$$\mathfrak{R} = \{(i, d(i)) \in [N]^2 : i \in M\}.$$

Equation (8.11) implies that the distortion of $\mathfrak{R}$ equals zero,

$$\operatorname{dis} \mathfrak{R} = \max\{|r(i_1, i_2) - r'(j_1, j_2)| : (i_1, j_1), (i_2, j_2) \in \mathfrak{R}\} = 0.$$

We set

$$\hat{\mathfrak{R}} = \{((i, v(i)), (j, v'(j))) \in ([N] \times \mathbb{R}_+)^2 : (i, j) \in \mathfrak{R}\}.$$

There exists a coupling ν of the probability measures $N^{-1} \sum_{i=1}^N \delta_{(i,v(i))}$ and $N^{-1} \sum_{j=1}^N \delta_{(j,v'(j))}$ on $[N] \times \mathbb{R}_+$ with

$$\nu(\hat{\mathfrak{R}}) \geq N^{-1} \# M = 1 - N^{-1}(N - \# M).$$

By equation (8.10), it holds $|v(i) - v'(j)| \leq c_N$ for $(i, j) \in \mathfrak{R}$. By Proposition 7.3, which also holds for marked semi-metric measure spaces, it follows

$$d_{\mathrm{mGP}}(\hat{\chi}^N_k, \hat{\chi}^N_{k+1}) \leq N^{-1}(N - \# M) + c_N.$$

It remains to show

$$N - \# M \leq 2(N - \# L). \quad (8.12)$$

The assertion then follows by equation (8.7).

For $i \in [N]$, let $I_i = \mathbf{1}\{i \in L, C_i \subset L^c\}$. (Then I_i is the indicator variable that individual i reproduces as the only individual of its minimal clade.) Let $i, j \in [N]$ with $i \neq j$ and consider the case that there exists $\ell \in C_i \cap C_j$. W. l. o. g., we assume $\rho^N_k(j, \ell) \leq \rho^N_k(i, \ell)$ (if this does not hold, we transpose i and j). As $\rho^N_k \in \mathfrak{U}$, we obtain

$$\rho^N_k(i, j) \leq \rho^N_k(i, \ell) \vee \rho^N_k(j, \ell) = \rho^N_k(i, \ell).$$

As $\ell \in C_i$, it follows $j \in C_i$. If $I_i = 1$, then it follows that $j \in L^c$ and $I_j = 0$. Hence, in any case, the elements of the set $\mathcal{A} := \{C_i : i \in [N], I_i = 1\}$ are nonempty disjoint subsets of L^c, or it holds $\mathcal{A} = \emptyset$. This implies $\#\{C_i : i \in [N], I_i = 1\} \leq N - \# L$. Furthermore, generation k contains at most $N - \# L$ many individuals with more than one offspring in generation $k + 1$. The claim (8.12) follows by definition of M. □

Proof of Theorem 3.5 (end). We assume $\Xi \in \mathcal{M}_{\mathrm{dust}}$. To show relative compactness of the sequence of processes $((\hat{\chi}^N_{\lfloor c_N^{-1} t \rfloor}, t \in \mathbb{R}_+), N \in \mathbb{N})$, it suffices to verify condition (b) in Theorem 3.8.6 of [36]. Condition (a) in this theorem is satisfied as the one-dimensional distributions converge.

Using Lemma 8.3 and the bound

$$N - \# \pi \leq \# \cup \sigma$$

for $\pi \in \mathcal{P}_N$ and $\sigma = \{B \in \pi : \#B \geq 2\}$, we obtain

$$\mathbb{E}[d_{\mathrm{mGP}}(\hat{\chi}_k^N, \hat{\chi}_{k+1}^N)] \leq 2N^{-1}\mathbb{E}[\# \cup \sigma_1^N] + c_N$$

for all $k \in \mathbb{N}_0$. By exchangeability,

$$\mathbb{E}[\# \cup \sigma_1^N] = \sum_{i=1}^{N} \mathbb{E}[\mathbf{1}\{i \in \cup \sigma_1^N\}] = N\mathbb{P}(\gamma_1(\sigma_1^N) = \{\{1\}\}).$$

Let $(\mathcal{F}_t^N, t \in \mathbb{R}_+)$ be the filtration induced by the process $(\hat{\chi}^N_{\lfloor c_N^{-1} t \rfloor}, t \in \mathbb{R}_+)$. For $t \in \mathbb{R}_+$, $\delta > 0$, $u \in [0, \delta]$, and $s \in [0, \delta \wedge t]$, the Markov property of $(\hat{\chi}_k^N, k \in \mathbb{N}_0)$ at $\lfloor c_N^{-1} t \rfloor$ yields

$$\begin{aligned}
&\mathbb{E}[d_{\mathrm{mGP}}(\hat{\chi}^N_{\lfloor c_N^{-1}(t+u) \rfloor}, \hat{\chi}^N_{\lfloor c_N^{-1} t \rfloor}) | \mathcal{F}_t^N] d_{\mathrm{mGP}}(\hat{\chi}^N_{\lfloor c_N^{-1}(t-s) \rfloor}, \hat{\chi}^N_{\lfloor c_N^{-1} t \rfloor}) \\
&\leq \mathbf{1}\{\delta \geq c_N/2\}(\lfloor c_N^{-1} \delta \rfloor + 1)(2N^{-1}\mathbb{E}[\# \cup \sigma_1^N] + c_N) \\
&\leq \mathbf{1}\{\delta \geq c_N/2\}(\delta + c_N)(2c_N^{-1}\mathbb{P}(\gamma_1(\sigma_1^N) = \{\{1\}\}) + 1) \quad \text{a.s.}
\end{aligned} \tag{8.13}$$

In the first inequality, we also use $d_{\mathrm{mGP}} \leq 1$, and that if $\delta < c_N/2$, then at least one of the distances on the left-hand side of (8.13) equals zero.

It follows from the assumptions that the right-hand side of (8.13) converges to the expression $\delta(2\lambda_{1,\{\{1\}\}} + 1)$ as N tends to infinity. Now we show that the right-hand side of (8.13) converges to zero uniformly in N as δ tends to zero. For each $\varepsilon > 0$, there exists $N_\varepsilon \geq 2$ such that for all $N \geq N_\varepsilon$, it holds

$$c_N^{-1}\mathbb{P}(\gamma_1(\sigma_1^N) = \{\{1\}\}) \leq 2\lambda_{1,\{\{1\}\}}$$

and $c_N < \varepsilon$. Hence the right-hand side of (8.13) is bounded from above by $(\delta + \varepsilon)(4\lambda_{1,\{\{1\}\}} + 1)$ for $N \geq N_\varepsilon$. For δ sufficiently small and $N \leq N_\varepsilon$, the right-hand side of (8.13) equals zero.

As the right-hand side of (8.13) does not depend on t, we have verified (8.28) and (8.29) in Theorem 3.8.6 of [36]. To verify also (8.30), hence condition (b) in Theorem 3.8.6 of [36], we estimate as above

$$\mathbb{E}[d_{\mathrm{mGP}}(\hat{\chi}^N_{\lfloor c_N^{-1}\delta \rfloor}, \hat{\chi}_0^N)] \leq \lfloor c_N^{-1}\delta \rfloor (2N^{-1}\mathbb{E}[\# \cup \sigma_1^N] + c_N) \leq \delta(2c_N^{-1}\mathbb{P}(\gamma_1(\sigma_1^N) = \{\{1\}\}) + 1).$$

Also this expression converges to zero uniformly in N as δ tends to zero. □

List of notation

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

Miscellaneous
$\mathbb{R}_+ = [0, \infty)$, $\mathbb{N} = \{1, 2, 3, \ldots\}$, $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$, $[N] = \{1, \ldots, n\}$ for $N \in \mathbb{N}$
γ_n: restriction map in various contexts (p. 157, p. 163, p. 177)
d_{P}: Prohorov metric

(Marked) distance matrices

$\mathfrak{U}_N$, $\mathfrak{U}$: space of semi-ultrametrics on $[N]$, on $\mathbb{N}$ (p. 148)
$\hat{\mathfrak{U}}_N$, $\hat{\mathfrak{U}}$: space of decomposed semi-ultrametrics on $[N]$, on $\mathbb{N}$ (p. 148)
α: retrieves the semi-ultrametric from a decomposed semi-ultrametric (p. 148)
$\beta : \mathfrak{U}_N \to \hat{\mathfrak{U}}_N$: decomposition map into the external branches and the remaining subtree (p. 148)
$\Upsilon(\rho)$: vector of the lengths of the external branches in the coalescent tree associated with ρ (p. 148)

(Marked) metric measure spaces

$\mathbb{U}_N$, $\mathbb{U}$: spaces of isomorphy classes of ultrametric measure spaces (p. 149)
$\hat{\mathbb{U}}_N$, $\hat{\mathbb{U}}$: spaces of isomorphy classes of marked metric measure spaces (p. 150)
d_{GP}, d_{mGP}: Gromov-Prohorov metric, marked Gromov-Prohorov metric (pp. 149, 150)
ν^{χ}: distance matrix distribution of $\chi \in \mathbb{U}$, or marked distance matrix distribution of $\chi \in \hat{\mathbb{U}}$ (pp. 149, 150)
$\nu^{N,\chi}$: N-distance matrix distribution of $\chi \in \mathbb{U}$, or N-marked distance matrix distribution of $\chi \in \hat{\mathbb{U}}$ (pp. 149, 151)
$\mathcal{U}^{\mathrm{erg}}$: space of distance matrix distributions (p. 150)
$\psi_N : \mathfrak{U}_N \to \mathbb{U}_N$, $\hat{\psi}_N : \hat{\mathfrak{U}}_N \to \hat{\mathbb{U}}_N$: construction of (marked) metric measure spaces (pp. 149, 151)
$\alpha : \hat{\mathbb{U}}_N \to \mathbb{U}_N$: maps a decomposed unlabeled tree to an unlabeled tree (p. 171)
$\beta : \mathbb{U}_N \to \hat{\mathbb{U}}_N$: decomposes an unlabeled tree at the external branches (p. 155)
$\beta_0 : \mathbb{U}_N \to \hat{\mathbb{U}}_N$: adds the zero mark (p. 159)
$\mathcal{C}_n$, $\hat{\mathcal{C}}_n$: sets of bounded differentiable functions with bounded uniformly continuous derivative (p. 158)
Π: set of polynomials on $\mathbb{U}$ (p. 158)
$\hat{\Pi}$: set of marked polynomials on $\hat{\mathbb{U}}$ (p. 158)
$\mathscr{C}$: a set of test functions on $\mathcal{U}^{\mathrm{erg}}$ (p. 158)

Partitions and semi-partitions

$\mathcal{P}_N$: Set of partitions of $[N]$, associated transformations (equation (3.5))
$\mathbf{0}_n = \{\{1\}, \ldots, \{n\}\} \in \mathcal{P}_n$
$\#\pi$: number of blocks of a partition π
$\mathcal{S}_n$ set of semi-partitions of $[n]$ (p. 156), associated transformations (p. 156)
$\Delta = \{x = (x(1), x(2), \ldots) : x(1) \geq x(2) \geq \ldots 0, |x|_1 \leq 1\}$
$\Delta^N = \{x \in \Delta : |x|_1 = 1, Nx(i) \in \mathbb{N}_0 \text{ for all } i \in \mathbb{N}\}$
$\Delta_c = \{x \in \Delta : x(1) > c\}$

Genealogy in the Cannings model

$(x_k^N, k \in \mathbb{N})$: sequence in Δ^N that gives the family sizes (p. 151)
$(\pi_k^N, k \in \mathbb{N})$ sequence in $\mathcal{P}_N$ that gives the families (p. 151)
$A_j(k, i)$: label of the ancestor in generation j of the individual i in generation k (p. 151)
$\rho_k^N(i, j)$: genealogical distance (p. 152)
c_N: pairwise coalescence probability (equation (3.1))
b_N: probability that a randomly sampled individual is in a non-singleton family (p. 154)
$\chi_k^N = \psi_N(\rho_k^N)$: unlabeled genealogical tree (equation (3.4))
$\hat{\chi}_k^N = \hat{\psi}_N(\beta(\rho_k^N))$: unlabeled genealogical trees, decomposed at the external branches (p. 154)

(r_k^N, v_k^N), $\tilde{\chi}_k^N = \hat{\psi}_N(r_k^N, v_k^N)$: another decomposition of the genealogical trees (p. 156)

Tree-valued Fleming-Viot processes

$\mathcal{M}_1(\Delta)$, $\mathcal{M}_{\text{dust}}$, $\mathcal{M}_{\text{nd}}$: Set of probability measures on Δ, subsets of the measures with and without dust (p. 152)
$\Xi = \Xi_0 + \Xi\{0\}\delta_0$ (equation (3.3))
λ_π, $\lambda_{n,\sigma}$: reproduction rates (pp. 159, 160)

Acknowledgments

Partial support from the DFG Priority Programme 1590 "Probabilistic Structures in Evolution" is acknowledged.

Bibliography

[1] D. Aldous, *The continuum random tree. III*, Ann. Probab. **21** (1993), no. 1, 248–289. MR1207226

[2] S. Athreya, W. Löhr, and A. Winter, *The gap between Gromov-vague and Gromov–Hausdorff-vague topology*, Stochastic Process. Appl. **126** (2016), no. 9, 2527–2553. MR3522292

[3] T. Austin, *Exchangeable random measures*, Ann. Inst. Henri Poincaré Probab. Stat. **51** (2015), no. 3, 842–861. MR3365963

[4] J. Berestycki and N. Berestycki, *Kingman's coalescent and Brownian motion*, ALEA Lat. Am. J. Probab. Math. Stat. **6** (2009), 239–259. MR2534485

[5] J. Berestycki, N. Berestycki, and J. Schweinsberg, *Beta-coalescents and continuous stable random trees*, Ann. Probab. **35** (2007), no. 5, 1835–1887. MR2349577

[6] ———, *Small-time behavior of beta coalescents*, Ann. Inst. Henri Poincaré Probab. Stat. **44** (2008), no. 2, 214–238. MR2446321

[7] J. Bertoin, *Random fragmentation and coagulation processes*, Cambridge University Press, Cambridge, 2006. MR2253162

[8] J. Bertoin and J.-F. Le Gall, *Stochastic flows associated to coalescent processes*, Probab. Theory Related Fields **126** (2003), no. 2, 261–288. MR1990057

[9] ———, *Stochastic flows associated to coalescent processes. II. Stochastic differential equations*, Ann. Inst. H. Poincaré Probab. Statist. **41** (2005), no. 3, 307–333. MR2139022

[10] ———, *Stochastic flows associated to coalescent processes. III. Limit theorems*, Illinois J. Math. **50** (2006), no. 1-4, 147–181 (electronic). MR2247827

[11] M. Birkner, *Stochastic models from population biology*, 2005. Lecture notes for a course at TU Berlin. Available at `http://www.staff.uni-mainz.de/birkner/lehre_archiv/smpb-30.6.05.pdf`.

[12] M. Birkner, J. Blath, M. Capaldo, A. Etheridge, M. Möhle, J. Schweinsberg, and A. Wakolbinger, *Alpha-stable branching and beta-coalescents*, Electron. J. Probab. **10** (2005), no. 9. MR2120246

[13] 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. MR2485878

[14] D. Blount and M. A. Kouritzin, *On convergence determining and separating classes of functions*, Stochastic Process. Appl. **120** (2010), no. 10. MR2673979

[15] M. G. B. Blum and O. François, *Minimal clade size and external branch length under the neutral coalescent*, Adv. in Appl. Probab. **37** (2005), no. 3, 647–662. MR2156553

[16] R. M. Blumenthal, *Excursions of Markov processes*, Probability and its Applications, Birkhäuser Boston, 1992. MR1138461

[17] D. Burago, Y. Burago, and S. Ivanov, *A course in metric geometry*, Graduate Studies in Mathematics, vol. 33, American Mathematical Society, Providence, RI, 2001. MR1835418

[18] C. Cannings, *The latent roots of certain Markov chains arising in genetics: a new approach. I. Haploid models*, Advances in Appl. Probability **6** (1974), 260–290. MR0343949

[19] ———, *The latent roots of certain Markov chains arising in genetics: a new approach. II. Further haploid models*, Advances in Appl. Probability **7** (1975), 264–282. MR0371430

[20] 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. MR1654531

[21] R. V. Chacon, Y. Le Jan, E. Perkins, and S. J. Taylor, *Generalised arc length for Brownian motion and Lévy processes*, Z. Wahrsch. Verw. Gebiete **57** (1981), no. 2, 197–211. MR626815

[22] I. Dahmer and G. Kersting, *The total external length of the evolving Kingman coalescent*, Probab. Theory Related Fields (2016). `http://dx.doi.org/10.1007/s00440-016-0703-7`.

[23] I. Dahmer, R. Knobloch, and A. Wakolbinger, *The Kingman tree length process has infinite quadratic variation*, Electron. Commun. Probab. **19** (2014), no. 87. MR3298276

[24] A. Depperschmidt, A. Greven, and P. Pfaffelhuber, *Marked metric measure spaces*, Electron. Commun. Probab. **16** (2011), 174–188. MR2783338

[25] ———, *Tree-valued Fleming-Viot dynamics with mutation and selection*, Ann. Appl. Probab. **22** (2012), no. 6, 2560–2615. MR3024977

[26] ———, *Path-properties of the tree-valued Fleming-Viot process*, Electron. J. Probab. **18** (2013), no. 84. MR3109623

[27] P. Donnelly and T. G. Kurtz, *A countable representation of the Fleming-Viot measure-valued diffusion*, Ann. Probab. **24** (1996), no. 2, 698–742. MR1404525

[28] ———, *Particle representations for measure-valued population models*, Ann. Probab. **27** (1999), no. 1, 166–205. MR1681126

[29] X. Duhalde, *Uniform Hausdorff measure of the level sets of the Brownian tree*, ALEA Lat. Am. J. Probab. Math. Stat. **11** (2014), 885–916.

[30] T. Duquesne and J.-F. Le Gall, *Probabilistic and fractal aspects of Lévy trees*, Probab. Theory Related Fields **131** (2005), no. 4, 553–603. MR2147221

[31] ———, *The Hausdorff measure of stable trees*, ALEA Lat. Am. J. Probab. Math. Stat. **1** (2006), 393–415. MR2291942

[32] A. Etheridge, *An introduction to superprocesses*, University Lecture Series, vol. 20, American Mathematical Society, Providence, RI, 2000. MR1779100

[33] ———, *Some mathematical models from population genetics*, Lecture Notes in Mathematics, vol. 2012, Springer-Verlag, 2011. Lectures from the 39th Probability Summer School held in Saint-Flour, 2009. MR2759587

[34] A. Etheridge and T. G. Kurtz, *Genealogical constructions of population models* (2016). arXiv:1402.6724.

[35] A. Etheridge and P. March, *A note on superprocesses*, Probab. Theory Related Fields **89** (1991), no. 2, 141–147. MR1110534

[36] S. N. Ethier and T. G. Kurtz, *Markov processes*: *Characterization and convergence*, Wiley, New York, 1986. MR838085

[37] ———, *Fleming-Viot processes in population genetics*, SIAM J. Control Optim. **31** (1993), no. 2, 345–386. MR1205982

[38] S. N. Evans, *Kingman's coalescent as a random metric space*, Stochastic models (Ottawa, ON, 1998), 2000, pp. 105–114. MR1765005

[39] ———, *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. MR2351587

[40] 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. MR3601650

[41] S. N. Evans and A. Winter, *Subtree prune and regraft: a reversible real tree-valued Markov process*, Ann. Probab. **34** (2006), no. 3, 918–961. MR2243874

[42] N. Forman, C. Haulk, and J. Pitman, *A representation of exchangeable hierarchies by sampling from real trees* (2017). arXiv:1101.5619.

[43] 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. MR3069373

[44] P. C. Glöde, *Dynamics of genealogical trees for autocatalytic branching processes*, PhD thesis, Universität Erlangen-Nürnberg, 2012.

[45] 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. MR2520129

[46] ———, *Tree-valued resampling dynamics Martingale problems and applications*, Probab. Theory Related Fields **155** (2013), no. 3-4, 789–838. MR3034793

[47] A. Greven, L. Popovic, and A. Winter, *Genealogy of catalytic branching models*, Ann. Appl. Probab. **19** (2009), no. 3, 1232–1272. MR2537365

[48] M. Gromov, *Metric structures for Riemannian and non-Riemannian spaces*, Progress in Mathematics, vol. 152, Birkhäuser, Boston, MA, 1999. MR1699320

[49] S. Gufler, *Pathwise construction of tree-valued Fleming-Viot processes* (2016). arXiv:1404.3682.

[50] ———, *A representation for exchangeable coalescent trees and generalized tree-valued Fleming-Viot processes* (2016). arXiv:1608.08074.

[51] ———, *Invariance principles for tree-valued Cannings chains* (2017). arXiv:1608.08203.

[52] O. Hénard, *The fixation line in the Λ-coalescent*, Ann. Appl. Probab. **25** (2015), no. 5, 3007–3032. MR3375893

[53] P. Herriger and M. Möhle, *Conditions for exchangeable coalescents to come down from infinity*, ALEA Lat. Am. J. Probab. Math. Stat. **9** (2012), no. 2, 637–665. MR3069379

[54] B. Hughes, *Trees and ultrametric spaces: a categorical equivalence*, Adv. Math. **189** (2004), no. 1, 148–191. MR2093482

[55] O. Kallenberg, *Foundations of modern probability*, Second Edition, Springer, New York, 2002. MR1876169

[56] ———, *Probabilistic symmetries and invariance principles*, Springer, New York, 2005. MR2161313

[57] A. S. Kechris, *Classical descriptive set theory*, Graduate Texts in Mathematics, vol. 156, Springer-Verlag, New York, 1995. MR1321597

[58] G. Kersting, J. Schweinsberg, and A. Wakolbinger, *The evolving beta coalescent*, Electron. J. Probab. **19** (2014), no. 64. MR3238784

[59] J. F. C. Kingman, *The coalescent*, Stochastic Process. Appl. **13** (1982), no. 3, 235–248. MR671034

[60] ———, *Exchangeability and the evolution of large populations*, Exchangeability in probability and statistics (Rome, 1981), 1982, pp. 97–112. MR675968

[61] S. Kliem and W. Löhr, *Existence of mark functions in marked metric measure spaces*, Electron. J. Probab. **20** (2015), no. 73. MR3371432

[62] T. G. Kurtz, *Martingale problems for conditional distributions of Markov processes*, Electron. J. Probab. **3** (1998), no. 9, 29 pp. MR1637085

[63] T. G. Kurtz and G. Nappo, *The filtered martingale problem*, The Oxford handbook of nonlinear filtering, 2011, pp. 129–165. MR2884595

[64] C. Labbé, *From flows of Λ-Fleming-Viot processes to lookdown processes via flows of partitions*, Electron. J. Probab. **19** (2014), no. 55. MR3227064

[65] J.-F. Le Gall, *Marches aléatoires, mouvement brownien et processus de branchement*, Séminaire de Probabilités, XXIII, 1989, pp. 258–274. MR1022916

[66] ______, *Brownian excursions, trees and measure-valued branching processes*, Ann. Probab. **19** (1991), no. 4, 1399–1439. MR1127710

[67] ______, *Random trees and applications*, Probab. Surv. **2** (2005), 245–311. MR2203728

[68] ______, *Random real trees*, Ann. Fac. Sci. Toulouse Math. (6) **15** (2006), no. 1, 35–62. MR2225746

[69] 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. MR3037215

[70] G. Miermont, *Tessellations of random maps of arbitrary genus*, Ann. Sci. Éc. Norm. Supér. (4) **42** (2009), no. 5, 725–781. MR2571957

[71] 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. MR2684740

[72] M. Möhle and S. Sagitov, *A classification of coalescent processes for haploid exchangeable population models*, Ann. Probab. **29** (2001), no. 4, 1547–1562. MR1880231

[73] J. Neveu, *Arbres et processus de Galton-Watson*, Ann. Inst. H. Poincaré Probab. Statist. **22** (1986), no. 2, 199–207. MR850756

[74] J. Neveu and J. Pitman, *The branching process in a Brownian excursion*, Séminaire de Probabilités, XXIII, 1989, pp. 248–257. MR1022915

[75] ______, *Renewal property of the extrema and tree property of the excursion of a one-dimensional Brownian motion*, Séminaire de Probabilités, XXIII, 1989, pp. 239–247. MR1022914

[76] E. A. Perkins, *Conditional Dawson-Watanabe processes and Fleming-Viot processes*, Seminar on Stochastic Processes, 1991 (Los Angeles, CA, 1991), 1992, pp. 143–156. MR1172149

[77] 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. MR2307061

[78] 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. MR2851692

[79] J. Pitman, *Coalescents with multiple collisions*, Ann. Probab. **27** (1999), no. 4, 1870–1902. MR1742892

[80] ______, *Combinatorial stochastic processes*, Lecture Notes in Mathematics, vol. 1875, Springer-Verlag, Berlin, 2006. Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002. MR2245368

[81] D. Revuz and M. Yor, *Continuous martingales and Brownian motion*, Third edition, Springer-Verlag, Berlin, 1999. MR1725357

[82] L. C. G. Rogers and J. W. Pitman, *Markov functions*, Ann. Probab. **9** (1981), no. 4, 573–582. MR624684

[83] L. C. G. Rogers and D. Williams, *Diffusions, Markov processes, and martingales. Vol. 1*, Cambridge University Press, 2000. MR1796539

[84] S. Sagitov, *The general coalescent with asynchronous mergers of ancestral lines*, J. Appl. Probab. **36** (1999), no. 4, 1116–1125. MR1742154

[85] ———, *Convergence to the coalescent with simultaneous multiple mergers*, J. Appl. Probab. **40** (2003), no. 4, 839–854. MR2012671

[86] J. Schweinsberg, *Coalescents with simultaneous multiple collisions*, Electron. J. Probab. **5** (2000), no. 4. MR1781024

[87] T. Shiga, *A stochastic equation based on a Poisson system for a class of measure-valued diffusion processes*, J. Math. Kyoto Univ. **30** (1990), no. 2, 245–279. MR1068791

[88] L. Stournaras, *The tree-valued Wright-Fisher model*, 2012. Diploma thesis, Universität Freiburg.

[89] V. S. Varadarajan, *Groups of automorphisms of Borel spaces*, Trans. Amer. Math. Soc. **109** (1963), 191–220. MR0159923

[90] 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. MR3335017

[91] A. M. Vershik, *Random and universal metric spaces*, Dynamics and randomness II, Kluwer Acad. Publ., Dordrecht, 2004, pp. 199–228. MR2153330

Zusammenfassung

1 Hintergrund: Metrische Maßräume

In der vorliegenden Arbeit werden Genealogien auf eine Weise beschrieben, in der die Beschriftung der Individuen entfernt ist. Eine Beschriftung der Individuen ist nämlich oft nicht relevant: Werden beispielsweise unabhängig und identisch verteilte Stichproben aus einer Population gezogen, so gibt der unbeschriftete genealogische Baum dieser Stichproben nicht weniger Auskunft über die Genealogie der gesamten Population als derjenige, in denen die Stichproben entsprechend der Reihenfolge beschriftet sind, in der sie gezogen wurden. Zudem kann die evolvierende Genealogie einer Population, die durch das Moran-Modell beschreiben wird, als Prozess von ungeordneten Bäumen aus dem lookdown-Modell von Donnelly und Kurtz [28] abgelesen werden, welches auch für eine unendliche Populationsgröße definiert ist.

Während sich aus dem genealogischen Baum einer endlichen Population die Beschriftung der Individuen leicht entfernen lässt, indem die Isomorphieklasse dieses Baumes unter Permutationen der Beschriftungen betrachtet wird, ist bei unendlichen Populationen zu beachten, dass durch die Isomorphieklasse unter unendlichen Permutationen die asymptotischen Häufigkeiten von Individuen mit interessierenden Eigenschaften nicht bestimmt sind, wie beispielsweise die asymptotischen Familiengrößen der Individuen, die einen gemeinsamen Ahnen zu einer bestimmten Zeit in der Vergangenheit haben. Um lediglich diese Information zusätzlich zu den genealogischen Abständen zwischen den Individuen zu behalten, eignet sich die Beschreibung von genealogischen Bäumen durch Isomorphieklassen von metrischen Maßräumen.

Von besonderem Interesse sind hier semi-ultrametrische Räume. Eine Semi-Ultrametrik ist eine Halbmetrik ρ, welche die verschärfte Dreiecksungleichung

$$\max\{\rho(x,y),\rho(y,z)\} \geq \rho(x,z)$$

erfüllt. Semi-Ultrametriken auf $\mathbb{N}$ stehen in eineindeutiger Beziehung zu rechtsstetigen Pfaden mit Werten im Raum der Partitionen von $\mathbb{N}$, in denen die Zustände nur gröber werden können. Solche Pfade sind Realisierungen von Koaleszenten, welche wichtige genealogische Modelle in der Populationsgenetik sind. Sie beschreiben eine Genealogie einer mit den Elementen von $\mathbb{N}$ beschrifteten Stichprobe, in der $i, j \in \mathbb{N}$ einen gemeinsamen Vorfahren zu der Zeit in der Vergangenheit haben, zu der i und j in einem gleichen Block sind. Jede Ultrametrik spannt auch einen reellen Baum auf und lässt sich als Unterraum der Blätter in diesen einbetten. Ein reeller Baum ist ein metrischer Raum (T, d), welcher in dem Sinne baumartig ist, dass (i) für alle $x, y \in T$ eine Isometrie von dem reellen Intervall $[0, d(x, y)]$ nach T existiert, und (ii) kein Unterraum von T homöomorph zum Einheitskreis ist. Ein kombinatorischer Baum mit gegebenen Kantenlängen lässt sich durch Aneinanderkleben von reellen Intervallen mit diesen Kantenlängen als reeller Baum auffassen.

Den Familien in einem Koaleszenten, die einen gemeinsamen Vorfahren zu einer bestimmten Zeit in der Vergangenheit haben, entsprechen die Kugeln des zugehörigen ultrametrischen Raumes. Evans [38] untersucht die Vervollständigung des dem Kingman-Koaleszenten entsprechenden ultrametrischen Raumes, welchen er mit einem Wahrscheinlichkeitsmaß versieht, das jede Kugel mit der entsprechenden asymptotischen Familiengröße gewichtet.

Ein Tripel (X, r, μ), bestehend aus einem vollständigen und separablen metrischen Raum (X, r) und einem Wahrscheinlichkeitsmaß μ auf der Borelschen Sigma-Algebra wird metrischer Maßraum genannt. Vershik [91] gibt eine notwendige und hinreichende Bedingung an eine zufällige Halbmetrik ρ auf $\mathbb{N}$ an, unter der sich mit jeder typischen Realisierung von ρ ein metrischer Maßraum (X, r, μ) assoziieren lässt, welcher die Eigenschaft besitzt, dass die Abstände $(r(x_i, x_j))_{i,j\in\mathbb{N}}$ zwischen den Gliedern einer u. i. v. Folge $(x_i)_{i\in\mathbb{N}}$ aus (X, r, μ) wieder wie ρ verteilt sind. Die Verteilung der Matrix $(r(x_i, x_j))_{i,j\in\mathbb{N}}$ wird Distanzmatrixverteilung von (X, r, μ) genannt. Der Gromovsche Rekonstruktionssatz [48, 91] besagt, dass metrische Maßräume (X, r, μ), (X', r', μ') genau dann die gleiche Distanzmatrixverteilung haben, wenn zwischen den Trägern von μ und μ' eine maßerhaltende Isometrie existiert. In diesem Fall nennen wir die metrischen Maßräume isomorph. Greven, Pfaffelhuber und Winter [45] zeigen, dass der Raum der Isomorphieklassen von metrischen Maßräumen polnisch ist, wenn er mit der Gromov-schwachen Topologie versehen wird, in der metrische Maßräume genau dann konvergieren, wenn ihre Distanzmatrixverteilungen schwach konvergieren.

Greven, Pfaffelhuber und Winter [45] behandeln auch die Isomorphieklassen von ultrametrischen Maßräumen, welche Λ-Koaleszenten entsprechen, und geben so eine beschriftungsfreie Beschreibung dieser Koaleszenten an, aus der sich die asymptotischen Familiengrößen ablesen lassen. In [45] ist auch gezeigt, dass dieser Ansatz genau dann funktioniert, wenn die Λ-Koaleszenten staubfrei sind. Eine Interpretation von Staubfreiheit ist, dass der ultrametrische Raum keine isolierten Punkte aufweist.

2 Ziehen von austauschbaren Ultrametriken aus zufälligen markierten metrischen Maßräumen

In der vorliegenden Arbeit definieren wir die Distanzmatrixverteilung eines markierten metrischen Maßraumes und konstruieren für jede austauschbare Semi-Ultrametrik ρ auf $\mathbb{N}$ einen zufälligen markierten metrischen Maßraum, dessen Distanzmatrixverteilung nach Mittelung über die Zufälligkeit des markierten metrischen Maßraumes wieder die Verteilung von ρ ist. Dies erlaubt es, auch Koaleszenten mit Staub durch Isomorphieklassen von markierten metrischen Maßräumen zu beschreiben.

Markierte metrische Maßräume wurden von Depperschmidt, Greven und Pfaffelhuber

[24] eingeführt, um zusätzlich zu einer Genealogie mithilfe von Marken auch Alleltypen zu beschreiben [25, 26]. Ein markierter metrischer Maßraum ist ein Tripel (X, r, m), bestehend aus einem vollständigen und separablen metrischen Raum (X, r) und einem Wahrscheinlichkeitsmaß auf der Borelschen Sigma-Algebra auf dem Produktraum $X \times \mathbb{R}_+$, dessen zweiter Faktor Markenraum heißt und in der vorliegenden Arbeit stets der Raum $\mathbb{R}_+$ sein wird. Depperschmidt et al. [24] definieren die markierte Distanzmatrixverteilung eines markierten metrischen Maßraumes (X, r, m) als die Verteilung von $((r(x_i, x_j))_{i,j\in\mathbb{N}}, (v_i)_{i\in\mathbb{N}})$, wobei $(x_i, v_i)_{i\in\mathbb{N}}$ eine gemäß m u. i. v. Folge in $X \times \mathbb{R}_+$ ist. Wir nennen markierte metrische Maßräume mit der gleichen markierten Distanzmatrixverteilung isomorph. Wie in [25] gezeigt, ist der Raum $\hat{\mathbb{M}}$ der Isomorphieklassen von markierten metrischen Maßräumen polnisch, wenn er mit der markierten Gromov-schwachen Topologie versehen ist, in der Konvergenz durch die schwache Konvergenz der markierten Distanzmatrixverteilungen definiert ist.

Die Distanzmatrixverteilung eines markierten metrischen Maßraumes χ definieren wir in der vorliegenden Arbeit als die Verteilung der durch

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

gegebenen Matrix $(\rho(i, j))_{i,j\in\mathbb{N}}$, wobei $((r(i, j))_{i,j\in\mathbb{N}}, (v(i))_{i\in\mathbb{N}})$ die markierte Distanzmatrixverteilung von χ besitze.

Ist umgekehrt eine Semi-Ultrametrik ρ auf $\mathbb{N}$ gegeben, können wir diese in eine markierte Distanzmatrix $\beta(\rho) := (r, v) \in \mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}}$ zerlegen, so dass in dem von ρ aufgespannten reellen Baum

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

der Länge des externen Astes entspricht, der in Blatt i endet, und

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

den Abstand des Anfangsknotens dieses externen Astes zu dem Anfangsknoten des in Blatt j endenden externen Astes angibt.

Wir identifizieren die Punkte mit r-Abstand Null in dem halbmetrischen Raum $(\mathbb{N}, r)$ und bilden die metrische Vervollständigung, welche wir hier mit (X, r) bezeichnen wollen und uns dabei mit $i \in \mathbb{N}$ auch auf das entsprechende Element von X beziehen werden. Wir nehmen Austauschbarkeit von ρ an, d. h. Invarianz der Verteilung von ρ unter endlichen Permutationen p, deren Wirkung durch $p(\rho) = (\rho(p(i), p(j)))_{i,j\in\mathbb{N}}$ gegeben sei. Für jede typische Realisierung von ρ zeigen wir dann, sowohl durch eine approximative Konstruktion als auch mit dem Satz von de Finetti, die schwache Konvergenz der Wahrscheinlichkeitsmaße $\frac{1}{n}\sum_{i=1}^{n} \delta_{(i,v(i))}$ auf dem Raum $X \times \mathbb{R}_+$ zu einem Samplingmaß m. Wir bezeichnen mit $\hat{\psi}$ die (messbare) Funktion, welche die markierte Distanzmatrix (r, v) auf die Isomorphieklasse des markierten metrischen Maßraumes (X, r, m) abbildet.

Wir erhalten aus der obigen Konstruktion f. s. einen markierten metrischen Maßraum (X, r, m), dessen Isomorphieklasse $\chi = \hat{\psi}(\beta(\rho))$ eine $\hat{\mathbb{M}}$-wertige Zufallsvariable mit der Eigenschaft ist, dass $\mathbb{E}[\nu]$ wieder die Verteilung von ρ ist, wobei das zufällige Maß ν die Distanzmatrixverteilung von χ bezeichne.

Wir halten außerdem fest, dass (analog zu Vershik [91, Lemma 7]) die Distanzmatrixverteilung eines markierten metrischen Maßraumes invariant und ergodisch bezüglich endlicher Permutationen ist. Die Verteilung von ρ wird durch die obige Darstellung daher in ihre ergodische Komponenten zerlegt. Aus jeder typischen Realisierung von ρ lässt sich die ergodische Komponente als die Distanzmatrixverteilung von $\hat{\psi} \circ \beta(\rho)$ ablesen.

3 Baumwertige Fleming-Viot-Prozesse

Wie maßwertige Fleming-Viot-Prozesse [36, 37] Modelle für die Evolution der Typenverteilung in evolvierenden Populationen sind, so beschreiben baumwertige Fleming-Viot-Prozesse die Evolution der Verteilung der genealogischen Abstände von zufällig gezogenen Individuen. Greven, Pfaffelhuber und Winter [46] konstruieren einen baumwertigen Fleming-Viot-Prozess im Fall von binären Reproduktionsereignissen als Lösung eines wohlgestellten Martingalproblems.

In der vorliegenden Arbeit verallgemeinern wir baumwertige Fleming-Viot-Prozesse auf den Fall mit simultanen multiplen Reproduktionsereignissen, insbesondere schließen wir den Fall mit Staub ein. Dazu lesen wir aus einem geeignet im Sinne von Donnelly und Kurtz [28] definierten lookdown-Modell einen Prozess $(\rho_t, t \in \mathbb{R}_+)$ mit Werten im Raum der Semi-Ultrametriken auf $\mathbb{N}$ ab, welcher die evolvierenden genealogischen Abstände beschreibt. Den Bildprozess $(\hat{\psi}(\beta(\rho_t)), t \in \mathbb{R}_+)$ unter der Abbildung $\hat{\psi} \circ \beta$ aus Abschnitt 2 bezeichnen wir als einen Fleming-Viot-Prozess mit Werten im Raum der Isomorphieklassen von markierten metrischen Maßräume, die sich als Bäume interpretieren lassen. Auch den Prozess der zugehörigen Distanzmatrixverteilungen, also der ergodischen Komponenten, bezeichnen wir als baumwertigen Fleming-Viot-Prozess. Im staubfreien Fall lässt sich der baumwertige Fleming-Viot-Prozess zudem als ein Prozess mit Werten im Raum der Isomorphieklassen von metrischen Maßräumen definieren. Mithilfe des Kriteriums von Rogers und Pitman [82] zeigen wir die Markoveigenschaft der baumwertigen Fleming-Viot-Prozesse und geben wohlgestellte Martingalprobleme für diese Prozesse an.

Bertoin und Le Gall [8] beschreiben evolvierende Genealogien durch einen Fluss von Brücken. Sie konstruieren daraus maßwertige Fleming-Viot-Prozesse und zeigen, dass die Genealogie zu einer festen Zeit durch einen Ξ-Koaleszenten von Schweinsberg [86] gegeben ist. In der vorliegenden Arbeit lesen wir aus dem Fluss von Brücken einen Prozess mit Werten im Raum der Verteilungen von Semi-Ultrametriken ab und charakterisieren diesen als einen Fleming-Viot-Prozess mit Werten im Raum der Distanzmatrixverteilungen.

Die in den Abschnitten 1 – 3 zusammengefassten Inhalte finden sich in Kapitel 2 der Dissertation, welches der zur Veröffentlichung eingereichten Arbeit [50] entspricht.

4 Pfadweise Konstruktion aus dem lookdown-Modell

Bei der in Abschnitt 3 beschriebenen Vorgehensweise zur Konstruktion von baumwertigen Fleming-Viot-Prozessen mittels der Darstellung aus Abschnitt 2 kann das fast sichere Ereignis, auf dem das Sampling-Maß existiert, vom Zeitpunkt abhängen. Wir konstruieren Sampling-Maße jedoch auch zu allen Zeitpunkten gleichzeitig auf einem fast sicheren

Ereignis und erhalten daraus eine pfadweise Konstruktion von baumwertigen Fleming-Viot-Prozessen.

Zur Beschreibung der pfadweisen Konstruktion skizzieren wir nun das lookdown-Modell. Die Zeitachse ist $\mathbb{R}_+$ und es gibt abzählbar viele levels, welche mit den natürlichen Zahlen indiziert sind. Zu jeder Zeit befindet sich auf jedem level genau ein Teilchen. Mit fortschreitender Zeit nehmen die Teilchen an Reproduktionsereignissen teil, in denen die Nachkommen auf höhere levels gesetzt werden als ihr Erzeuger, und alle Teilchen ihr level nur erhöhen können. Wir bezeichnen das Teilchen auf level i zu einem Zeitpunkt t als das Individuum $(s,i) \in \mathbb{R}_+ \times \mathbb{N}$ und definieren den genealogischen Abstand $\rho((s,i),(t,j))$ zwischen zwei Individuen mithilfe der Reproduktionsereignisse und der gegebenen genealogischen Abstände zwischen den Individuen zur Zeit Null. Wir identifizieren Individuen mit genealogischem Abstand Null und bilden die metrische Vervollständigung, welche wir als den lookdown-Raum (Z,ρ) bezeichnen, wobei wir die Individuen $(t,i) \in \mathbb{R}_+ \times \mathbb{N}$ auch als Elemente des lookdown-Raumes auffassen.

Unter der Annahme, dass die Genealogie zu festen Zeitpunkten durch einen staubfreien Ξ-Koaleszenten gegeben ist, zeigen wir f. s., dass die uniformen Wahrscheinlichkeitsmaße

$$\mu_t^n := \frac{1}{n}\sum_{i=1}^{n} \delta_{(t,i)}$$

auf den Individuen auf den ersten n levels uniform in kompakten Zeitintervallen bezüglich der Prohorov-Metrik d_{P}^Z auf dem lookdown-Raum konvergieren:

$$\lim_{n\to\infty} \sup_{t\in[0,T]} d_{\mathrm{P}}^Z(\mu_t^n, \mu_t) = 0 \quad \text{f. s. für alle } T \in \mathbb{R}_+. \tag{4.1}$$

Durch diesen Ausdruck sind auch die Sampling-Maße μ_t definiert. Aus der Konstruktion können wir nun den baumwertigen Fleming-Viot-Prozess f. s. pfadweise als den Prozess $(\chi_t, t \in \mathbb{R}_+)$ ablesen, wobei χ_t die Isomorphieklasse des metrischen Maßraumes (Z,ρ,μ_t) bezeichne.

Aus der Konstruktion folgern wir zudem, dass der Prozess $(\chi_t, t \in \mathbb{R}_+)$ f. s. càdlàg Pfade in der Gromov-schwachen Topologie besitzt und nur zu den Zeiten springen kann, zu denen große Reproduktionsereignisse (mit unendlich vielen Neugeborenen) auftreten. Außerdem zeigen wir, dass zu den Zeiten t von großen Reproduktionsereignissen μ_t Atome besitzt, nicht aber der linksseitige Grenzwert μ_{t-}. Daraus folgt, dass der Prozess $(\chi_t, t \in \mathbb{R}_+)$ genau zu den Zeiten von großen Reproduktionsereignissen springt.

Im Falle des Herunterkommens von unendlich, d. h. wenn für alle $s < t$ die Ahnenmenge zur Zeit s der Individuen zur Zeit t endlich ist, ist der Abschluss X_t des Unterraums $\{t\} \times \mathbb{N} \subset Z$ der Individuen zur Zeit t kompakt. In diesem Fall lässt sich auch der Prozess $(\mathcal{X}_t, t \in \mathbb{R}_+)$ betrachten, wobei $\mathcal{X}_t$ die maßerhaltende Isometrieklasse von (X_t, ρ, μ_t) sei, und der Zustandsraum mit der stärkeren Gromov-Hausdorff-Prohorov-Topologie versehen sei. Diese Topologie hebt die metrische Struktur von X_t auch unabhängig von dem Maß μ_t hervor, was f. s. zu zusätzlichen Sprüngen zu den Zeitpunkten führt, zu denen alle Nachfahren eines Individuums aussterben. Wir zeigen zudem f. s., dass für alle t der abgeschlossene Träger des Maßes μ_t den gesamten Raum X_t umfasst. Wir nennen den Prozess $(\mathcal{X}_t, t \in \mathbb{R}_+)$ einen baumwertigen evolvierenden Ξ-Koaleszenten.

Wir konstruieren die baumwertigen Fleming-Viot-Prozesse auch im Fall mit Staub pfadweise. Dazu definieren wir für jedes Individuum (t,i) einen Elter $z(t,i)$ als das jüngste reproduzierende Individuum auf der Ahnenlinie, falls ein solches existiert. Andernfalls definieren wir $z(t,i)$ als den Elter von dessen Vorfahren zur Zeit Null, wofür wir den lookdown-Raum um die Eltern der Individuen zur Zeit Null erweitern und von dem erweiterten lookdown-Raum $(\hat{Z},\rho)$ sprechen. Den genealogischen Abstand zwischen (t,i) und $z(t,i)$ bezeichnen wir mit $v_t(i)$. Wir versehen den Raum $\hat{Z}\times\mathbb{R}_+$ mit der Maximum-Produktmetrik und betrachten darauf die Wahrscheinlichkeitsmaße

$$m_t^n = \frac{1}{n}\sum_{i=1}^n \delta_{(z(t,i),v_t(i))}$$

welche für jedes der Individuen auf den ersten n levels den Elter und den genealogischen Abstand zu diesem mit der Masse $1/n$ gewichten. Unter der Annahme, dass die Genealogie zu festen Zeitpunkten durch einen Ξ-Koaleszenten gegeben ist, zeigen wir f. s. die uniforme Konvergenz in kompakten Zeitintervallen bezüglich der Prohorov-Metrik auf dem Produktraum $\hat{Z}\times\mathbb{R}_+$,

$$\lim_{n\to\infty}\sup_{t\in[0,T]} d_{\mathrm{P}}^{\hat{Z}\times\mathbb{R}_+}(m_t^n, m_t) = 0 \quad \text{f. s. für alle } T\in\mathbb{R}_+, \tag{4.2}$$

welche auch die Sampling-Maße m_t definiert. Den baumwertigen Ξ-Fleming-Viot-Prozess lesen wir dann f. s. pfadweise als den Prozess $(\hat{\chi}_t, t\in\mathbb{R}_+)$ ab, wobei $\hat{\chi}_t$ die Isomorphieklasse des markierten metrischen Maßraumes $(\hat{Z},\rho,m_t)$ bezeichne.

Aus dieser Konstruktion erhalten wir für einen geeigneten Anfangszustand, dass auch die Pfade von $(\hat{\chi}_t, t\in\mathbb{R}_+)$ càdlàg sind mit Sprüngen zu den Zeiten von großen Reproduktionereignissen. Im Fall mit Staub besitzt zu den Zeiten von großen Reproduktionsereignissen die $\mathbb{R}_+$-Komponente von m_t ein Atom in 0, nicht aber der linksseitige Grenzwert m_{t-}. Wir zeigen ferner mittels einer expliziten Darstellung der Maße m_t, dass diese rein atomar sind, wenn das für m_0 gilt. Wir betrachten auch den Fleming-Viot-Prozess mit Werten im Raum der Distanzmatrixverteilungen und zeigen, dass dieser, unabhängig vom Anfangszustand, immer f. s. càdlàg Pfade besitzt.

Für den Beweis der Konvergenzen (4.1) und (4.2) approximieren wir die Maße μ_t bzw. m_t durch atomare Maße, die wir mithilfe von Familien von Partitionen konstruieren. Im Fall ohne Staub verwenden wir den Fluss der Partitionen von Foucart [43] und Labbé [64]. Für den Fall mit Staub definieren wir eine Familie von Partitionen, in denen die Eltern von gleichzeitig lebenden Individuen, deren levels sich im gleichen Block befinden, einen geringen Abstand zueinander haben, wenn es sich nicht um einen Ausnahmeblock handelt. Die Untersuchung der asymptotischen Frequenzen dieser beiden Familien von Partitionen beruht auf Techniken, mittels derer Donnelly und Kurtz [28] maßwertige Fleming-Viot-Prozesse pfadweise konstruieren.

Die in diesem Abschnitt zusammengefassten Resultate finden sich in Kapitel 3 der Dissertation, welches der zur Veröffentlichung eingereichten Arbeit [49] entspricht.

5 Konstruktion aus einer Brownschen Exkursion

In dem Kingmans Koaleszenten entsprechenden Fall, dass in jedem Reproduktionsereignis genau ein Teilchen genau ein weiteres Teilchen erzeugt, konstruieren wir den lookdown-Raum in zweiseitiger Zeit mitsamt den Sampling-Maßen, und damit auch einen stationären baumwertigen Fleming-Viot-Prozess, auch aus einer Brownschen Exkursion $(B(s), s \in \mathbb{R}_+)$, welche wir auf Höhe größer als 1 bedingen. Die Brownsche Exkursion kodiert einen zufälligen reellen Baum (T, d), welcher bis auf die Normierung Aldous' [1] continuum random tree ist und durch die Halbmetrik

$$d(x, y) = B(x) + B(y) - 2 \min_{[x,y]} B$$

auf $\mathbb{R}_+$ und Identifikation der Punkte mit Abstand Null gegeben ist. Der Ausdruck $B(x)$ ist auch für $x \in T$ durch Repräsentanten von x wohldefiniert und kann als Höhe von x in dem Baum (T, d) interpretiert werden. Ferner bezeichnen wir mit g_x und d_x das kleinste bzw. größte Element in der Äquivalenzklasse $x \in T$. Ein Punkt $x \in T$ wird Vorfahre von $y \in T$ genannt $(x \prec y)$, falls $[g_y, d_y] \subset [g_x, d_x]$ gilt. Die Metrik d auf T erfüllt auch

$$d(x, y) = B(x) + B(y) - 2 \max\{B(v) : v \in T, v \prec x, v \prec y\}, \quad x, y \in T.$$

Wir entfernen nun von (T, d) die Wurzel und das höchste Blatt, um zu dem Unterbaum

$$T' = \{x \in T : B(x) \in (0, H)\}$$

zu gelangen, wobei $H := \sup B$ der Höhe des Baumes entspricht. Wir bezeichnen mit $\zeta(u)$ die Lokalzeit von B in Höhe u. Durch eine Zeittransformation

$$\tau(u) = \int_1^u \frac{4}{\zeta(v)} dv, \quad u \in (0, H),$$

mit der Shiga [87] auch den maßwertigen Fleming-Viot-Prozess aus dem maßwertigen Dawson-Watanabe-Prozess erhält, strecken wir die Metrik d zu der durch

$$d'(x, y) = \tau(B(x)) + \tau(B(y)) - 2 \max\{\tau(B(v)) : v \in T', v \prec x, v \prec y\}, \quad x, y \in T'.$$

gegebenen Metrik.

Wir bezeichnen mit m_u das normierte Lokalzeitmaß auf T in Höhe $u \in (0, H)$. Wir können nun einen stationären baumwertigen Fleming-Viot-Prozess $(\chi_t, t \in \mathbb{R})$ aus der Brownschen Exkursion ablesen, indem wir χ_t als die Isomorphieklasse des metrischen Maßraumes $(T', d', m_{U(t)})$ definieren, wobei $U(t) = \inf\{u \in (0, H) : \tau(u) > t\}$ die inverse Zeittransformation zu τ bezeichne.

Um $(\chi_t, t \in \mathbb{R})$ als einen baumwertigen Fleming-Viot-Prozess zu identifizieren, konstruieren wir zu $(T', d', m_{U(t)})$ maßerhaltend isometrische metrische Maßräume (Z, ρ, μ_t) auf die gleiche Weise aus Poissonprozessen wie einen zweiseitigen lookdown-Raum mit den Sampling-Maßen. Die Punkte dieser Poissonprozesse entsprechen lookdown-Ereignissen, welche wir (Berestycki et al. [4, 5] folgend) aus der Brownschen Exkursion ablesen, in der wir Teilexkursionen nach ihrer Höhe anordnen. Um diese Punktprozesse als Poissonprozesse zu identifizieren, geben wir eine lookdown-Darstellung für den nach Neveu und Pitman [74, 75] sowie Le Gall [65] in eine reflektierte Brownsche Bewegung einbeschriebenen binären Verzweigungswald an. Bei ihrer Konstruktion von Kingmans Koaleszenten aus einer Brownschen Exkursion beschreiten Berestycki und Berestycki [4] hier einen alternativen Weg über direkte Berechnungen mittels Exkursionstheorie. Die Konvergenz der Sampling-Maße auf (Z, ρ) folgern wir aus der uniformen downcrossing-Darstellung der Lokalzeit von Chacon et al. [21].

Das in diesem Abschnitt beschriebene Resultat findet sich in Kapitel 4 der Dissertation.

6 Invarianzprinzipien für baumwertige Cannings-Ketten

Koaleszenten mit simultanen multiplen Verschmelzungsereignissen sind robuste Grenzprozesse für partitionswertige Prozesse, welche die Genealogie zu festen Zeitpunkten in Cannings-Modellen beschreiben (Möhle und Sagitov [72]). Greven, Pfaffelhuber, und Winter [46, Remark 2.21] vermuteten eine entsprechende Universalität auch für den baumwertigen Fleming-Viot-Prozess. Im Folgenden wollen wir mit (*) die Bedingungen von Möhle und Sagitov [72] und Sagitov [85] bezeichnen, unter denen die Genealogie in Cannings-Modellen gegen einen Ξ-Koaleszenten konvergiert.

In der vorliegenden Arbeit betrachten wir die genealogischen Abstände $\rho_k^N(i,j)$ der Individuen i, j in Generation k in einem Cannings-Modell mit (endlicher) Populationsgröße N, welche wir mit der Paarverschmelzungswahrscheinlichkeit $c_N > 0$ reskalieren. Wir beschreiben die evolvierenden genealogischen Bäume durch eine Markovkette $(\chi_k, k \in \mathbb{N}_0)$, wobei χ_k^N die Isomorphieklasse des metrischen Maßraumes $(\{1, \dots, N\}, \rho_k^N, N^{-1} \sum_{i=1}^N \delta_i)$ bezeichne. Unter der Voraussetzung (*) und der Annahme, dass die Anfangszustände konvergieren, zeigen wir im staubfreien Fall die Konvergenz der Prozesse $(\chi^N_{\lfloor c_N^{-1} t \rfloor}, t \in \mathbb{R}_+)$ gegen einen baumwertigen Ξ-Fleming-Viot-Prozess in der Skorohod-Topologie über der Gromov-schwachen Topologie.

Insbesondere, um den Fall mit Staub einzuschließen, betrachten wir auch den Prozess $(\xi_k^N, k \in \mathbb{N}_0)$, wobei ξ_k^N die Distanzmatrixverteilung von χ_k^N bezeichne. Wir zeigen, dass die Prozesse $(\xi^N_{\lfloor c_N^{-1} t \rfloor}, t \in \mathbb{R}_+)$ im allgemeinen Fall unter der Voraussetzung (*) und der Annahme über die Anfangszustände gegen einen Ξ-Fleming-Viot-Prozess mit Werten im Raum der Distanzmatrixverteilungen in der Skorohod-Topologie über der schwachen Topologie konvergieren.

Wir betrachten zudem Markovketten $(\hat{\chi}_k^N, k \in \mathbb{N}_0)$ mit Werten im Raum der Isomorphieklassen von markierten metrischen Maßräumen, in denen $\hat{\chi}_k^N$ der Zerlegung des durch χ_k^N gegebenen Baumes an den Anfangsknoten der externen Äste entspricht: Sei $\hat{v}_k^N(i)$ die Länge des in Individuum i endenden externen Astes in dem mit c_N reskalierten genealogischen Baum von Generation k, und $\hat{r}_k^N(i,j)$ der Abstand der Anfangsknoten der in den Individuen i und j endenden externen Äste. Dann können wir $\hat{\chi}_k^N$ als die Isomorphieklasse des markierten metrischen Maßraumes $(\{1,\dots,N\},\hat{r}^N_k,N^{-1}\sum_{i=1}^N \delta_{(i,\hat{v}^N_k(i))})$ definieren. Unter der Voraussetzung (*) und einer geeigneten Annahme über den Anfangszustand konvergieren im staubfreien Fall auch die Prozesse $(\hat{\chi}^N_{\lfloor c_N^{-1} t\rfloor}, t\in\mathbb{R}_+)$ gegen einen baumwertigen Fleming-Viot-Prozess in der Skorohod-Metrik über der markierten Gromov-schwachen Topologie. Im Fall mit Staub zeigen wir die entsprechende Konvergenz unter einer Zusatzannahme an die Wahrscheinlichkeit b_N, dass ein uniform aus einer Generation im Cannings-Modell gezogenes Individuum zu einer Familie der Größe mindestens 2 gehört: Wir nehmen an, dass b_N/c_N gegen den Parameter der Exponentialverteilung konvergiert, welche als die Grenzverteilung der Länge eines uniform aus einem Ξ-Koaleszenten gezogenen externen Astes für unendliche Populationsgröße bekannt ist. Wir geben auch ein Beispiel an, in dem alle Voraussetzungen bis auf diese Zusatzannahme erfüllt sind und zeigen, dass die entsprechenden Prozesse dann nicht konvergieren. In dem Beispiel konvergiert zwar die Genealogie gegen einen Ξ-Koaleszenten mit Staub, also mit positiven externen Astlängen, die aus den Cannings-Modellen abgelesenen externen Astlängen konvergieren aber gegen Null. Der Grund hierfür liegt in Reproduktionsereignissen, die mit hoher Rate auftreten, aber jeweils einen so geringen Anteil der Population umfassen, dass sie in der Genealogie nach dem Grenzübergang nicht sichtbar sind.

Der Beweis der erwähnten Invarianzprinzipien beruht auf der Konvergenz der Übergangskerne der Markovketten zu den Generatoren von baumwertigen Fleming-Viot-Prozessen auf einem Kern im Sinne von Ethier und Kurtz [36, Kapitel 1.3]. Wir verwenden das Resultat aus Kapitel 2, dass die Definitionsbereiche der Martingalprobleme für baumwertige Fleming-Viot-Prozesse Kerne sind. Dies ermöglicht es, den Konvergenzsatz [36, Corollary 4.8.9] anzuwenden. Die relative Kompaktheit der approximierenden Prozesse muss für die Anwendung dieses Konvergenzsatzes nicht überprüft werden. Das Invarianzprinzip für die Markovketten $(\hat{\chi}^N_k, k\in\mathbb{N}_0)$ zeigen wir indirekt über die Zerlegung der genealogischen Bäume an den jüngsten Ahnen mit mindestens zwei Kindern. Mit dem Satz von Slutzky vergleichen wir dann die approximieren Prozesse. Im Fall mit Staub vergleichen wir hier nur die endlichdimensionalen Verteilungen und zeigen zusätzlich die relative Kompaktheit der Folge der approximierenden Prozesse $(\hat{\chi}^N_{\lfloor c_N^{-1} t\rfloor}, t\in\mathbb{R}_+)$.

Die in diesem Abschnitt zusammengefassten Resultate finden sich in Kapitel 5 der Dissertation, welches der zur Veröffentlichung eingereichten Arbeit [51] entspricht.